Circulas from Fourier series

Shogo Kato,
Institute of Statistical Mathematics, Tokyo, Japan
Arthur Pewsey[1]
University of Extremadura, Cáceres, Spain
and M.C. Jones
The Open University, Milton Keynes, UK

Summary. This article considers a general Fourier series based approach to obtaining the bivariate circular analogues of copulas recently coined 'circulas'. In contrast to standard copulas, circulas are periodic and have marginals that are circular uniform. As examples of the general construction we consider a number of classes of circulas arising from different patterns of non-zero Fourier coefficients. The shape and sparsity of such arrangements are found to play a key role in determining the properties of the resultant models. The special cases of the circulas we consider all have simple closed-form expressions for their densities which involve no computationally demanding normalizing constants. Methods for simulation, model identification, parameter estimation and testing for independence and goodness-of-fit are developed and applied in illustrative analyses of toroidal data sets.

Keywords: bivariate circular distributions; copulas; dependence; torus

1 Introduction

Circular data, corresponding to directions or the time of the day, month or year, arise in scientific disciplines as wide-ranging as biology, meteorology, astronomy, geology and the political and behavioural sciences (Pewsey *et al.*, 2013). The description of such data as 'circular' reflects the fact that their natural support is the unit circle, $\mathbb{S}^1$. In many contexts, the relationship between two circular random variables will be of interest. The natural support for bivariate circular vectors is the unit torus $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$, data on them sometimes being referred to as 'toroidal'.

This paper considers a general construction, based on Fourier series, for the circular analogues of bivariate copulas recently termed 'circulas' by Jones *et al.* (2015). An analogous result to Sklar's theorem (Sklar, 1959) states that the distribution function F of a bivariate vector (Θ_1, Θ_2) of continuous circular random variables distributed on $\mathbb{T}^2$ can be related to the marginal distribution functions of Θ_1 and Θ_2, F_1 and F_2, through the circula C as

$$F(\theta_1, \theta_2) = C(2\pi F_1(\theta_1), 2\pi F_2(\theta_2)), \qquad 0 \leq \theta_1, \theta_2 < 2\pi. \tag{1}$$

C is itself a bivariate circular distribution function, with marginal distributions that are circular uniform rather than the (linear) uniform ones of copulas. (For discussion of the most appropriate

[1] *Address for correspondence*: Arthur Pewsey, Department of Mathematics, Escuela Politécnica, University of Extremadura, 10003 Cáceres, Spain.
Email: apewsey@unex.es

starting points for circular distribution functions, see Section 4.2.) Taking partial derivatives, the density of (Θ_1, Θ_2) is

$$f(\theta_1, \theta_2) = 4\pi^2 f_1(\theta_1) f_2(\theta_2) c(2\pi F_1(\theta_1), 2\pi F_2(\theta_2)), \tag{2}$$

where f_1 and f_2 are the marginal densities of Θ_1 and Θ_2 and c is the circula density corresponding to C. Importantly, circula densities are periodic with

$$c(\theta_1 \pm 2\pi k, \theta_2 \pm 2\pi l) = c(\theta_1, \theta_2), \qquad k, l = 0, 1, \ldots .$$

In addition, circula densities are usually assumed to be continuous at $\theta_1 = 2k\pi$ and $\theta_2 = 2l\pi$ as well as all other points on the circle. It follows from (2) that the conditional density of $\Theta_2|\Theta_1 = \theta_1$ is

$$f(\theta_2|\theta_1) = 4\pi^2 f_2(\theta_2) c(2\pi F_1(\theta_1), 2\pi F_2(\theta_2)), \tag{3}$$

likewise for $f(\theta_1|\theta_2)$.

Jones *et al.* (2015) provide an in-depth treatment of what, as far as we are aware, has been the only explicitly proposed circula to date. Often ascribed to Wehrly & Johnson (1980), that circula has density $c(2\pi F_1(\theta_1), 2\pi F_2(\theta_2)) = g(2\pi F_1(\theta_1) \pm 2\pi F_2(\theta_2))/2\pi$, g being a circular density referred to as the 'binding' density. Our purpose here is to introduce a far more general, Fourier series based, construction for flexible circula distributions with simple closed-form expressions for their densities.

Section 2 provides the details of the proposed general construction together with results for three circular dependence measures and the conditional mean directions and mean resultant lengths of circulas generated using it. The way dependence measures, in particular, depend on the Fourier coefficients on which the circulas are based is especially simple and attractive. Section 3 introduces a number of fundamental classes of circulas obtained using the general construction and provides details of their basic properties. The second of the classes corresponds to the circula of Wehrly & Johnson (1980). Section 4 discusses simulation, model identification, parameter estimation and tests for independence and goodness-of-fit, both for the proposed circulas and the models for toroidal data derived from them. Two of the new classes and the proposed inferential methodology are applied in illustrative analyses of two toroidal datasets in Section 5. The paper ends with Section 6 devoted to discussion. An appendix provides proofs of Theorems 1 and 2, and a Supplementary Materials document available from the journal website presents some extended classes of circulas and various complementary figures. To distinguish them from the ones in this main document, the labels of the figures and sections in the Supplementary Materials document include an initial letter S prior to their numbers.

2 Fourier series based circula densities

It is well-known (Mardia & Jupp, 1999, Sec. 3.3.2) that any continuous circular density, f, can be expressed in the form of a Fourier series as

$$f(\theta) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \phi(m)\, e^{-im\theta}, \quad -\pi \le \theta < \pi,$$

for appropriately chosen Fourier coefficients $\phi(m)$, $m \in \mathbb{Z}$. Here we consider a family of continuous distributions on the torus whose density can be expressed analogously as

$$f(\theta_1, \theta_2) = \frac{1}{4\pi^2} \sum_{m,n=-\infty}^{\infty} \phi(m, n)\, e^{-i(m\theta_1 + n\theta_2)}, \quad -\pi \le \theta_1, \theta_2 < \pi, \tag{4}$$

where the Fourier coefficients $\phi(m,n) \in \mathbb{C}$, $m,n \in \mathbb{Z}$ are appropriately defined so that $f(\theta_1,\theta_2) \geq 0$ and $\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(\theta_1,\theta_2)\mathrm{d}\theta_1\mathrm{d}\theta_2 = 1$. As the following theorem shows, $\phi(m,n)$ is related to the characteristic function of the distribution with density (4).

Theorem 1 *If a random vector (Θ_1,Θ_2) has density (4) then*

$$E\left[e^{i(m\Theta_1+n\Theta_2)}\right] = \phi(m,n), \quad m,n \in \mathbb{Z}.$$

It follows from Theorem 1 that, for any $\phi(m,n)$ in (4),

$$\overline{\phi(m,n)} = \phi(-m,-n), \quad |\phi(m,n)| \leq 1 \quad \text{and} \quad \phi(0,0) = 1, \tag{5}$$

where $\overline{z}$ denotes the complex conjugate of z. So, if $\phi(m,n)$ is real, which is the case for all the classes of circulas proposed in Section 3, $\phi(-m,-n) = \phi(m,n)$. Henceforth we refer to the series on the right-hand side of (4) as a Fourier series.

The following theorem identifies those densities in (4) that are circula densities, i.e. for which both marginal densities are circular uniform.

Theorem 2 *A density in family (4) is a circula density if and only if*

$$\phi(m,0) = \begin{cases} 1, & m = 0, \\ 0, & m \neq 0, \end{cases} \quad \textit{and} \quad \phi(0,n) = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0, \end{cases} \quad m,n \in \mathbb{Z}. \tag{6}$$

Such circula densities can be expressed as

$$\begin{aligned} c(\theta_1,\theta_2) &= \frac{1}{4\pi^2}\left[1 + 2\,\mathrm{Re}\left\{\sum_{m=1}^{\infty}\sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \phi(m,n)\, e^{-i(m\theta_1+n\theta_2)}\right\}\right] \\ &= \frac{1}{4\pi^2}\left[1 + 2\,\mathrm{Re}\left\{\sum_{n=1}^{\infty}\sum_{\substack{m=-\infty \\ m\neq 0}}^{\infty} \phi(m,n)\, e^{-i(m\theta_1+n\theta_2)}\right\}\right]. \end{aligned} \tag{7}$$

Note that if $\phi(m,n) = 0$ for all $m,n \neq 0$, the distribution is the bivariate circular uniform. The remainder of the paper focuses on developing circula densities using Equations (4) and/or (7), their use in deriving flexible models through Equation (2), and the application of such models in the modelling of toroidal data.

2.1 Circular dependence measures

Here we provide general results for three existing signed circular dependence measures when applied to circulas. The measures considered are those of Rivest (1982), Jammalamadaka & Sarma (1988) and Fisher & Lee (1983) which we denote by ρ_{R}, ρ_{JS} and ρ_{FL}, respectively.

For a circula, the circular dependence measures of Rivest (1982) and Jammalamadaka & Sarma (1988) are the same and have a very simple and close relationship with the circula densities in Equation (7), being given in terms of $\phi(m,n)$ as $\rho_{\mathrm{R}} = \rho_{\mathrm{JS}} = |\phi(1,-1)| - |\phi(1,1)|$. Clearly, when only one of $\phi(1,-1)$ and $\phi(1,1)$ is non-zero, then

$$\rho_{\mathrm{R}} = \rho_{\mathrm{JS}} = \begin{cases} |\phi(1,-1)| & \text{if } \phi(1,-1) \neq 0, \\ -|\phi(1,1)| & \text{if } \phi(1,1) \neq 0. \end{cases}$$

For a circula it is easily shown that the circular dependence measure of Fisher & Lee (1983) is given by $\rho_{\mathrm{FL}} = |\phi(1,-1)|^2 - |\phi(1,1)|^2$. Hence, when only one of $\phi(1,-1)$ and $\phi(1,1)$ is non-zero,

$$\rho_{\mathrm{FL}} = \begin{cases} |\phi(1,-1)|^2 & \text{if } \phi(1,-1) \neq 0, \\ -|\phi(1,1)|^2 & \text{if } \phi(1,1) \neq 0, \end{cases}$$

and thus the *magnitude* of ρ_{FL} is the square of the magnitude of $\rho_{\mathrm{R}} = \rho_{\mathrm{JS}}$.

These findings are appealing because of their simplicity.

2.2 Conditional mean directions and resultant lengths

Using density (4) with circular uniform marginals, the first trigonometric moment of $\Theta_1|\Theta_2 = \theta_2$ for the corresponding circula can be expressed as

$$\begin{aligned} E(e^{i\Theta_1}|\theta_2) &= \int_{-\pi}^{\pi} e^{i\theta_1} c(\theta_1|\theta_2) \mathrm{d}\theta_1 \\ &= \int_{-\pi}^{\pi} e^{i\theta_1} \left(\frac{1}{2\pi} \sum_{m,n=-\infty}^{\infty} \phi(m,n)\, e^{-i(m\theta_1 + n\theta_2)} \right) \mathrm{d}\theta_1 \\ &= \frac{1}{2\pi} \sum_{m,n=-\infty}^{\infty} \phi(m,n) \int_{-\pi}^{\pi} e^{-i\{(m-1)\theta_1 + n\theta_2\}} \mathrm{d}\theta_1 \\ &= \sum_{n=-\infty}^{\infty} \phi(1,n) e^{-in\theta_2}, \end{aligned} \tag{8}$$

where $c(\theta_1|\theta_2)$ denotes the conditional density of $\Theta_1|\Theta_2 = \theta_2$. Similarly,

$$E(e^{i\Theta_2}|\theta_1) = \sum_{m=-\infty}^{\infty} \phi(m,1) e^{-im\theta_1} \tag{9}$$

$$= \sum_{m=1}^{\infty} \phi(m,1) e^{-im\theta_1} + \overline{\sum_{m=1}^{\infty} \phi(m,-1) e^{-im\theta_1}}. \tag{10}$$

Expression (8) implies that, in order to calculate the mean direction and mean resultant length of $\Theta_1|\Theta_2 = \theta_2$, it suffices to focus on the non-zero Fourier coefficients in the $m = 1$ column of the planar plot of $\{\phi(m,n)\}_{m,n\in\mathbb{Z}}$. (See Figure 1 for examples of such plots.) It then follows that:

(a) If all the coefficients in the $m = 1$ column are zero, the mean resultant length of the conditional distribution of $\Theta_1|\Theta_2 = \theta_2$, $R(\Theta_1|\theta_2) \equiv |E(e^{i\Theta_1}|\theta_2)|$, is zero;

(b) If there is only one non-zero coefficient in the $m = 1$ column, say at $(m,n) = (1,d)$, the mean direction of the conditional distribution of $\Theta_1|\Theta_2 = \theta_2$, $M(\Theta_1|\theta_2) \equiv \arg(E(e^{i\Theta_1}|\theta_2))$, is $-d\theta_2$ and $R(\Theta_1|\theta_2) = |\phi(1,d)|$;

(c) If there are two or more non-zero coefficients in the $m = 1$ column, $M(\Theta_1|\theta_2)$ is nonlinear and $R(\Theta_1|\theta_2)$ is in general heteroscedastic.

Analogous results for the conditional distribution of $\Theta_2|\Theta_1 = \theta_1$ follow from considering Equation (9). From the equivalent expression (10), the mean direction and mean resultant length of $\Theta_2|\Theta_1 = \theta_1$ only depend on the non-zero coefficients for $m \geq 1$ in the $n = \pm 1$ rows.

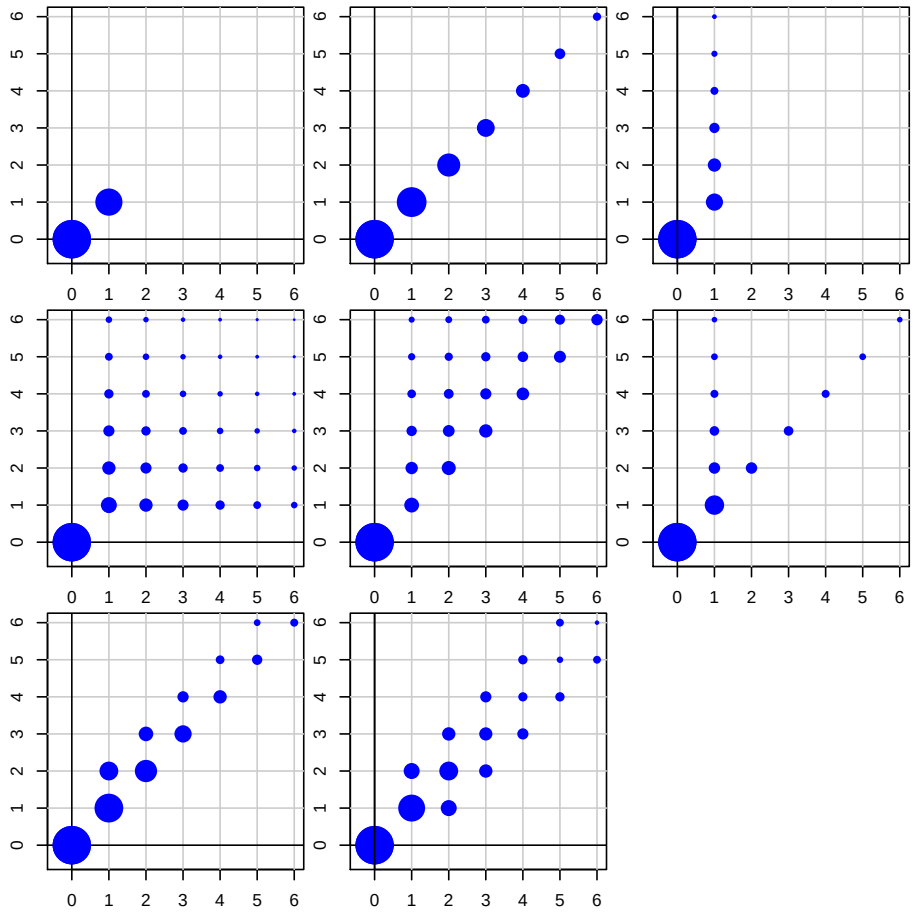

Figure 1: Examples of patterns of non-zero Fourier coefficients for Classes 1–8 with $q = -1$ and $m, n \in \{0, 1, ..., 6\}$. The values of m appear on the horizontal axes, and those for n on the vertical axes. The area of a dot at (m, n) is proportional to the value of $\phi(m, n)$. The panels for the eight classes are ordered from top left to bottom right, reading from left to right.

3 Classes of Fourier series based circula densities

In this section we consider eight classes of (7) with simple patterns for their non-zero Fourier coefficients and provide details of their basic properties. Given Equation (5) and Theorem 2, unless explicitly stated otherwise, we consider classes of circula densities defined through non-zero Fourier coefficients on the $\mathbb{Z}^+ \times (\mathbb{Z}^+ \cup \mathbb{Z}^-)$ lattice. We begin with densities generated using very simple arrangements for their non-zero Fourier coefficients before moving on to others derived using more elaborate patterns. Figure 1 illustrates patterns of the non-zero Fourier coefficients corresponding to Classes 1–8. The special cases of the various classes of circulas considered are derived using geometric series of non-zero Fourier coefficients to generate flexible models with closed-form expressions for their densities. In order to obtain circula densities exhibiting pointwise symmetry about the origin, all of the non-zero Fourier coefficients of the eight classes are assumed to be real. A circula density, c, is pointwise symmetric about $(0, 0)$ if $c(\theta_1, \theta_2) = c(-\theta_1, -\theta_2)$ for all $(\theta_1, \theta_2) \in [-\pi, \pi)^2$.

Since all of the classes of circulas we consider have at most one of $\phi(1, -1)$ and $\phi(1, 1)$ which is non-zero, we only quote values of ρ_{R}, those for ρ_{JS} and ρ_{FL} following from the relations identified in Section 2.1.

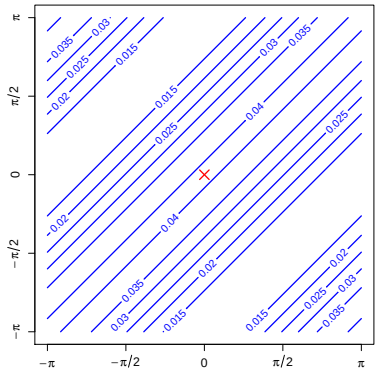

Figure 2: Planar contour plot of circula density (11) with parameter vector $(q, \gamma) = (1, 0.3)$. The cross identifies $(\theta_1, \theta_2) = (0, 0)$.

3.1 Class 1: Single (non-zero) point

Let

$$\phi(m,n) = \begin{cases} \gamma, & (m,n) = (1,-q), \\ 0, & \text{otherwise}, \end{cases}$$

where $0 \leq \gamma \leq 1/2$, and $q \in \{-1, 1\}$. Thus, this class is generated by a single non-zero real-valued Fourier coefficient on the $\mathbb{Z}^+ \times (\mathbb{Z}^+ \cup \mathbb{Z}^-)$ lattice.

In this case, the circula density (7) reduces to

$$c(\theta_1, \theta_2) = \frac{1}{4\pi^2} \left\{1 + 2\gamma \cos(\theta_1 - q\theta_2)\right\}. \tag{11}$$

The mode and antimode of (11) occur at every point along the lines $\theta_1 = q\theta_2$ and $\theta_1 = q\theta_2 + \pi$, respectively, where $-\pi \leq \theta_1, \theta_2 < \pi$. Figure 2 portrays a planar contour plot of density (11) when $q = 1$ and $\gamma = 0.3$.

For this class, $\rho_{\mathrm{R}} = q\gamma$. The strength of dependence between Θ_1 and Θ_2 is thus controlled by γ, and its sign by q. This role of q is the same for all the classes considered in this section. The conditional distributions of $\Theta_1|\Theta_2 = \theta_2$ and $\Theta_2|\Theta_1 = \theta_1$ are cardioid distributions on the circle with mean directions $M(\Theta_1|\theta_2) = q\theta_2$ and $M(\Theta_2|\theta_1) = q\theta_1$, respectively, and mean resultant lengths $R(\Theta_1|\theta_2) = R(\Theta_2|\theta_1) = \gamma$. So this class affords 'linear-homoscedastic' circular-circular regression in both directions.

3.2 Class 2: Diagonal line

Here we consider a class of circula densities generated using non-zero Fourier coefficients on a diagonal of the $\mathbb{Z}^+ \times (\mathbb{Z}^+ \cup \mathbb{Z}^-)$ lattice. Specifically, let

$$\phi(m,n) = \begin{cases} \psi(m), & n = -qm, \\ 0, & n \neq -qm, \end{cases}$$

where the $\psi(m)$ are the Fourier coefficients of any circular distribution, for which $\psi(0) = 1$, and $q \in \{-1, 1\}$. Then the circula density is of the form

$$c(\theta_1, \theta_2) = \frac{1}{2\pi} g(\theta_1 - q\theta_2), \tag{12}$$

where $g(\theta) = (2\pi)^{-1}[1 + 2\mathrm{Re}\{\sum_{m=1}^{\infty} \psi(m)e^{-im\theta}\}]$ is a density on the unit circle. This is the class of circulas discussed in Jones *et al.* (2015). For all but the case when g is circular uniform, (12) has linear contours parallel to the $q\pi/4$ diagonal. If the polar representation of $g(\theta)$ is unimodal with mode (antimode) at $\theta = \omega_g^*$, then the maxima (minima) of (12) are situated along the line $\theta_1 = q\theta_2 + \omega_g^*$ $(-\pi \leq \theta_1, \theta_2 < \pi)$.

When applying Equation (8),

$$\phi(1, n) = \begin{cases} \psi(1) = R_g e^{iM_g}, & n = -q, \\ 0, & \text{otherwise}, \end{cases}$$

where M_g and R_g are the mean direction and mean resultant length, respectively, of the distribution with density g. Therefore, $E(e^{i\Theta_1}|\theta_2) = R_g e^{i(q\theta_2 + M_g)}$, and hence $M(\Theta_1|\theta_2) = q\theta_2 + M_g$ and $R(\Theta_1|\theta_2) = R_g$. Using Equation (10), $M(\Theta_2|\theta_1) = q(\theta_1 - M_g)$ and $R(\Theta_2|\theta_1) = R_g$. This class thus affords a general class of circulas with 'linear-homoscedastic' circular-circular regression in both directions, R_g taking the role of γ in Class 1. For this model, $\rho_{\mathrm{R}} = qR_g$. Density (12) reduces to (11) when $\psi(1) = \gamma$ and $\psi(m) = 0$ otherwise, and $0 \leq \gamma \leq 1/2$.

When $\psi(m) = \gamma\rho^{m-1}$ $(m = 1, 2, ..., 0 \leq \gamma < 1,\ 2\gamma - 1 \leq \rho < 1)$, density (12) can be expressed in closed form as

$$c(\theta_1, \theta_2) = \frac{1}{4\pi^2}\left\{1 + 2\gamma\,\frac{\cos(\theta_1 - q\theta_2) - \rho}{1 + \rho^2 - 2\rho\cos(\theta_1 - q\theta_2)}\right\}. \tag{13}$$

For this subclass, all three parameters, $\gamma = \phi(1, -q) = \psi(1)$, ρ and q, affect the dependence between θ_1 and θ_2 although, as elsewhere, $\rho_R = q\gamma$. Figure S4 presents planar contour plots of density (13) for $q = -1$, $\gamma = 0.7$ and three values of ρ. Note that density (13) can also be represented as the following two component mixture with toroidal uniform and wrapped Cauchy-like component densities and mixing probability γ/ρ,

$$c(\theta_1, \theta_2) = \left(1 - \frac{\gamma}{\rho}\right)\frac{1}{4\pi^2} + \left(\frac{\gamma}{\rho}\right)\frac{1}{4\pi^2}\frac{1 - \rho^2}{1 + \rho^2 - 2\rho\cos(\theta_1 - q\theta_2)}.$$

The parameter ρ regulates the concentration of the wrapped Cauchy-like distribution. The maxima and minima of density (13) are given along the lines $\theta_1 = q\theta_2$ and $\theta_1 = q\theta_2 + \pi$ $(-\pi \leq \theta_1, \theta_2 < \pi)$, respectively, for any γ and $\rho \geq 0$. Both conditional distributions are special cases of the circular distribution of Kato and Jones (2015) which are two component mixtures with circular uniform and wrapped Cauchy components. In addition, it immediately follows from the general results given in the last paragraph that the mean directions and mean resultant lengths for both conditionals are $M(\Theta_1|\theta_2) = q\theta_2$, $M(\Theta_2|\theta_1) = q\theta_1$ and $R(\Theta_1|\theta_2) = R(\Theta_2|\theta_1) = \gamma$.

3.3 Class 3: Vertical line

Now consider the class of circula densities generated using non-zero Fourier coefficients on the vertical line of the $\mathbb{Z}^+ \times (\mathbb{Z}^+ \cup \mathbb{Z}^-)$ lattice given by

$$\phi(m, n) = \begin{cases} \psi(n), & m = 1,\ qn \leq -1, \\ 0, & \text{otherwise}, \end{cases} \tag{14}$$

for some non-zero Fourier coefficients ψ and $q \in \{-1, 1\}$.

For example, if $\psi(n) = \gamma\rho^{|n|-1}$ $(0 \le \gamma \le 1/2,\ 0 \le \rho \le 1-2\gamma)$, then

$$c(\theta_1, \theta_2) = \frac{1}{4\pi^2}\left\{1 + 2\gamma\frac{\cos(\theta_1 - q\theta_2) - \rho\cos\theta_1}{1 + \rho^2 - 2\rho\cos\theta_2}\right\}. \tag{15}$$

The mode and antimode of density (15) occur at $(\theta_1, \theta_2) = (0, 0)$ and $(\theta_1, \theta_2) = (-\pi, 0)$, respectively. As for density (13), $\gamma = \phi(1, -q) = \psi(1)$ so that $\rho_{\mathrm{R}} = q\gamma$. Panels (a) and (b) of Figure S3 present planar contour plots of density (15) with $q = 1$ and two (γ, ρ) combinations. The dependence between Θ_1 and Θ_2 is clearly regulated by all three parameters: q, γ and ρ. In particular, γ regulates the strength of the dependence between Θ_1 and Θ_2, and ρ the degree of deformation of the density's shape around the main diagonal.

As was the case for density (13), density (15) can also be represented as a two-component mixture, in this case, as a mixture of the toroidal uniform density and (15) with its $\gamma = 1/2$. Specifically,

$$c(\theta_1, \theta_2) = (1 - 2\gamma)\frac{1}{4\pi^2} + 2\gamma\frac{1}{4\pi^2}\left\{1 + \frac{\cos(\theta_1 - q\theta_2) - \rho\cos\theta_1}{1 + \rho^2 - 2\rho\cos\theta_2}\right\},$$

the mixing probability being 2γ.

Using Equation (8), the first trigonometric moment of the conditional distribution of $\Theta_1|\Theta_2 = \theta_2$ is $E(e^{i\Theta_1}|\theta_2) = \gamma/(e^{-iq\theta_2} - \rho)$. Hence, $M(\Theta_1|\theta_2) = \arg(\gamma/(e^{-iq\theta_2} - \rho)) = \mathrm{atan2}(q\sin\theta_2, \cos\theta_2 - \rho)$, where $\mathrm{atan2}(y, x)$ returns the angle measured anticlockwise from the x-axis to the vector connecting the origin with (x, y). This regression curve is a special case of the regression curve of Rivest (1997). Also, $R(\Theta_1|\theta_2) = |\gamma/(e^{-iq\theta_2} - \rho)| = \gamma/(1 + \rho^2 - 2\rho\cos\theta_2)^{1/2}$. Using Equation (10), $M(\Theta_2|\theta_1) = q\theta_1$ and $R(\Theta_2|\theta_1) = \gamma$. The conditional distribution of $\Theta_1|\Theta_2 = \theta_2$ is a cardioid distribution and that of $\Theta_2|\Theta_1 = \theta_1$ follows a special case of the distribution of Kato & Jones (2015).

It is possible to extend the domain of ρ to include $[2\gamma - 1, 0)$. However this extension does not generalise the density (15) apart from location shifts because of the relationship $c(\theta_1, \theta_2; \gamma, \rho, q) = c(\theta_1 + \pi, \theta_2 + \pi; \gamma, -\rho, q)$.

3.4 Class 4: Square

This class arises from using a square pattern of non-zero Fourier coefficients. As an attractive three-parameter example of this, let

$$\phi(m, n) = \begin{cases} \gamma\rho_1^{m-1}\rho_2^{|n|-1}, & m, -qn \in \mathbb{Z}^+, \\ 0, & \text{otherwise}, \end{cases}$$

where $0 < \rho_1, \rho_2 < 1$ and $q \in \{-1, 1\}$. Density (7) then becomes

$$c(\theta_1, \theta_2) = \frac{1}{4\pi^2}\left\{1 + 2\gamma\frac{\cos(\theta_1 - q\theta_2) - \rho_2\cos\theta_1 - \rho_1\cos\theta_2 + \rho_1\rho_2}{(1 + \rho_1^2 - 2\rho_1\cos\theta_1)(1 + \rho_2^2 - 2\rho_2\cos\theta_2)}\right\}, \quad -\pi \le \theta_1, \theta_2 < \pi, \tag{16}$$

where

$$0 \le \gamma \le \begin{cases} \rho_1\rho_2/[\{(1-\rho_1^2)(1-\rho_2^2)\}^{-1} - 1], & |\rho_1 - \rho_2| \le \rho_1\rho_2, \\ (1-\rho_1)(1+\rho_2)/2, & |\rho_1 - \rho_2| > \rho_1\rho_2,\ \rho_1 > \rho_2, \\ (1+\rho_1)(1-\rho_2)/2, & |\rho_1 - \rho_2| > \rho_1\rho_2,\ \rho_1 < \rho_2. \end{cases} \tag{17}$$

Figure S5 presents a contour plot of the maximum value of γ as a function of ρ_1 and ρ_2. The maximum of γ tends to 0.5 as $(\rho_1, \rho_2) \to (0, 0)$.

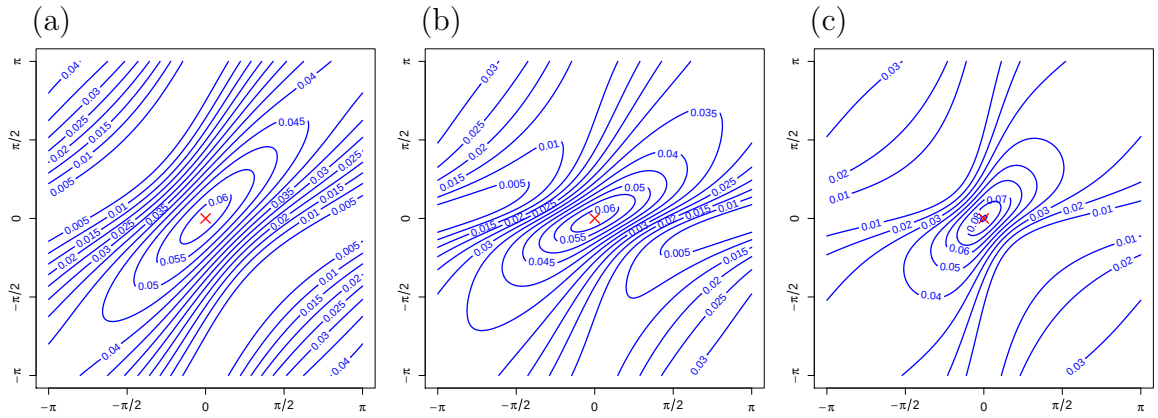

Figure 3: Planar contour plots of circula density (16) with $q = 1$ and (ρ_1, ρ_2, γ) equal to: (a) $(0.2, 0.2, 0.470)$; (b) $(0.2, 0.5, 0.300)$; (c) $(0.5, 0.5, 0.321)$. For each panel, the value of γ corresponds to the upper bound of inequality (17).

The mode of density (16) occurs at $(\theta_1, \theta_2) = (0, 0)$. If $|\rho_1 - \rho_2| > \rho_1\rho_2$, then there is only one antimode of (16) which occurs at $(\theta_1, \theta_2) = (0, -\pi)$ for $\rho_1 > \rho_2$ and $(\theta_1, \theta_2) = (-\pi, 0)$ for $\rho_1 < \rho_2$. When $|\rho_1 - \rho_2| \leq \rho_1\rho_2$, there are two antimodes which occur at $(\theta_1, \theta_2) = \pm(\alpha_1, -q\alpha_2)$, where $\alpha_j = \arg\{\rho_j + (1 - \rho_j^2)/(\rho_j + e^{ix_j})\} = \text{atan2}\{-(1 - \rho_j^2)\sin x_j, \rho_j(2 + \rho_j \cos x_j)\}$. For this model, $\rho_{\text{R}} = q\gamma$.

Planar contour plots of circula density (16) designed to illustrate the roles of ρ_1 and ρ_2 are displayed in Figure 3. When $\rho_1 = \rho_2$ the density is symmetric about the main diagonal and increasingly concentrated in the neighbourhood of the origin as $\rho_1 = \rho_2$ increases. For a fixed value of ρ_1, as ρ_2 increases the main axis of the central elliptical contour tilts increasingly away from the main diagonal towards $\theta_2 = 0$ and the dispersion increases in the neighbourhood of $(-\pi, -\pi) = (\pi, \pi)$. Due to the symmetry of (16), for a fixed value of ρ_2 the main axis tilts increasingly towards $\theta_1 = 0$ as ρ_1 increases.

The conditional density of $\Theta_1|\Theta_2 = \theta_2$ can be seen to be that of a special case of the Kato & Jones (2015) distribution. Proceeding as for Class 3, $M(\Theta_1|\theta_2) = \text{atan2}(q \sin\theta_2, \cos\theta_2 - \rho_2)$ and $R(\Theta_1|\theta_2) = \gamma/(1 + \rho_2^2 - 2\rho_2 \cos\theta_2)^{1/2}$. The conditional distribution of $\Theta_2|\Theta_1 = \theta_1$ is also a special case of the Kato & Jones (2015) distribution, with $M(\Theta_2|\theta_1) = q\theta_1$ and $R(\Theta_2|\theta_1) = \gamma$.

3.5 Class 5: Upper triangle

The circula density (13) of Class 2 can be extended using an arrangement of non-zero Fourier coefficients forming an upper triangular pattern with the following particular specification:

$$\phi(m, n) = \begin{cases} \gamma\rho^{m-1}\lambda^{-(m+qn)}, & 1 \leq m \leq -qn, \\ 0, & \text{otherwise}, \end{cases}$$

where $0 < \rho < 1$, $0 \leq \lambda < 1$, and $q \in \{-1, 1\}$. The constraints on γ are

$$0 \leq \gamma \leq \begin{cases} \rho(1 - \rho^2)(1 - \lambda^2)/[1 - \rho^2(1 - \lambda^2)], & \lambda \geq (1 - \rho)/\rho, \\ (1 - \lambda)(1 + \rho)/2, & \lambda < (1 - \rho)/\rho. \end{cases} \tag{18}$$

A contour plot of the maximum value of γ as a function of ρ and λ is presented in Figure S6. The maximum of γ tends to 1 as $(\rho, \lambda) \to (1, 0)$.

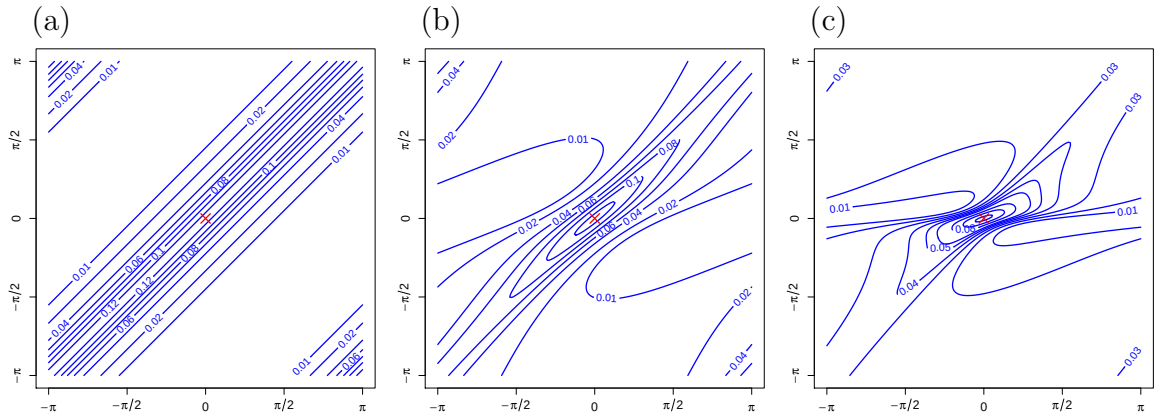

Figure 4: Planar contour plots of circula density (19) for $q = 1$, $\rho = 0.6$ and (λ, γ) equal to: (a) $(0, 0.8)$; (b) $(0.4, 0.48)$; (c) $(0.8, 0.16)$. For each panel, the value of γ corresponds to the upper bound of inequality (18).

For this case, the circula density is

$$c(\theta_1, \theta_2) = \frac{1}{4\pi^2}\left[1 + 2\gamma \frac{\cos(\theta_1 - q\theta_2) - \lambda\cos\theta_1 + \rho\lambda\cos\theta_2 - \rho}{\{1 + \rho^2 - 2\rho\cos(\theta_1 - q\theta_2)\}(1 + \lambda^2 - 2\lambda\cos\theta_2)}\right], \quad -\pi \le \theta_1, \theta_2 < \pi. \tag{19}$$

The mode of the density is located at $(\theta_1, \theta_2) = (0, 0)$, and the antimode(s) at $(\theta_1, \theta_2) = (-\pi, 0)$ if $\lambda < (1 - \rho)/\rho$ or at $(\theta_1, \theta_2) = \pm(\alpha_1, -q\alpha_2)$ if $\lambda \ge (1 - \rho)/\rho$. Here $\alpha_1 = -\alpha_2 + \text{atan2}\{(1 - \rho^2)\sin x_1, \rho(2 + \rho\cos x_1)\}$, $\alpha_2 = \text{atan2}\{-(1 - \lambda^2)\sin x_2, 2\lambda + (1 + \lambda^2)\cos x_2\}$, $x_1 = \arccos\{(-1 - \rho^2 + \rho^2\lambda^2)/(2\rho)\}$, and $x_2 = \arccos\{(1 - \rho^2 - \rho^2\lambda^2)/(2\rho^2\lambda)\}$. When $\lambda = 0$, (19) reduces to (13) of Class 2. For this model too, $\rho_R = q\gamma$.

Figure 4 presents planar contour plots of the circula density (19) for $q = 1$, $\rho = 0.6$ and three combinations of λ and γ. When $\lambda = 0$ the contours are straight lines. This is because (19) reduces to (13) when $\lambda = 0$. As λ increases, the axis of the central elliptical contour tilts increasingly away from the main diagonal towards $\theta_2 = 0$ and the other contours tend to be increasingly asymmetric about the main diagonal.

The conditional distribution of Θ_1 given $\Theta_2 = \theta_2$ is a special case of the Kato & Jones (2015) distribution, while that of Θ_2 given $\Theta_1 = \theta_1$ does not seem to be a well-known distribution in general. For this model, $M(\Theta_1|\theta_2) = \text{atan2}(q\sin\theta_2, \cos\theta_2 - \lambda)$ and $R(\Theta_1|\theta_2) = \gamma/(1 + \lambda^2 - 2\lambda\cos\theta_2)^{1/2}$. Their counterparts for the conditional distribution of $\Theta_2|\Theta_1 = \theta_1$ are $M(\Theta_2|\theta_1) = q\theta_1$ and $R(\Theta_2|\theta_1) = \gamma$. Thus, as is also the case for Class 3, this class affords 'nonlinear-heteroscedastic' circular-circular regression for Θ_1 as a function of Θ_2, and linear-homoscedastic circular-circular regression for Θ_2 as a function of Θ_1. On the contrary, Classes 1 and 2 provide models for only linear-homoscedastic circular-circular regression. Class 4 offers nonlinear-heteroscedastic circular-circular regression for both conditionals.

3.6 Class 6: Diagonal and vertical lines

Here we consider a class of circulas defined by the Fourier coefficients

$$\phi(m, n) = p\phi_2(m, n) + (1 - p)\phi_3(m, n), \quad m \in \mathbb{Z}^+, \ n \in \mathbb{Z}^+ \cup \mathbb{Z}^-, \tag{20}$$

where

$$\phi_2(m,n) = \begin{cases} \gamma\rho^{m-1}, & n = -qm, \\ 0, & n \neq -qm, \end{cases} \qquad \phi_3(m,n) = \begin{cases} \gamma\rho^{|n|-1}, & m = 1,\ qn \leq -1, \\ 0, & \text{otherwise}, \end{cases} \tag{21}$$

$0 \leq p, \rho < 1,\ q \in \{-1, 1\}$, and $0 \leq \gamma \leq (1-\rho^2)/2(1+\rho-2p\rho)$. From (20) and (21), it can be seen that this class of circulas arises as a two-component mixture of the previously considered special cases of Classes 2 and 3, namely model (13) and model (15).

As a mixture, the shapes of the density contours are weighted combinations of those in Figures S4 and S3, the value of p determining their relative weights. Figure S7 presents planar contour plots of examples of this density when $q = 1$ and the two component circula densities are weighted equally. Its mode and antimode are located at $(\theta_1, \theta_2) = (0, 0)$ and $(\theta_1, \theta_2) = (-\pi, 0)$, respectively. As for both of its component models, $\gamma = |\phi(1, -q)|$ for this class too and $\rho_R = q\gamma$. Once again, ρ is a shape parameter. Clearly, the conditional distributions of the mixture model are mixtures of the conditional distributions of its two components.

An alternative representation of the parameter space of the circula density under consideration is $q \in \{-1, 1\}$, $0 \leq \gamma < 1$,

$$p \in \begin{cases} [0, 1], & \gamma \leq 1/2, \\ \left[(1 + \sqrt{2\gamma - 1}/\gamma)/2, 1\right], & \gamma > 1/2, \end{cases}$$

and $\max(0, \alpha_1 - \alpha_2) \leq \rho \leq \alpha_1 + \alpha_2$, where $\alpha_1 = \gamma(2p-1)$ and $\alpha_2 = (\gamma^2(2p-1)^2 - 2\gamma + 1)^{1/2}$. This representation enables us to determine the ranges of p and ρ for a fixed value of γ.

3.7 Class 7: Bidiagonal

Here we consider an alternative two-component mixture, this time of Class 2 and a shifted version of Class 2 (both are special cases of the more general diagonal line structure explored in Subsection S1.2 of the Supplementary Material). Its Fourier coefficients are

$$\phi(m,n) = p\phi_2(m,n) + (1-p)\phi_{\mathrm{E2}}(m,n), \quad m \in \mathbb{Z}^+,\ n \in \mathbb{Z}^+ \cup \mathbb{Z}^-, \tag{22}$$

where $\phi_2(m,n)$ is as in (21) and

$$\phi_{\mathrm{E2}}(m,n) = \begin{cases} \gamma\rho^{m-1}, & n = -q(m+1), \\ 0, & \text{otherwise}. \end{cases}$$

This class of circulas has density

$$c(\theta_1, \theta_2) = \frac{1}{4\pi^2}\left[1 + 2\gamma\frac{p\{\cos(\theta_1 - q\theta_2) - \rho\} + (1-p)\{\cos(\theta_1 - 2q\theta_2) - \rho\cos\theta_2\}}{1 + \rho^2 - 2\rho\cos(\theta_1 - q\theta_2)}\right], \quad -\pi \leq \theta_1, \theta_2 < \pi. \tag{23}$$

Here the ranges of the parameters are $0 \leq p, \rho < 1$ and

$$0 \leq \gamma \leq \begin{cases} (1-\rho)/\{2(1-2p)\}, & 0 \leq p < \rho/(1+2\rho), \\ (1-\rho^2)p\rho/\{p^2 + (1-2p)\rho^2\}, & \rho/(1+2\rho) \leq p < \rho, \\ (1+\rho)/2, & \rho \leq p < 1. \end{cases} \tag{24}$$

Figure S8 provides a contour plot of the maximum possible value of γ as a function of ρ and p. From that plot it can be seen that, for a fixed value of ρ, the maximum possible value of γ is monotonically nondecreasing with respect to p.

Figure S9 presents planar contour plots for examples of density (23) when $q = 1$ and the two components are weighted equally. Although a superficial visual inspection of its panel (a) might suggest the contrary, all three circula densities portrayed are unimodal on the torus. The mode of (23) is located at $(\theta_1, \theta_2) = (0, 0)$, and its antimode(s) at $(\theta_1, \theta_2) = (-\pi, -\pi)$ if $0 \leq p < \rho/(1 + 2\rho)$, $(\theta_1, \theta_2) = (g(\alpha), \alpha)$ and $(\theta_1, \theta_2) = (g(-\alpha), -\alpha)$ if $\rho/(1 + 2\rho) \leq p < \rho$, and $(\theta_1, \theta_2) = (-\pi, 0)$ if $\rho < p < 1$, where $\alpha = \arccos[\{p(1 - \rho^2) - (1 - p)^2\rho^2\}/\{2(1 - p)p\rho^2\}]$ and $g(\alpha) = 2\alpha - \pi/2 - \arccos(\rho \sin \alpha)$.

For this model, $\rho_{\mathrm{R}} = qp\gamma$, with $p\gamma$ — rather than γ alone — regulating the strength of the dependence and q its direction.

3.8 Classes with reflected patterns of Fourier coefficients

Let $\{\phi(m, n)\}_{m,n\in\mathbb{Z}}$ denote the Fourier coefficients of a circula with density $c(\theta_1, \theta_2)$. Then the distribution with Fourier coefficients

$$\phi^*(m, n) = \phi(n, m), \quad m, n \in \mathbb{Z}, \tag{25}$$

is also a circula, with density

$$c^*(\theta_1, \theta_2) = c(\theta_2, \theta_1), \quad -\pi \leq \theta_1, \theta_2 < \pi.$$

This fact can be used to obtain new classes of circulas from those previously discussed. When the Fourier coefficients of such a circula are plotted as in Figure 1, the patterns they adopt are reflections about the main diagonal of those for the original class. For example, if $\{\phi(m, n)\}$ are the Fourier coefficients of a Class 3 circula, as in the top right panel of Figure 1, then the plot of $\{\phi^*(m, n)\}$ has non-zero Fourier coefficients along the horizontal line with $n = -q$. Similarly, the non-zero coefficients of $\phi^*(m, n) = \phi(n, m)$ for Classes 5, 6 and 7 define patterns in the form of a lower triangle, a combination of diagonal line and a horizontal line, and two diagonal lines with $n = -qm$ and $n = -q(m - 1)$ $(m \geq 2)$, respectively. The patterns adopted by the non-zero coefficients of Classes 1, 2 and 4 are unchanged by transformation (25).

Many properties of the transformed classes follow straightforwardly from those of the original classes. For example, the planar contour plot of the density of a transformed circula is a reflection about the $\theta_1 = \theta_2$ diagonal of that for the original circula. Also, the ranges of the parameters of a transformed class are the same as those for the original class.

3.9 Class 8: Tridiagonal

We employ transformation (25) in the definition of our final class of circulas, which turns out to play a major role in the application of Section 5.2: it is a special case of a wider family defined through a mixture of a Class 2 and two shifted versions of Class 2 circulas. The Fourier coefficients of the latter are given by

$$\phi(m, n) = p\phi_2(m, n) + \frac{1 - p}{2}\phi_{\mathrm{E}2}(m, n) + \frac{1 - p}{2}\phi_{\mathrm{E}2}(n, m), \tag{26}$$

where

$$\phi_2(m,n) = \begin{cases} \gamma_1\rho_1^{m-1}, & n = -qm, \\ 0, & \text{otherwise}, \end{cases}, \quad \phi_{\mathrm{E}2}(m,n) = \begin{cases} \gamma_2\rho_2^{m-1}, & n = -q(m+1), \\ 0, & \text{otherwise}, \end{cases}$$

$q \in \{-1, 1\}$ and the parameters $0 \le \gamma_1, \gamma_2, \rho_1, \rho_2 < 1$ and $0 \le p \le 1$ satisfy conditions that are, in general, not easily identified.

The density of the circula with Fourier coefficients (26) is given by

$$c(\theta_1,\theta_2) = \frac{1}{4\pi^2}\Bigg[1 + 2p\gamma_1\frac{\cos(\theta_1 - q\theta_2) - \rho_1}{1 + \rho_1^2 - 2\rho_1\cos(\theta_1 - q\theta_2)} + (1-p)\gamma_2\frac{\cos(\theta_1 - 2q\theta_2) + \cos(\theta_2 - 2q\theta_1) - \rho_2(\cos\theta_1 + \cos\theta_2)}{1 + \rho_2^2 - 2\rho_2\cos(\theta_1 - q\theta_2)}\Bigg], \quad -\pi \le \theta_1, \theta_2 < \pi. \tag{27}$$

The mode of density (27) is $(\theta_1, \theta_2) = (0, 0)$ since this is the mode of each component of the mixture. Its antimodes are $\theta_1 = q\theta_2 + \pi$ $(-\pi \le \theta_2 < \pi)$ for $p = 1$ and $(\theta_1, \theta_2) = (\pi, \pi)$ for $p = 0$. For $0 < p < 1$, there are, in general, no closed-form expressions for the antimodes, but it can be shown that the minimum value of density (27) occurs at $(\theta_1, \theta_2) = (\theta, -q\theta), (-\theta, q\theta)$ for some θ.

The special case of model (27) employed in Section 5.2 is that with $p\gamma_1 = \rho_1 = \rho_2^2$. Reparametrizing, let $\rho = \rho_2$ and $\gamma = (1-p)\gamma_2/2$. The density of this two-parameter special case is given by

$$c(\theta_1,\theta_2;\rho,\gamma) = \frac{1}{4\pi^2}\Bigg[\frac{1-\rho^4}{1 + \rho^4 - 2\rho^2\cos(\theta_1 - q\theta_2)} + 2\gamma\frac{\cos(\theta_1 - 2q\theta_2) + \cos(\theta_2 - 2q\theta_1) - \rho(\cos\theta_1 + \cos\theta_2)}{1 + \rho^2 - 2\rho\cos(\theta_1 - q\theta_2)}\Bigg], \tag{28}$$

where $0 \le \rho < 1$ and

$$0 \le \gamma \le \gamma_{\max} = \frac{1-\rho^4}{4(\rho\cos\theta^* - \cos 3\theta^*)}\frac{1 + \rho^2 - 2\rho\cos 2\theta^*}{1 + \rho^4 - 2\rho^2\cos 2\theta^*}, \tag{29}$$

the latter obtained by solving $c(\theta, -q\theta; \rho, \gamma_{\max}) = 0$ and $dc(\theta, -q\theta; \rho, \gamma_{\max})/d\theta = 0$. In (29), θ^* $(\in (0, \pi])$ is a solution to the equation

$$\sin\theta^*(c_0 + c_1\cos 2\theta^* + c_2\cos^2 2\theta^* + c_3\cos^3 2\theta^*) = 0,$$

where $c_0 = -3+5\rho-3\rho^2-3\rho^3-7\rho^4+\rho^5+\rho^6+\rho^7$, $c_1 = -6+2\rho+6\rho^2-6\rho^3+4\rho^4-4\rho^5-4\rho^6$, $c_2 = 2\rho(2+10\rho-6\rho^2+12\rho^3+2\rho^4)$, $c_3 = -24\rho^3$, and θ^* satisfies $\theta^* = \mathrm{argmin}_{0<\theta\le\pi} c(\theta, -q\theta; \rho, \gamma_{\max}(\rho, \theta^*))$. Note that $\theta^* = \pi$ is a solution to this equation and the equation $c_0 + c_1 x + c_2 x^2 + c_3 x^3 = 0$ has closed-form solutions. Numerical investigations indicate that, for $\rho \in [0, 0.6)$, $\theta^* = \pi$ and thus

$$\gamma_{\max} = (1 + \rho^2)/\{4(1 + \rho)\}, \tag{30}$$

and that, for $\rho \in [0.6, 1)$,

$$\gamma_{\max} \simeq -0.0123 + 1.2986\rho - 1.9212\rho^2 + 0.6350\rho^3, \tag{31}$$

to close approximation. Figure S10 plots the relation between ρ and $\gamma_{\max}$ obtained using restriction (29). Equation (30) reproduces its upper curve, whilst its lower curve, for $\rho \in [0.6, 1)$, is visually indistinguishable from that obtained using Equation (31).

Figure S11 displays contour plots of density (28) for $q = 1$ and six combinations of (ρ, γ). The parameter ρ controls the strength of dependence between θ_1 and θ_2 (for this version of the model, $\rho_R = q\rho^2$), the number of modes and antimodes, and the shape of the density around the main mode and antimode(s). For small values of ρ the density has multiple modes and antimodes, whilst for large values of ρ it has a single mode and three antimodes. As γ increases, the value of the density at $(0, 0)$ increases, the value at (π, π) decreases and the contours around the mode(s) become tighter.

4 Simulation and model fitting

In this section we consider simulation, parameter estimation and testing for independence and goodness-of-fit, first for the circulas introduced in Section 3 and then for the bivariate circular models with density (2) generated from them.

4.1 Circula models

To simulate a random vector from $c(\theta_1, \theta_2)$, first simulate θ_A from the circular uniform distribution, where A is either 1 or 2, and then use it to simulate θ_B from the conditional distribution of $\Theta_B|\theta_A$ where $B = 3 - A$. Efficient algorithms for doing the latter may exist or it might be necessary to apply the inverse probability integral transform, either using a closed-form expression or numerically. See Kato & Jones (2015) for the relevant details for their model. Of course, A can be chosen so that the distribution of $\Theta_B|\theta_A$ is the easier of the two conditional distributions from which to simulate. Alternatively, vectors from $c(\theta_1, \theta_2)$ can be simulated using a simple acceptance-rejection approach based on uniform random simulation in $[-\pi, \pi]^2 \times [0, c_{\max}]$, where $c_{\max}$ denotes the maximum value of c.

Let $\{(\theta_{1k}, \theta_{2k}), k = 1, ..., n\}$ denote an i.i.d. sample of random vectors from a circular density c where n from here on denotes sample size. If, as will generally be the case in practice, the distributional form of c is unknown, and n is moderate to large, an inspection of a scatterplot of the data will usually be sufficient to identify q and can provide insight as to the form of the underlying generating circula density. Also, a level plot of the absolute values of the empirical Fourier coefficients, $\tilde{\phi}(r, -qs) = \frac{1}{n}\sum_{k=1}^{n} e^{i(r\theta_{1k} - qs\theta_{2k})}$, for $r, s = 1, 2, ..., 6$, say, can be inspected for patterns, like those in Figure 1, indicative of the structure of the Fourier coefficients of the underlying circula density. Level plots of this type are available in R's `lattice` package and are illustrated in Section 5. Note that computation of the empirical Fourier coefficients is extremely fast. In practice, a range of potential c's might be explored and the best fitting model established using information criteria and goodness-of-fit testing (see below).

All of the circula densities in Section 3 have, by design, relatively simple closed-form expressions involving no computationally demanding normalizing constants. As a consequence, their log-likelihood functions are easy to compute. However, maximum likelihood (ML) estimation must invariably be conducted numerically, constrained optimization techniques generally being required so as to respect the constraints on the parameters. We have made successful use of R's `nlminb` routine, which allows for box constraints, when performing ML estimation. Method of moments (MM) estimation can be used to obtain potential starting values from

which to begin the optimization process. For Class 5, for example, $\gamma = ||\phi(1,-1)| - |\phi(1,1)||$, $q = \text{sgn}(|\phi(1,-1)| - |\phi(1,1)|)$, $\rho = |\phi(2,-2q)|/\gamma$ and $\lambda = |\phi(1,-2q)|/\gamma$. Moment estimates of the parameters are calculated sequentially from these equations, substituting the $\phi(r,s)$ by their sample analogues $\tilde{\phi}(r,s)$. For the calculation of confidence intervals and regions we favour the use of profile log-likelihood methods based on standard asymptotic chi-squared theory for likelihood ratios because they can be easily programmed to incorporate parameter constraints. Confidence regions calculated by inverting the information matrix will generally not respect such constraints.

Independence can be tested for using the permutation approach proposed in Section 3.3 of Kato & Pewsey (2015) based on the likelihood ratio statistic $L = 2(\ell_1 - \ell_0)$, where ℓ_1 denotes the maximum of the log-likelihood for the chosen circula and $\ell_0 = -2n\log(2\pi)$ the log-likelihood of the circular uniform random variables Θ_1 and Θ_2 under independence. Alternatively, at least for the unimodal circulas considered here, for which $\rho_R = 0$ implies the circular uniform marginals are independent, the more easily calculated moment estimate of ρ_R, $\tilde{\rho}_R = |\tilde{\phi}(1,-1)| - |\tilde{\phi}(1,1)|$, can be used as the test statistic.

The distribution function of a circula can be represented as $C(\theta_1,\theta_2) = C_{1|2}(\theta_1|\theta_2)\ C_2(\theta_2)$, where C_2 is the marginal distribution function of Θ_2, i.e. the circular uniform distribution function, and $C_{1|2}$ is the conditional distribution function of $\Theta_1|\Theta_2 = \theta_2$. The latter is defined as

$$C_{1|2}(\theta_1|\theta_2) = \int_0^{\theta_1} c_{1|2}(\phi|\theta_2)\mathrm{d}\phi, \tag{32}$$

where $c_{1|2}$ denotes the conditional density of $\Theta_1|\Theta_2 = \theta_2$, and, for simplicity, it is assumed that $\theta_1 = \theta_1 \bmod(2\pi) \in [0, 2\pi)$. Then $U_1 = C_2(\Theta_2) = \Theta_2/2\pi$ and $U_2 = C_{1|2}(\Theta_1|\Theta_2 = \theta_2)$ are independent U(0, 1) random variables. As Θ_1 and Θ_2 are circular uniform random variables, $\{(\theta_{2k}, 2\pi C_{1|2}(\theta_{1k}|\theta_{2k})),\ k = 1, ..., n\}$ will be a uniformly distributed sample on the torus. The same applies to $\{(\theta_{1k}, 2\pi C_{2|1}(\theta_{2k}|\theta_{1k})), k = 1, ..., n\}$. When the parameters of the conditional distributions are estimated, we obtain the estimated conditional distribution functions $\hat{C}_{1|2}$ and $\hat{C}_{2|1}$, and $\{(\theta_{2k}, 2\pi\hat{C}_{1|2}(\theta_{1k}|\theta_{2k})),\ k = 1, ..., n\}$ and $\{(\theta_{1k}, 2\pi\hat{C}_{2|1}(\theta_{2k}|\theta_{1k})),\ k = 1, ..., n\}$ will be samples of pseudo-uniform vectors on the torus. In order to test the goodness-of-fit of the fitted circula density, one can apply the obvious adaptation of the parametric bootstrap based approach described in Section 3.2 of Pewsey & Kato (2016) to the two pseudo-uniform samples.

4.2 Bivariate circular models

The shapes of the bivariate circular densities obtained using density (2) depend heavily on the reference points from which the marginal densities are integrated in the definitions of the marginal distribution functions F_1 and F_2. Traditionally, the reference point used has been the origin, 0. However, for this choice, changes in the location parameters of the marginal distributions result in shape changes, not just location shifts, in the densities obtained using Equation (2). To avoid such shape changes, we define $F_j(\theta), j = 1, 2$, as

$$F_j(\theta) = \int_{\omega_j}^{\theta} f_j(\psi)\mathrm{d}\psi, \qquad \omega_j \leq \psi < \omega_j + 2\pi,$$

where ω_j denotes the antimode of the circular density f_j.

A feature of copulas often referred to as being appealing (Mikosch, 2006), which from an inspection of (2) is clearly shared by circulas, is that the bivariate densities derived from them

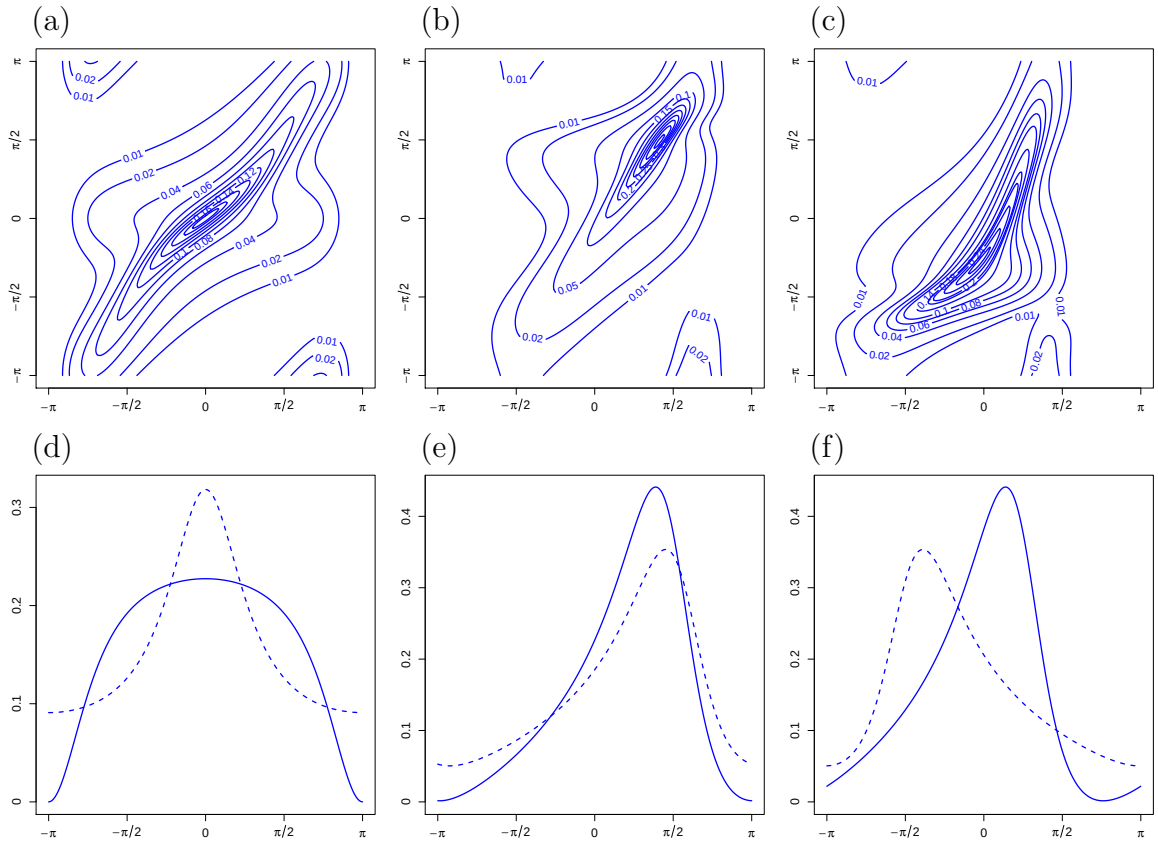

Figure 5: Planar contour plots of the bivariate circular density (2) when c is the circula density depicted in Figure 4(b) and the marginal densities are the Kato & Jones (2015) marginal densities depicted in the panels immediately below each of them. The values of the parameters of the marginal Kato–Jones densities for Θ_1 and Θ_2, $(\mu_1, \gamma_1, \bar{\alpha}_{2,1}, \bar{\beta}_{2,1})$ and $(\mu_2, \gamma_2, \bar{\alpha}_{2,2}, \bar{\beta}_{2,2})$, respectively, are: (d) $(0, 0.3, -0.12, 0)$ and $(0, 0.3, 0.12, 0)$; (e) $(\pi/4, 0.58, 0.164, 0.164)$ and $(\pi/4+0.2, 0.4, 0.113, 0.113)$; (f) $(0, 0.58, 0.164, 0.164)$ and $(-\pi/4, 0.4, 0.113, -0.113)$. In panels (d)–(f) the solid (dashed) curve represents the marginal density of Θ_1 (Θ_2).

can be decomposed into their marginal densities and a dependence regulating copula or circula density. However, as the decomposition arises from the product of such densities, it is not immediately obvious what shapes the resulting bivariate densities might adopt. As the panels in the left-hand column of Figure 3 of Jones *et al.* (2015) illustrate, the shapes of bivariate circular densities generated using the circula density of Class 2 can be far from what one might expect even for the seemingly innocuous choice of von Mises marginal and binding densities. Panels (a)–(c) of Figure 5 illustrate the flexibility of density (2) when the circula density in Figure 4(b) is used together with the pairs of marginal Kato & Jones (2015) densities depicted in panels (d)–(f) immediately below them. In Figure 5(d) both marginal densities are symmetric, that for Θ_1 (Θ_2) being relatively flat-topped (peaked). Both densities in Figure 5(e) are skewed in the same direction, whilst, in Figure 5(f), they are skewed in opposite directions.

Set A to whichever one of 1 and 2 proves more convenient again, and $B = 3 - A$ to the other. A random vector from (2) can be simulated by first simulating θ_A from the marginal distribution of Θ_A with distribution function F_A. The circular uniform random variate $\psi_A = 2\pi F_A(\theta_A)$ is then calculated and another circular uniform random variate, $\psi_B = 2\pi F_B(\theta_B)$, simulated

from $c_{B|A}(\psi_B|\psi_A)$. Finally, $\theta_B = F_B^{-1}(\psi_B/(2\pi))$. Alternatively, a simple acceptance-rejection approach based directly on (2) might be employed.

We advocate the following ML approach to fitting densities of the form (2). If the distributional forms of the marginal densities, f_1 and f_2, are not known or specified beforehand, histograms and/or kernel density estimates of the data on each variable can provide insight into potential underlying classes for them. Once potential distributional forms for f_1 and f_2 have been specified, their parameters are initially estimated separately using ML. Denoting the marginal distribution functions corresponding to those parameter estimates by $\hat{F}_1^*$ and $\hat{F}_2^*$, next the 'pseudo-sample' $\{(2\pi\hat{F}_1^*(\theta_{1k}), 2\pi\hat{F}_2^*(\theta_{2k})), k = 1, ..., n\}$ is computed. For moderate to large n, a scatterplot of the pseudo-sample provides a visual aid in the identification of a potential model for the circula density c. In addition, the inspection of a level plot of the absolute values of the empirical Fourier coefficients for the pseudo-sample will usually provide insight into the structure of the Fourier coefficients of the underlying circula density when the latter can be assumed to be pointwise symmetric. We recommend the inspection of the absolute values of the empirical Fourier coefficients because, unlike their real parts, they are invariant to location shifts. The imaginary parts of the empirical Fourier coefficients are not invariant to such location shifts either, and non-random patterns in level plots of them are indicative of potential location misspecification arising from the estimation of the parameters of the marginal distributions alone. This is a consequence of the fact that if a random vector (Θ_1, Θ_2) follows a circula with real Fourier coefficients $\phi(r, s)$ for any (r, s) then the random vector $(\Theta_1 + \mu_1, \Theta_2 + \mu_2)$ has Fourier coefficients $e^{i(r\mu_1+s\mu_2)}\phi(r, s)$. In the next estimation step, ML estimation as described in Section 4.1 is then applied to the pseudo-sample to obtain estimates of the parameters of the chosen circula density c. Finally, the estimates from the previous two stages are used as starting values in the maximization of the full log-likelihood function derived from (2).

The independence of Θ_1 and Θ_2 can be tested for using the permutation approach described in Section 4.1, but now with ℓ_0 denoting the maximum of the log-likelihood under independence, $\sum_{k=1}^n \log f_1(\theta_{1k}) + \sum_{k=1}^n \log f_2(\theta_{2k})$. An alternative approach if c is one of our unimodal circulas is to apply the permutation test based on $\tilde{\rho}_R$, also described in Section 4.1, to the pseudo-sample from c after maximization of the full log-likelihood function. We denote the latter by $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{F}_2(\theta_{2k})), k = 1, ..., n\}$.

Proceeding as in Section 4.1, $\{(2\pi F_2(\theta_{2k}), 2\pi C_{1|2}(2\pi F_1(\theta_{1k})|2\pi F_2(\theta_{2k}))), k = 1, ..., n\}$ and $\{(2\pi F_1(\theta_{1k}), 2\pi C_{2|1}(2\pi F_2(\theta_{2k})|2\pi F_1(\theta_{1k}))), k = 1, ..., n\}$ are two uniformly distributed samples on the torus. When the parameters of (2) are estimated using the approach described above, the samples $\{(2\pi\hat{F}_2(\theta_{2k}), 2\pi\hat{C}_{1|2}(2\pi\hat{F}_1(\theta_{1k})|2\pi\hat{F}_2(\theta_{2k}))), k = 1, ..., n\}$ and $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{C}_{1|2}(2\pi\hat{F}_2(\theta_{2k})|2\pi\hat{F}_1(\theta_{1k}))), k = 1, ..., n\}$ will be pseudo-uniformly distributed on the torus, and the goodness-of-fit of the fitted circular density can be tested for using the parametric bootstrap approach referred to at the end of Section 4.1.

5 Illustrative applications

In this section we present analyses of two toroidal datasets that illustrate the use of the proposed Fourier series based circulas, the bivariate circular models derived from them and the methodology described in Section 4.

Table 1: Parameters, true parameter values used to simulate the data in Figure 6(a) (TV), MM and ML estimates of the parameters of Kato & Jones (2015) distributions fitted to each marginal data set (MM_M and ML_M), MM and ML estimates of the parameters of the Class 5 circula with $q = 1$ fitted to the values of $\{(2\pi\hat{F}_1^*(\theta_{1k}), 2\pi\hat{F}_2^*(\theta_{2k})), k = 1, ..., n\}$ (MM_C and ML_C), and ML estimates of the parameters of the full model with $q = 1$ used to simulate the data (ML_F).

Parameter	TV	MM_M	ML_M	MM_C	ML_C	ML_F
μ_1	0	−0.07	−0.07			−0.04
γ_1	0.58	0.57	0.57			0.58
$\bar{\alpha}_{2,1}$	0.16	0.13	0.14			0.15
$\bar{\beta}_{2,1}$	0.16	0.16	0.16			0.16
μ_2	−0.79	−0.84	−0.84			−0.84
γ_2	0.40	0.43	0.43			0.42
$\bar{\alpha}_{2,2}$	0.11	0.14	0.13			0.14
$\bar{\beta}_{2,2}$	−0.11	−0.12	−0.13			−0.13
ρ	0.60			0.61	0.60	0.60
λ	0.40			0.41	0.42	0.42
γ	0.48			0.45	0.46	0.47

5.1 Simulated data

We first consider the random sample of size $n = 2000$ plotted in Figure 6(a) simulated from the bivariate circular density depicted in Figure 5(c). In our analysis we assume that c, f_1 and f_2 are correctly specified as the Class 5 circula density and Kato & Jones (2015) marginal densities, respectively, but, of course, that their parameters are unknown.

Proceeding as described in Section 4.2, the MM and ML estimates of the parameters of the assumed Kato & Jones (2015) marginal densities are presented in the third and fourth columns of Table 1.

Figure 6(b) provides a scatterplot of the pseudo-values $\{(2\pi\hat{F}_1^*(\theta_{1k}), 2\pi\hat{F}_2^*(\theta_{2k})), k = 1, ..., n\}$ calculated using the ML estimates of the parameters of the two marginal Kato & Jones (2015) densities. Panels (c), (d) and (e) of the same figure are level plots of the real parts, imaginary parts and absolute values, respectively, of the empirical Fourier coefficients for those pseudo-values. The imaginary parts of the empirical Fourier coefficients, in panel (d), are all small in absolute value and appear to manifest a random pattern. The patterns in panels (c) and (e) are consistent with the triangular structure of the Fourier coefficients of the underlying Class 5 circula represented graphically in the central panel of Figure 1. In practice, when the distributional form of c is unknown, such plots provide insight into potential choices for c.

The fifth and sixth columns of Table 1 present the MM and ML estimates of the parameters ρ, λ and γ of the Class 5 circula with $q = 1$ fitted to the values of $\{(2\pi\hat{F}_1^*(\theta_{1k}), 2\pi\hat{F}_2^*(\theta_{2k})), k = 1, ..., n\}$.

Finally, the ML estimates obtained for the Kato & Jones (2015) marginal densities and the Class 5 circula were used as starting values in the maximization of the full log-likelihood. The ML estimates obtained for the full model are denoted by ML_F in the seventh column of Table 1. Figure 6(f) provides a scatterplot of the values of $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{F}_2(\theta_{2k})), k = 1, ..., n\}$ corresponding to those ML estimates. (It is very similar to Figure 6(b).) The value of $\tilde{\rho}_R$ for

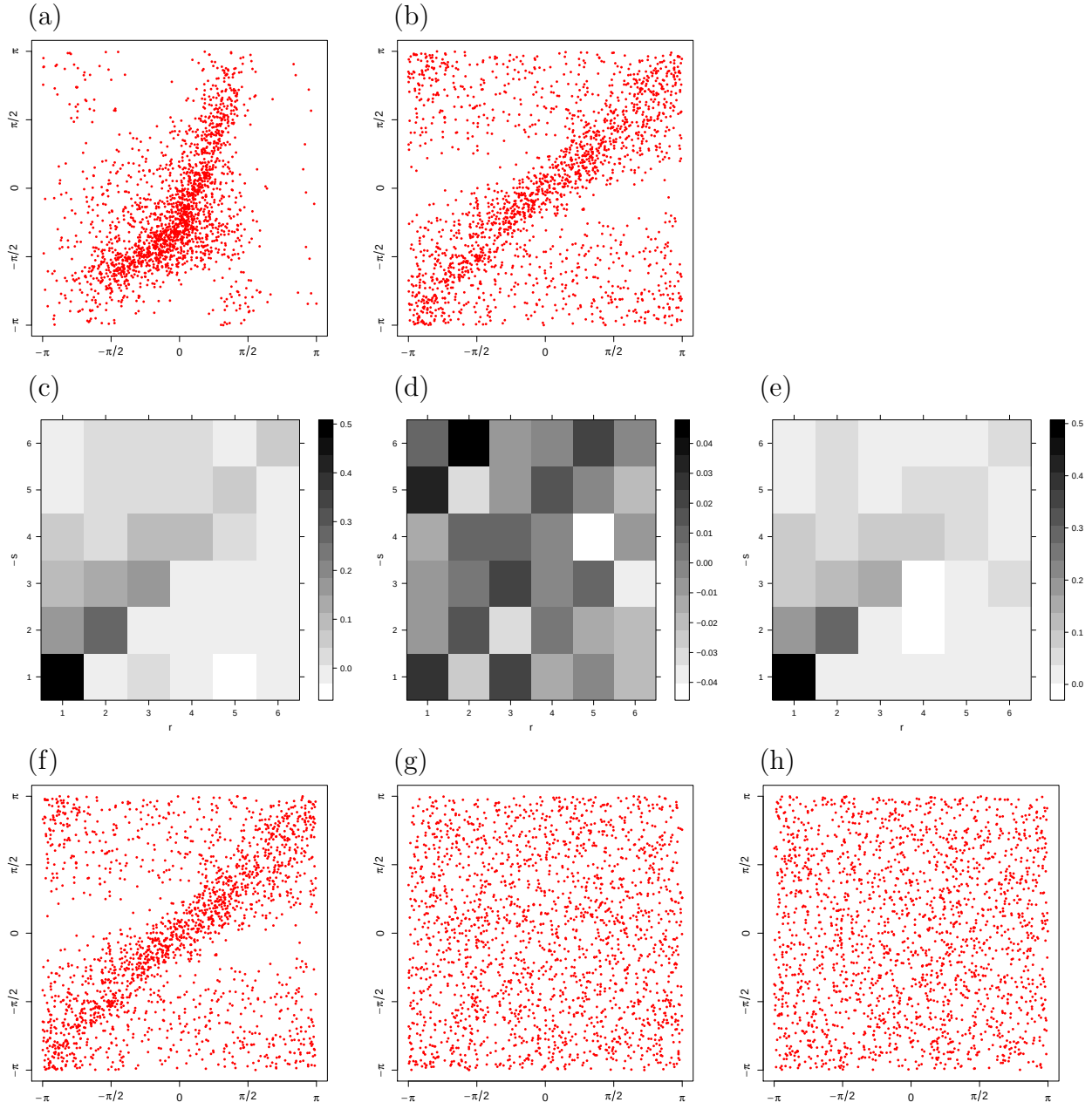

Figure 6: Planar scatterplots of: (a) a random sample of size $n = 2000$ simulated from the bivariate circular density depicted in Figure 5(c), $\{(\theta_{1k}, \theta_{2k}), k = 1, \ldots, n\}$; (b) $\{(2\pi\hat{F}_1^*(\theta_{1k}), 2\pi\hat{F}_2^*(\theta_{2k})), k = 1, ..., n\}$; (f) $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{F}_2(\theta_{2k})), k = 1, ..., n\}$; (g) $\{(2\pi\hat{F}_2(\theta_{2k}), 2\pi\hat{C}_{1|2}(2\pi\hat{F}_1(\theta_{1k})|2\pi\hat{F}_2(\theta_{2k}))), k = 1, ..., n\}$; (h) $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{C}_{2|1}(2\pi\hat{F}_2(\theta_{2k})|2\pi\hat{F}_1(\theta_{1k}))), k = 1, ..., n\}$. Panels (c), (d) and (e) are level plots of the real parts, imaginary parts and absolute values, respectively, of the empirical Fourier coefficients of the pseudo-values in panel (b) for $q = 1$.

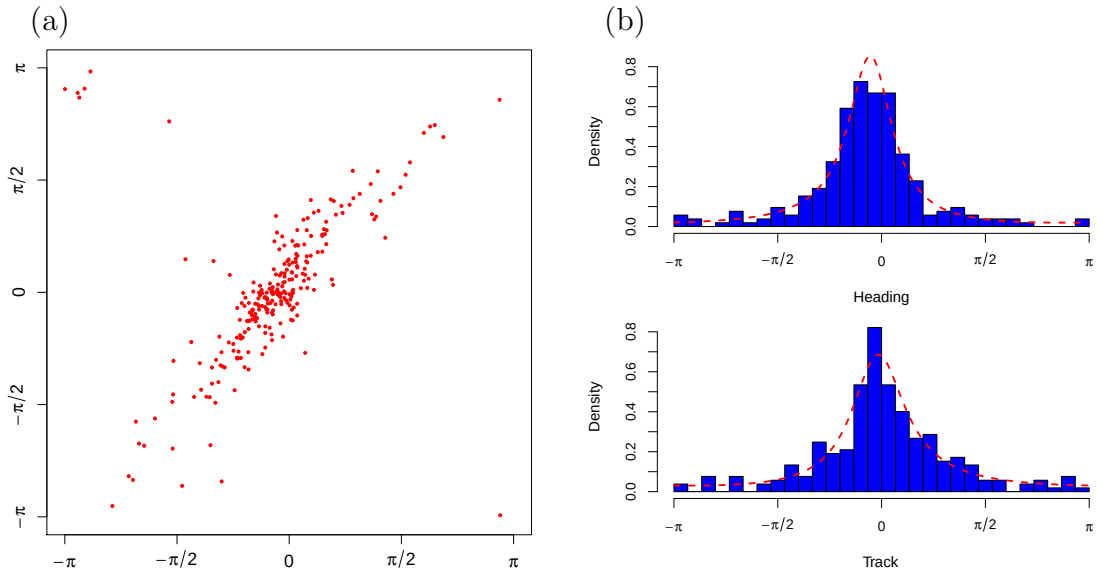

Figure 7: (a) Planar scatterplot of the track versus heading measurements of $n = 250$ migrating birds, $\{(\theta_{1k}, \theta_{2k}), k = 1, \ldots, n\}$, (b) Histograms of the heading and track measurements with ML fitted Kato & Jones (2015) densities superimposed.

those values is 0.451 and the p-value of the test of independence based on $\tilde{\rho}_R$ using 999 random permutations of $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{F}_2(\theta_{2k})), k = 1, ..., n\}$ was 0.001. Hence, independence of Θ_1 and Θ_2 is emphatically rejected by the test.

Panels (g) and (h) of Figure 6 are scatterplots of $\{(2\pi\hat{F}_2(\theta_{2k}), 2\pi\hat{C}_{1|2}(2\pi\hat{F}_1(\theta_{1k})|2\pi\hat{F}_2(\theta_{2k}))), k = 1, ..., n\}$ and $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{C}_{2|1}(2\pi\hat{F}_2(\theta_{2k})|2\pi\hat{F}_1(\theta_{1k}))), k = 1, ..., n\}$, respectively. Applying the adaptation of the goodness-of-fit testing approach of Pewsey & Kato (2016) to them using the Bingham type test statistic of Wellner (1979) and $B = 99$ parametric bootstrap samples, the value of the test statistic was 3.91 and the estimated p-value 0.23. Hence, there is no significant evidence to reject the ML fitted model as the underlying distribution from which the data were generated. From inspection of Table 1, the ML estimates of the parameters of that model are very close to the true values used in the simulation of the data.

5.2 Bird migration data

Our second analysis is of a random subsample of size $n = 250$ drawn from a much larger sample of 5916 heading and track measurements of migrating birds, the primary motivation for the subsampling being to illustrate the effectiveness of our methodology for a relatively small-sized data set. The data were collected using tracking radars located in the Negev Highlands of southern Israel during April of 1992 and are reported in Liechti & Bruderer (1995). The *heading*, θ_1, of a bird is the direction of its body axis during flight, whereas the *track*, θ_2, is its resulting flight direction. Both measurements were recorded to the nearest degree in $[0, 359]$. Figure 7(a) presents a planar scatterplot of the data converted to radians in $[-\pi, \pi)$, and Figure 7(b) histograms for the two marginal variables. The histograms are both close to unimodal and perhaps somewhat skew, so we investigated the fit of the flexible Kato & Jones (2015) family of unimodal circular distributions to them. The second and third columns of Table 2 contain the MM and ML point estimates of the parameters for those fits. The densities for the ML fits are

Table 2: Parameters, MM and ML estimates of the parameters of Kato & Jones (2015) distributions fitted to the heading and track measurements ($\mathrm{MM_M}$ and $\mathrm{ML_M}$), ML estimates of the parameters of a Class 8 circula with $q = 1$ fitted to the values of $\{(2\pi\hat{F}_1^*(\theta_{1k}), 2\pi\hat{F}_2^*(\theta_{2k})), k = 1, ..., n\}$ ($\mathrm{ML_C}$), and ML estimates of the parameters of the full model ($\mathrm{ML_F}$).

Parameter	$\mathrm{MM_M}$	$\mathrm{ML_M}$	$\mathrm{ML_C}$	$\mathrm{ML_F}$
μ_1	-0.21	-0.21		-0.15
γ_1	0.73	0.73		0.68
$\bar{\alpha}_{2,1}$	0.49	0.48		0.38
$\bar{\beta}_{2,1}$	0.02	0.02		-0.06
μ_2	0.01	0.01		0.07
γ_2	0.65	0.65		0.60
$\bar{\alpha}_{2,2}$	0.38	0.39		0.27
$\bar{\beta}_{2,2}$	-0.02	-0.03		-0.04
ρ			0.81	0.85
γ			0.11	0.09

superimposed upon the histograms in Figure 7(b) and appear to model the data rather well.

The pseudo-sample $\{(2\pi\hat{F}_1^*(\theta_{1k}), 2\pi\hat{F}_2^*(\theta_{2k})), k = 1, ..., n\}$ is portrayed in Figure 8(a). We immediately deduce from its structure that $q = 1$. Panels (b) and (c) of the same figure are level plots of the imaginary parts and absolute values, respectively, of the empirical Fourier coefficients for that pseudo-sample. The larger imaginary parts form a diagonal pattern consistent with the centre of the pseudo-values not being $(0, 0)$, as the cross in Figure 8(a) corroborates. The absolute values of the empirical Fourier coefficients in Figure 8(c) strongly suggest some form of diagonal pattern, and consideration of the numerical values of the coefficients located on the main diagonal and the two diagonals immediately above and below it suggested that the Class 8 circula with density (28) might be a possible model for the underlying c. The fourth column of Table 2 contains the ML point estimates of the parameters of that model fitted to the pseudo-sample.

Using the ML estimates in the third and fourth columns of Table 2 as starting values, we obtained the ML estimates of all 10 parameters of the full model presented in the fifth column of Table 2. Figure 8(d) provides a scatterplot of the $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{F}_2(\theta_{2k})), k = 1, ..., n\}$ pseudo-sample corresponding to that fit. Note how the centre of this pseudo-sample appears to be closer to $(0, 0)$. In Figure 8(e) the same pseudo-sample is superimposed on the density of the Class 8 circula corresponding to the full model fit.

A scatterplot of the original data is superimposed on the ML fitted density for the full model in Figure 8(f). The fitted model appears to describe the distribution of the majority of the observations, situated around that part of the diagonal in the neighbourhood of the origin, reasonably well. More formally, panels (g) and (h) of Figure 8 provide scatterplots of $\{(2\pi\hat{F}_2(\theta_{2k}), 2\pi\hat{C}_{1|2}(2\pi\hat{F}_1(\theta_{1k})|2\pi\hat{F}_2(\theta_{2k}))),\ k = 1, ..., n\}$ and $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{C}_{2|1}(2\pi\hat{F}_2(\theta_{2k})|$ $2\pi\hat{F}_1(\theta_{1k}))), k = 1, ..., n\}$, respectively. The distributions of the points within those scatterplots appear to be reasonably consistent with toroidal uniformity. Applying the adaptation of the goodness-of-fit testing approach of Pewsey & Kato (2016) to them using the Bingham type test statistic of Wellner (1979) and $B = 99$ parametric bootstrap samples, the value of the test statistic was 5.64 and the estimated p-value 0.06. At least at the 5% significance level, then,

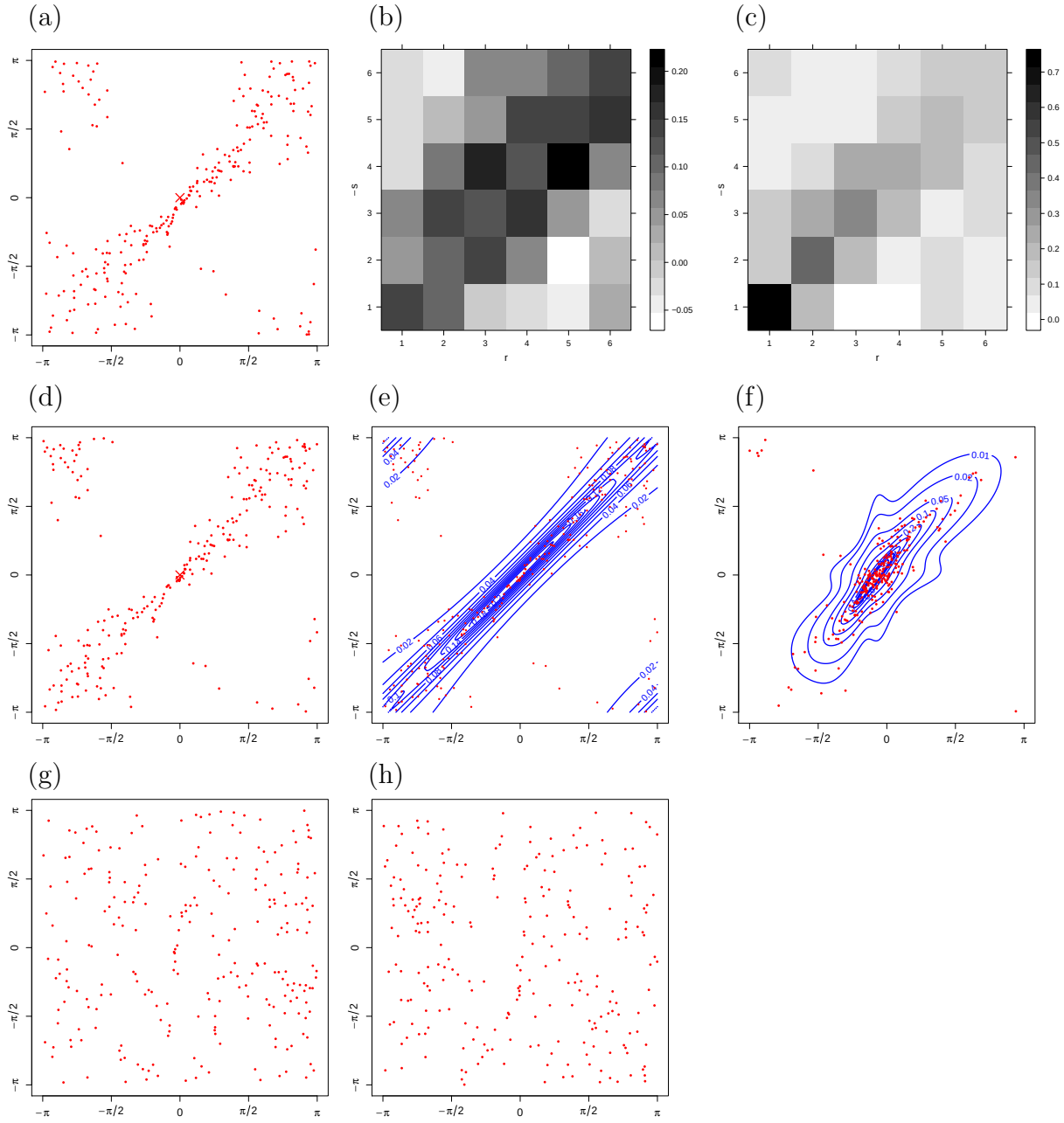

Figure 8: Planar scatterplots of: (a) $\{(2\pi\hat{F}_1^*(\theta_{1k}), 2\pi\hat{F}_2^*(\theta_{2k})), k = 1, ..., n\}$; (d) $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{F}_2(\theta_{2k})), k = 1, ..., n\}$; (e) $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{F}_2(\theta_{2k})), k = 1, ..., n\}$ superimposed on the Class 8 density corresponding to the full ML solution; (f) the original data $\{(\theta_{1k}, \theta_{2k}), k = 1, \ldots, n\}$ superimposed on the ML fitted density for the full model; (g) $\{(2\pi\hat{F}_2(\theta_{2k}), 2\pi\hat{C}_{1|2}(2\pi\hat{F}_1(\theta_{1k})|2\pi\hat{F}_2(\theta_{2k}))), k = 1, ..., n\}$; (h) $\{(2\pi\hat{F}_1(\theta_{1k}), 2\pi\hat{C}_{2|1}(2\pi\hat{F}_2(\theta_{2k})|2\pi\hat{F}_1(\theta_{1k}))), k = 1, ..., n\}$. The crosses in panels (a) and (d) identify $(0, 0)$. Panels (b) and (c) are level plots of the imaginary parts and absolute values, respectively, of the empirical Fourier coefficients of the pseudo-values in panel (a) for $q = 1$.

there is no significant evidence to reject the ML fitted distribution as a potential model for the data.

6 Discussion

As the many classes of circulas considered in Section 3 illustrate, the proposed general Fourier series construction provides a means of generating circula densities with simple closed-form expressions and wide-ranging distributional shapes. Moreover, when combined with the Kato & Jones (2015) family of circular distributions through Equation (2) they provide highly flexible models for bivariate circular data.

Stationary Markov models for circular time series can be defined from those models using an analogous approach to that of Wehrly & Johnson (1980). The flexibility of the circula models can be made even greater by allowing the Fourier series coefficients to be complex. The main effect of such an extension is to skew the circula distributions in various ways. Also, in principle, the advocated approach for bivariate circular data can be extended to produce d-dimensional circulas using the multivariate analogue of Equation (4) and patterns of non-zero Fourier coefficients distributed in d dimensions. Another possibility is, as mentioned in Jones *et al.* (2015), to model multivariate circular data using the circular analogues of pair copulas. Analogues of copulas on other compact Riemannian manifolds have been considered in Jupp (2015).

The approach we advocate should not be confused with those of Pertsemlidis et al. (2005) and Fernández-Durán (2007). Those authors make use of the Wehrly & Johnson (1980) circula combined with circular densities obtained from truncated Fourier series and nonnegative truncated Fourier series, respectively. In principle, such nonparametric marginal circular densities could be used together with the circulas proposed here to generate alternative models for toroidal data. We prefer to combine our circulas with Kato & Jones (2015) circular densities because of their parametric flexibility, unimodality and ease of interpretation. Multimodal toroidal data can be modelled using mixtures of the resulting bivariate circular models.

We stress again that our approach allows the use of level plots of the absolute values of empirical Fourier coefficients as a highly successful model identification tool: such plots prove to be more easily understood and used to suggest appropriate models than contour plots of estimated copula densities. Our Fourier-based approach also opens up the possibility of developing a fully nonparametric approach to copula estimation based on doubly truncated bivariate Fourier series, but this is beyond the scope of the current paper.

Alternative general approaches to obtaining bivariate circular distributions include wrapping and projecting, as in Jammalamadaka & Sarma (1988) and Saw (1983), respectively. Unlike the bivariate circular densities proposed here, those obtained using wrapping and projecting generally suffer from having no closed-form expression or are highly convoluted.

Acknowledgements

Financial support for this work was received by Shogo Kato from JSPS KAKENHI in the form of grants JP25400218 and JP17K05379 and by Arthur Pewsey from the Junta de Extremadura and the European Union in the form of grant GR15013.

Appendix

Proof of Theorem 1

For $m, n \in \mathbb{Z}$, we have

$$E\left[e^{i(m\Theta_1+n\Theta_2)}\right] = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{i(m\theta_1+n\theta_2)}\frac{1}{4\pi^2}\sum_{r,s=-\infty}^{\infty}\phi(r,s)e^{-i(r\theta_1+s\theta_2)}\mathrm{d}\theta_1\mathrm{d}\theta_2$$
$$= \frac{1}{4\pi^2}\sum_{r,s=-\infty}^{\infty}\phi(r,s)\left\{\left(\int_{-\pi}^{\pi}e^{i(m-r)\theta_1}\mathrm{d}\theta_1\right)\times\left(\int_{-\pi}^{\pi}e^{i(n-s)\theta_2}\mathrm{d}\theta_2\right)\right\}.$$

As

$$\int_{-\pi}^{\pi} e^{it\theta}\mathrm{d}\theta = \begin{cases} 2\pi, & t=0, \\ 0, & t\in\mathbb{Z}^+\cup\mathbb{Z}^-, \end{cases} \tag{33}$$

it follows that

$$E\left[e^{i(m\Theta_1+n\Theta_2)}\right] = \frac{1}{4\pi^2}4\pi^2\phi(m,n) = \phi(m,n).$$

Proof of Theorem 2

First we show that if $\phi(m,n)$ is given by (6), then the density (4) is a circula. The marginal density $f(\theta_1)$ of the distribution (4) can be expressed as

$$f(\theta_1) = \int_{-\pi}^{\pi} f(\theta_1,\theta_2)\mathrm{d}\theta_2 = \frac{1}{4\pi^2}\sum_{m,n=-\infty}^{\infty}\phi(m,n)e^{-im\theta_1}\int_{-\pi}^{\pi}e^{-in\theta_2}\mathrm{d}\theta_2.$$

Then it follows from the assumption (6) and Equation (33) that

$$f(\theta_1) = \frac{1}{2\pi}\sum_{m=-\infty}^{\infty}\phi(m,0)e^{-im\theta_1} = \frac{1}{2\pi}.$$

Similarly it can be seen that $f(\theta_2) = 1/(2\pi)$. Therefore the density (4) is a circula.

Next we prove that if a density in family (4) is a circula, then $\phi(m,n)$ is of the form (6). Assume that (Θ_1,Θ_2) has the density (4). Since the marginal of Θ_1 is the circular uniform, it follows from Theorem 1 that

$$\phi(m,0) = E(e^{im\Theta_1}) = \begin{cases} 1, & m=0, \\ 0, & m\neq 0, \end{cases} \qquad m\in\mathbb{Z}.$$

In a similar manner, we can see that $\phi(0,n)$ is given by the second equation of (6).

References

Fernández-Durán, J. J. (2007) Models for circular-linear and circular-circular data constructed from circular distributions based on nonnegative trigonometric sums. *Biometrics*, **63**, 579–585.

Fisher, N. I. and Lee, A. J. (1983) A correlation coefficient for circular data. *Biometrika*, **70**, 327–332.

Jammalamadaka, S. R. and Sarma, Y. R. (1988) A correlation coefficient for angular variables. In K. Matusita (Ed.), *Statistical Theory and Data Analysis II* (pp. 349–364). Elsevier, Amsterdam.

Jones, M.C., Pewsey, A. and Kato, S. (2015) On a class of circulas: copulas for circular distributions. *Annals of the Institute of Statistical Mathematics*, **67**, 843–862.

Jupp, P. E. (2015) Copulae on products of compact Riemannian manifolds. *Journal of Multivariate Analysis*, **140**, 92–98.

Kato, S. and Jones, M. C. (2015) A tractable and interpretable four-parameter family of unimodal distributions on the circle. *Biometrika*, **102**, 181–190.

Kato, S. and Pewsey, A. (2015) A Möbius transformation-induced distribution on the torus. *Biometrika*, **102**, 359–370.

Liechti, F. and Bruderer, B. (1995) Direction, speed and composition of nocturnal bird migration in the south of Israel. *Israel Journal of Zoology*, **41**, 501–515.

Mardia, K. V. and Jupp, P. E. (1999) *Directional Statistics*. Wiley, New York.

Mikosch, T. (2006) Copulas: tales and facts. *Extremes*, **9**, 3–20.

Pertsemlidis, A., Zelinka, J., Fondon, J. W., Henderson, R. K. and Otwinowski, Z. (2005) Bayesian statistical studies of the Ramachandran distribution. *Statistical Applications in Genetics and Molecular Biology*, **4**, 35.

Pewsey, A. and Kato, S. (2016) Parametric bootstrap goodness-of-fit testing for Wehrly–Johnson bivariate circular distributions. *Statistics & Computing*, **26**, 1307–1317.

Pewsey, A., Neuhäuser, M. and Ruxton, G. D. (2013) *Circular Statistics in R*. Oxford University Press, Oxford.

Rivest, L.-P. (1982) Some statistical methods for bivariate circular data. *Journal of the Royal Statistical Society Series B*, **44**, 81–90.

Rivest, L.-P. (1997) A decentred predictor for circular-circular regression. *Biometrika*, **84**, 717–726.

Saw, J. G. (1983) Dependent unit vectors. *Biometrika*, **70**, 665–671.

Sklar, A. (1959) Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, **8**, 229–231.

Wehrly, T. and Johnson, R. A. (1980) Bivariate models for dependence of angular observations and a related Markov process. *Biometrika*, **66**, 255–256.

Wellner, J. A. (1979) Permutation tests for directional data. *Annals of Statistics*, **7**, 929–943.