*Biometrika* (2023), **xx**, x, *pp*. 1–5
*Printed in Great Britain*

# Introduction of general distributions on sphere and torus in view of time series spectra

BY MASANOBU TANIGUCHI
*Waseda University*
taniguchi@waseda.jp

AND YUJIE XUE
*The Institute of Statistical Mathmatics*
xue.yujie@ism.ac.jp

SUMMARY

There are various fields where observations are taken on directions in three dimensions, e.g., sphere and torus. Here we will introduce a very general family of distributions on sphere and torus by use of time series spectra, which includes a lot of proposed classical one as special cases. Because time series spectra can be described by a lot of famous parametric models, e.g., AR, ARMA etc., we can develop the systematic model selection in this field by use of AIC, BIC, etc. Applications are very wide.

*Some key words*: Distributions on sphere and torus; Time series spectra; AR, ARMA models; Model selection.

## 1. INTRODUCTION

There are various fields where observations are taken on directions in three dimensions, e.g., molecular biology and physics. The von Mises-Fisher distribution is very fundamental with probability density function

$$f(\boldsymbol{x}) = c(\kappa, \boldsymbol{\mu})^{-1} \exp[\kappa \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{x}], \quad \boldsymbol{x} \in \mathbb{S}^2, \tag{1.1}$$

where $\mathbb{S}^2 = \{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\| = 1\}$, $\boldsymbol{\mu} \in \mathbb{S}^2$ and $c(\cdot)$ is the normalizing constant (e.g., Mardia and Jupp (2000)). The Fisher-Bingham distribution is given by density

$$f(\boldsymbol{x}) = c(\kappa, \boldsymbol{\mu}, \boldsymbol{A})^{-1} \exp[\kappa \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{x} + \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}], \tag{1.2}$$

where $\boldsymbol{A}$ is a symmetric $3 \times 3$ matrix (Mardia (1975)). For a special choice of $\boldsymbol{\mu}$ and $\boldsymbol{A}$, Kent (1982) derived a polar co-ordinates form of (1.2) by

$$f(\theta, \phi) = c(\kappa, \beta)^{-1} \exp[\kappa \cos\theta + \beta \sin^2\theta \cos 2\phi]. \tag{1.3}$$

Kato and McCullagh (2020) introduced a Cauchy family of distribution by density

$$f(\boldsymbol{x}) = c(\rho, \boldsymbol{\mu})^{-1} [\frac{1-\rho^2}{1+\rho^2 - 2\rho \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{x}}], \quad \boldsymbol{x} \in \mathbb{S}^2. \tag{1.4}$$

For modelling of torsional angles of molecules, Singh et al. (2002) introduced the following distribution on torus by

$$f(\theta, \phi) = c \exp[\kappa_1 \cos(\theta - \mu_1) + \kappa_2 \cos(\phi - \mu_2) + \lambda \sin(\theta - \mu_1) \sin(\phi - \mu_2)], \tag{1.5}$$

where $-\pi \leq \theta, \phi \leq \pi$, $\kappa_1, \kappa_2 \geq 0$, $-\infty < \lambda < \infty$, $-\pi \leq \mu_1, \mu_2 \leq \pi$ and $c$ is a normalization constant. Also, Kato and Pewsey (2015) introduced the following wrapped Cauchy type distribution by density

$$\begin{aligned} f(\theta, \phi) =& c[c_0 - c_1 \cos(\theta - \mu_1) - c_2 \cos(\phi - \mu_2) - c_3 \cos(\theta - \mu_1) \cos(\phi - \mu_2) \\ & - c_4 \sin(\theta - \mu_1) \sin(\phi - \mu_2)]^{-1}, \quad (-\pi \leq \theta, \phi \leq \pi), \end{aligned} \tag{1.6}$$

where $c_i$ and $\mu_i$ are constants.
For circular data, the wrapped Cauchy density is often used, and is defined by

$$f_W(\omega) = c \frac{1}{1 + \rho^2 + 2\rho \cos \omega}, \quad \omega \in [-\pi, \pi]. \tag{1.7}$$

Time series people understand that this is exactly the AR(1) spectral density

$$f_S(\omega) = c \frac{1}{|1 + \rho e^{i\omega}|^2}, \quad \omega \in [-\pi, \pi]. \tag{1.8}$$

Motivated by this, Taniguchi et al. (2020) introduced a very general family of joint circular distributions by a higher order spectral density

$$f_S(\omega_1, \omega_2, \ldots, \omega_n), \quad \omega_k \in [-\pi, \pi], \tag{1.9}$$

which can be decomposed to

$$\prod_{k=1}^{n} f_S(\omega_k), \tag{1.10}$$

if $\omega_1, \omega_2, \ldots, \omega_n$ are independent, where $f_S(\omega_k)$ is the spectral density of frequency $\omega_k$. Advantage of this approach is that we can introduce time series models for $f_S(\omega_k)$, i.e., AR($p$), ARMA($p, q$) models etc., then the systematic model selection in this field can be carried out.

In this paper, for distributions on sphere and torus, in view of above, we will introduce a very general family of the distributions by time series spectra, whose forms will be

$$f(\theta, \phi) = \sum_{k=-\infty}^{\infty} \{A_k e^{ik(\theta+\phi)} + B_k e^{ik(\theta-\phi)} + C_k e^{ik\theta}\}$$

for $(\theta, \phi) \in \mathbb{S}^2$. $f(\theta, \phi)$ is the sum of time series spectra, hence, we may use ARMA($p, q$) modelling etc. We can develop the model selection by use of AIC and BIC etc. Discussion on torus is similar. The applications are in various fields.

## 2. Distributions on Sphere

In this section we introduce a very general distribution on sphere in view of time series spectra. Let the 3-dim polar representation be given by

$$\begin{cases} x &= \sin\theta \cos\phi; \\ y &= \sin\theta \sin\phi; \\ z &= \cos\theta. \end{cases} \tag{2.1}$$

Substituting $\begin{cases} \cos\theta &= \frac{e^{i\theta}+e^{-i\theta}}{2} \\ \sin\theta &= \frac{e^{i\theta}-e^{-i\theta}}{2i} \end{cases}$ to (2.1), we obtain

$$\begin{cases} x &= \frac{e^{i\theta}-e^{-i\theta}}{2i} \times \frac{e^{i\phi}+e^{-i\phi}}{2} = \frac{e^{i(\theta+\phi)}-e^{-i(\theta+\phi)}+e^{i(\theta-\phi)}-e^{-i(\theta-\phi)}}{4i}; & (2.2) \\ y &= \frac{e^{i\theta}-e^{-i\theta}}{2i} \times \frac{e^{i\phi}-e^{-i\phi}}{2i} = \frac{e^{i(\theta+\phi)}+e^{-i(\theta+\phi)}-e^{i(\theta-\phi)}-e^{i(\phi-\theta)}}{-4}; & (2.3) \\ z &= \frac{e^{i\theta}+e^{-i\theta}}{2}. & (2.4) \end{cases}$$

Let $\delta_k = \begin{cases} \frac{1}{4i} & ,k=1 \\ \frac{-1}{4i} & ,k=-1 \end{cases}$, $\gamma_k = \begin{cases} -\frac{1}{4} & ,k=1 \\ -\frac{1}{4} & ,k=-1 \end{cases}$, and $\eta_k = \begin{cases} \frac{1}{2} & ,k=1 \\ \frac{1}{2} & ,k=-1 \end{cases}$. Then we can see that

$$x = \sum_{k=\pm1} \delta_k \{e^{ik(\phi+\theta)} + e^{ik(\phi-\theta)}\}; \tag{2.5}$$

$$y = \sum_{k=\pm1} \gamma_k \{e^{ik(\phi+\theta)} - e^{ik(\phi-\theta)}\}; \tag{2.6}$$

$$z = \sum_{k=\pm1} \eta_k e^{ik\theta}. \tag{2.7}$$

Next we introduce general Fourier coefficients:

(i) pure imaginary coefficient $a_k = -\bar{a}_k$ satisfying $a_{-k} = -a_k$;
(ii) real coefficients $b_k$, $k \in \mathbb{Z}$, $b_{-k} = b_k$;
(iii) real coefficients $c_k$, $k \in \mathbb{Z}$, $c_{-k} = c_k$.

It is seen that

$$[\sum_{k=\pm1} a_k\{e^{ik(\phi+\theta)} + e^{ik(\phi-\theta)}\} + \sum_{k=\pm1} b_k\{e^{ik(\phi+\theta)} - e^{ik(\phi-\theta)}\} + \sum_{k=\pm1} c_k e^{ik\theta}] \sin\theta \tag{2.8}$$

corresponds to the part $\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{x}$ in the von Mises-Fisher distribution (1.1) and Kato-McCullagh distribution (1.4). Here $\sin\theta$ is the Jacobian. Because $e^{ik(\phi+\theta)}$, $e^{ik(\phi-\theta)}$ and $e^{ik(\theta)}$ are Fourier basis at frequency $\phi+\theta$, $\phi-\theta$ and $\theta$, it is natural to think of the strength of these frequencies by the spectral densities

$$f_S(\theta,\phi) = [\sum_{k=-\infty}^{\infty} a_k\{e^{ik(\phi+\theta)} + e^{ik(\phi-\theta)}\} + \sum_{k=-\infty}^{\infty} b_k\{e^{ik(\phi+\theta)} - e^{ik(\phi-\theta)}\} + \sum_{k=-\infty}^{\infty} c_k e^{ik\theta}] \sin\theta. \tag{2.9}$$

Mardia (1975) introduced the Fisher-Bingham distribution (1.2) which includes $\kappa\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$. Hence, in (2.5) and (2.6), if we consider $x^2 - y^2$, it is not difficult to see that (2.9) becomes

$$[f_S(\theta,\phi) + \sum_{k=-\infty}^{\infty} d_k e^{ik\phi}] \sin\theta. \tag{2.10}$$

Therefore we can introduce our very general family of distributions on sphere as follows. For a smooth function $H[\cdot]$, we propose

$$F_S(\theta,\phi) = H[\sum_{k=-\infty}^{\infty} (A_k e^{ik(\theta+\phi)} + B_k e^{ik(\theta-\phi)} + C_k e^{ik\theta} + D_k e^{ik\phi})] \sin\theta, \tag{2.11}$$

as a probability density on sphere, where $A_k$, $B_k$, $C_k$ and $D_k$ are complex-valued, and $F_S(\theta,\phi) \geq 0$ so that $\int F_S(\theta,\phi)\,\mathrm{d}\mathbb{P} = 1$.

## 3. Distributions on Torus

In this section we introduce a very general distribution on torus $\mathbb{T}$ in view of time series spectra. Let the 3-dim polar representation be given by

$$\begin{cases} x &= R\cos\theta + r\cos\phi\cos\theta \\ y &= R\sin\theta + r\cos\phi\sin\theta \\ z &= r\sin\phi \end{cases}, \qquad (0 \leq \theta, \phi \leq 2\pi). \tag{3.1}$$

Write $\begin{cases} \cos\theta &= \frac{1}{2}(\mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta}) \\ \sin\theta &= \frac{1}{2\mathrm{i}}(\mathrm{e}^{\mathrm{i}\theta} - \mathrm{e}^{-\mathrm{i}\theta}) \end{cases}$, then we have

$$\begin{aligned} x &= R\frac{1}{2}(\mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta}) + r\frac{1}{4}(\mathrm{e}^{\mathrm{i}\phi} + \mathrm{e}^{-\mathrm{i}\phi})(\mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta}) \\ &= R\frac{1}{2}(\mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta}) + r\frac{1}{4}(\mathrm{e}^{\mathrm{i}(\theta+\phi)} + \mathrm{e}^{-\mathrm{i}(\theta+\phi)} + \mathrm{e}^{-\mathrm{i}(\phi-\theta)} + \mathrm{e}^{\mathrm{i}(\phi-\theta)}); \end{aligned} \tag{3.2}$$

$$\begin{aligned} y &= R\frac{1}{2\mathrm{i}}(\mathrm{e}^{\mathrm{i}\theta} - \mathrm{e}^{-\mathrm{i}\theta}) + r\frac{1}{4\mathrm{i}}(\mathrm{e}^{\mathrm{i}\phi} + \mathrm{e}^{-\mathrm{i}\phi})(\mathrm{e}^{\mathrm{i}\theta} - \mathrm{e}^{-\mathrm{i}\theta}) \\ &= R\frac{1}{2\mathrm{i}}(\mathrm{e}^{\mathrm{i}\theta} - \mathrm{e}^{-\mathrm{i}\theta}) + r\frac{1}{4\mathrm{i}}(\mathrm{e}^{\mathrm{i}(\theta+\phi)} - \mathrm{e}^{-\mathrm{i}(\theta+\phi)} + \mathrm{e}^{-\mathrm{i}(\phi-\theta)} - \mathrm{e}^{\mathrm{i}(\phi-\theta)}); \end{aligned} \tag{3.3}$$

$$z = r\frac{1}{2\mathrm{i}}(\mathrm{e}^{\mathrm{i}\phi} - \mathrm{e}^{-\mathrm{i}\phi}). \tag{3.4}$$

Similarly as in Section 2, we propose a very general family of distributions on $\mathbb{T}$ by

$$g(\theta,\phi) \propto [\sum_{k=-\infty}^{\infty} a_k \mathrm{e}^{\mathrm{i}k\theta} + \sum_{k=-\infty}^{\infty} b_k \mathrm{e}^{\mathrm{i}k\phi} + \sum_{k=-\infty}^{\infty} c_k \mathrm{e}^{\mathrm{i}k(\theta+\phi)} + \sum_{k=-\infty}^{\infty} d_k \mathrm{e}^{\mathrm{i}k(\theta-\phi)}]|J|. \tag{3.5}$$

$\{a_k\}$, $\{b_k\}$, $\{c_k\}$, $\{d_k\}$ are complex-valued coefficients, and are chosen so that $g(\theta,\phi)$ is real-valued. Here $|J| = r(r\cos\phi + R)$, the Jacobian.

Singh et al. (2002) introduced the distribution on $\mathbb{T}$ by (1.5). We can see that the exponent is of our form (3.5). Also Kato and Pewsey (2015) introduced the distribution on $\mathbb{T}$ by (1.6). It is seen that the inside of the reverse function is of our form (3.5).

Hence we introduce our very general family of distributions on $\mathbb{T}$ as follows. For a smooth function $G[\cdot]$, we propose

$$F_T(\theta,\phi) \equiv G[\sum_{k=-\infty}^{\infty} (A'_k \mathrm{e}^{\mathrm{i}k(\theta+\phi)} + B'_k \mathrm{e}^{\mathrm{i}k(\theta-\phi)} + C'_k \mathrm{e}^{\mathrm{i}k\theta} + D'_k \mathrm{e}^{\mathrm{i}k\phi})]|J|, \tag{3.6}$$

as a probability density on $\mathbb{T}$, where $A'_k$, $B'_k$, $C'_k$ and $D'_k$ are complex-valued, and $F_T(\theta,\phi) \geq 0$ so that $\int F_T(\theta,\phi)\,\mathrm{d}\mathbb{P} = 1$.

## 4. Summary and concluding remarks

We could introduce a very general distributions on $\mathbb{S}^2$ and $\mathbb{T}$ by time series spectra (2.11) and (3.6) respectively.The advantage is that we can develop the problem of model selection systematically because the time series spectra have a lot of famous finite parametric models, e.g., AR, MA, and ARMA models, i.e., for modelling, we can use

$$F(\theta,\phi) = H[s_1 f^{(1)}_{\text{ARMA}}(\theta+\phi) + s_2 f^{(2)}_{\text{ARMA}}(\theta-\phi) + s_3 f^{(3)}_{\text{ARMA}}(\theta) + s_4 f^{(4)}_{\text{ARMA}}(\phi)],$$

where $s_1, \ldots, s_4$ are real constants and $f^{(j)}_{\text{ARMA}}$ are ARMA spectral densities (e.g., Taniguchi and Kakizawa (2000)). Hence we can use AIC, BIC etc to select the model, which enriches applications for data from $\mathbb{S}^2$ and $\mathbb{T}$. Also, the systematic asymptotic estimation theory will be possible (e.g., Taniguchi and Kakizawa (2000)).

## Acknowledgement

The authors thank Professor Shogo Kato for his instructive comments.

## References

Kato, S. & Pewsey, A. (2015). A Möbius transformation–introduced distribution on the torus. *Biometrika* **102**(2), 359–370.

Kato, S. & McCullagh, P. (2020). Some properties of a Cauchy family on the sphere derived from the Möbius transformations. *Bernoulli* **26**(4), 3224–3248.

Kent, J. T. (1982). The Fisher-Bingham distribution on the sphere. *J. R. Statist. Soc. B* **44**(1), 71–80.

Mardia, K. V. (1975). Statistics of directional data (with discussion) *J. R. Statist. Soc. B* **37**, 349–393.

Mardia, K. V. & Jupp, P. E. (2000). *Directional Statistics*. New York: Wiley.

Singh, H., Hinzdo, V. & Demchuk, E. (2002). Probablistic model for two dependent circular variables. *Biometrika* **89**(3), 719–723.

Taniguchi, M. & Kakizawa, Y.(2000). *Asymptotic Theory of Statistical Inference for Time Series*. New York: Springer.

Taniguchi, M., Kato, S., Ogata, H. & Pewsey, A. (2020). Models for circular data from time series spectra. *J. Time Ser. Anal.* **41**, 808–829.