Necessary and sufficient condition for the existence of a Fréchet mean on the circle

Benjamin Charlier
Institut de Mathématiques de Toulouse
Université de Toulouse et CNRS (UMR 5219)
Benjamin.Charlier@math.univ-toulouse.fr

September 9, 2011

Abstract

Let $(\mathbb{S}^1, d_{\mathbb{S}^1})$ be the unit circle in $\mathbb{R}^2$ endowed with the arclength distance. We give a sufficient and necessary condition for a general probability measure μ to admit a well defined Fréchet mean on $(\mathbb{S}^1, d_{\mathbb{S}^1})$. This criterion allows to recover already known sufficient conditions of existence. We also derive a new sufficient condition without restriction on the support of the measure. Then, we study the convergence of the empirical Fréchet mean to the Fréchet mean. An algorithm to compute the empirical Fréchet mean is also given.

Keywords : circular data, Fréchet mean, uniqueness. *AMS classification*: 62H11.

1 Introduction

Statistics for non-Euclidean data In many fields of interest, results of an experiment are objects taking values in non-Euclidean spaces. A rather general framework to model such data is Riemannian geometry and more particularly quotient manifolds. As an illustration, in biology or geology, directional data are often used, see e.g. [16] or [7] and references therein. In this case, observations take their values in the circle or a sphere, that is, an Euclidean space quotiented by the action of scaling. Another well known example of non-Euclidean data in the statistical literature is Kendall's Shape Space, see e.g. [12] or [5] and references therein. The shape of a k-ads of the plane is invariant by the action of the group of similarity in the plane. After a convenient renormalization, the shape space is identified to a complex projective space, that is the quotient of a sphere by the action of a group of rotation.

Definition of basic statistical concepts, such as mean, must be adapted for random variables with values in non-Euclidean spaces such as manifolds. To describe the localization of a probability distribution, one needs to define a central value. In an Euclidean space, problems of existence and uniqueness may arise when one tries to defined such a central value. In general metric spaces, the situation is further complicated by extra phenomenons that do not happen in the Euclidean spaces. Therefore, there has been multiple attempts to give a definition of a mean in non Euclidean space, see among many others [2], [4], [14], [15], [10], [6], [18] or [8].

In this paper, we consider the so-called Fréchet mean, see [8], [10], [14] or [2] and references therein. We are particularly interested in the study of its uniqueness. The Fréchet mean is defined on general metric spaces by extending the fact that Euclidean mean minimizes the sum of square the distance to the data, see equation (2.2) below. To study the well definiteness of the Fréchet mean on a manifold, two facts must be taken into account: non uniqueness of geodesics from one point to another (existence of a cut locus) and the effect of curvature, see e.g. [4]. Due to the cut-locus, the distance function is no longer convex and finding conditions to ensure the uniqueness of the Fréchet mean is not obvious. Two main directions have been explored in the literature: bounding the support of the measure in

[10], [12], [14], [3] or [1], or consider special cases of absolutely continuous radial distributions, see [12], [13] or [11]. In a sense, these two conditions control the concentration of the probability measure. The philosophy behind these works is to ensure a convexity property of the Fréchet functional given equation (2.1) below, see e.g. the introduction of [1] for a review of the above cited papers.

Fréchet mean on the circle In this paper, we focus on the so-called Fréchet mean of a probability measure μ on the unit circle $\mathbb{S}^1$ of the plane,

$$\mathbb{S}^1 = \{x_1^2 + x_2^2 = 1,\ (x_1, x_2) \in \mathbb{R}^2\}.$$

We denote by $d_{\mathbb{S}^1}$ the arclength distance on $\mathbb{S}^1$ given for all $x = (x_1, x_2), p = (p_1, p_2) \in \mathbb{S}^1$ by

$$d_{\mathbb{S}^1}(x,p) = 2 \arcsin\left(\frac{\|x-p\|}{2}\right), \tag{1.1}$$

where $\|x - p\| = \sqrt{(x_1 - p_1)^2 + (x_2 - p_2)^2}$ is the Euclidean norm in $\mathbb{R}^2$. The advantage of dealing with a simple object such as the circle is that curvature problems disappear and we only face the cut-locus problem. In this sense, it allows us to completely understand its effect on the non-convexity of the distance function $d_{\mathbb{S}^1}$, and to give a complete answer about the problem of uniqueness. In what follows, we fully characterize probability measures that admit a well defined Fréchet mean on the circle $(\mathbb{S}^1, d_{\mathbb{S}^1})$. In particularly a necessary and sufficient condition is given in Theorem 5.1, which links the existence of a Fréchet mean for a measure μ to the comparison between the distribution μ and the uniform measure λ on $\mathbb{S}^1$. The surprising fact is that λ appears as a benchmark to discriminate measures having a well defined Fréchet mean. The uniform measure λ is the 'worst' possible case as all points of the circle is a Fréchet mean, indeed the Fréchet functional (2.1) is constant and equals to $\frac{\pi^3}{3}$.

In opposition to what have been done before we do not try to ensure convexity property on the Fréchet functional. Indeed, the definition of the Fréchet mean relies on the *global* optimization problem (2.2) which is, in general, non convex. The advantage of our approach is that we do not need to restrict the support or suppose restrictive conditions of symmetry on the density. As the geometry of flat manifold is simple, we can derive explicit form on the Fréchet functional and its derivative which can be hard to compute in non-flat manifolds such as n-dimensional spheres.

1.1 Organization of the paper

In Section 2, we recall the definition of the Fréchet mean, and we review already known sufficient conditions that ensure the well definiteness of the Fréchet mean. In Section 3, we introduce notations that will be used throughout the paper. In Section 4, we give explicit expressions for the Fréchet functional and its derivative. A study of the critical points of the Fréchet functional is also done. Section 5 contains the main result with the necessary and sufficient condition of Theorem 5.1 for the existence of the Fréchet mean for a general measure. We also recover the conditions introduced in Section 2, and we propose a new criterion that ensures the well definiteness of the Fréchet mean. In Section 6, we study the convergence of the empirical Fréchet mean to the Fréchet mean, and describe an algorithm to compute the empirical Fréchet mean.

2 Fréchet Mean

A standard way to extend the definition of mean in non euclidean metric space is to use the minimization property of the Euclidean mean. This definition is usually credited to M. Fréchet in [8] although some authors credit it to E. Cartan (see e.g. [12]). Let M be a metric measured space endowed with the metric d and the probability measure μ. The Fréchet functional is defined for all $p \in M$ by

$$F_\mu(p) = \frac{1}{2} \int_M d^2(x,p) d\mu(x). \tag{2.1}$$

Definition 2.1. *We say that the Fréchet mean of* μ *in* (M, d) *is well defined if* F_μ *is finite and admits a* unique *argmin. That is, there exists a unique* $p^* \in M$ *satisfying* $F_\mu(p^*) = \min_{p\in M} F_\mu(p)$*, and we note*

$$p^* = \underset{p\in M}{argmin}\, F_\mu(p). \tag{2.2}$$

When (M, d) is a Riemannian manifold, some authors call the argmins of F_μ the Riemanian center of mass in [18] or intrinsic mean in [2]. In what follows we restrict our attention to the case of the circle $\mathbb{S}^1$ endowed with the arclength distance $d_{\mathbb{S}^1}$ defined in (1.1). The circle $(\mathbb{S}^1, d_{\mathbb{S}^1})$ is a simple one dimensional compact Riemannian manifold of finite diameter $\max\{d_{\mathbb{S}^1}(x, p),\ x, p \in \mathbb{S}^1\} = \pi$. The Fréchet functional F_μ is thus finite for all $p \in \mathbb{S}^1$ and is continuous. Hence F_μ attains its minimum in at least one point and the only issue at hand is uniqueness.

2.1 Previous work on the uniqueness of the Fréchet mean

There exists few general conditions to ensure that a measure on a metric space admits a well defined Fréchet mean. As in Euclidean spaces, they are related to a certain notion of concentration. We refer to [9] for the definitions of notions in Riemannian geometry used in this part. To the best of our knowledge, the only case treated in the literature is when (M, d) is a complete Riemannian manifold. In this framework, the non-uniqueness of the Fréchet mean is due to two facts: existence of a cut locus and the effect of curvature, see [4] and references therein for more detailed discussions on this point. In the rest of this section, we review some criteria that ensure the well definiteness of the Fréchet mean of a manifold (M, d) endowed with a probability measure μ.

Manifold with negative curvature The most general result concerns the simply connected Riemannian manifold (M, d) with negative sectional curvature. They are usually called Cartan-Hadamard manifolds and are globally diffeomorphic to $\mathbb{R}^n$ with some change in the metric. In particular there is no cut locus, i.e there is a unique minimizing geodesic between two points. In this case, a general probability measure μ admits a Fréchet mean provided $F_\mu(p)$ is finite for some $p \in M$, see [2] Theorem 2.1.

The situation is more complex when (M, d) has a non negative curvature. In this setting, it exists two kinds of sufficient conditions that ensure the well definiteness of the Fréchet mean.

Bound on the support The first condition concerns the complete connected Riemannian manifold (M, d) with non negative scalar curvature. The condition consists in a restriction on the support of the measure that must be contained in a sufficiently small geodesic ball [10]. If (M, d) is a flat Riemannian manifold, a sufficient condition for the existence of the Fréchet mean is that the support of μ is contained in a geodesically convex open normal neighborhood of M and that $F_\mu(p)$ is finite for some p. For example, on the circle, this imposes μ to be supported in a half-circle. Suppose now that (M, d) has a scalar curvature bounded from above by $C > 0$ and that $R > 0$ denotes the infimum of the injectivity radius. Then, a sufficient condition to ensure the existence of the Fréchet mean of a measure μ is that the support of μ is contained in a geodesic ball of radius less than $\frac{1}{4}\min\left\{R, \frac{\pi}{\sqrt{C}}\right\}$. For a precise statement and definitions see [14] Theorem 1. In the special case of the n-dimensional unit sphere $\mathbb{S}^n = \{x_1^2 + \ldots + x_{n+1}^2 = 1,\ x_1, \ldots, x_{n+1} \in \mathbb{R}\}$ of $\mathbb{R}^{n+1}$, the Fréchet mean of a measure μ is well defined as soon as its support is strictly included in a geodesic ball of radius $\frac{\pi}{4}$. In [3], the authors improve the bound when the measure has finite support, i.e is a finite sum of Dirac's masses. The Fréchet mean of such a measure is well defined if the support of the measure lies in a closed hemisphere, that is a geodesic ball of radius $\frac{\pi}{2}$, and is not contained entirely in the boundary of this hemisphere. The philosophy behind this condition of boundedness of the support, is to guaranty the *convexity* of F_μ in the interior of the support of μ.

Radial distribution The second kind of conditions that ensure the well definiteness of the Fréchet mean concerns particular examples of Riemannian manifolds. We denote by (Σ_2^k, ρ) the Kendall shape space endowed with the Procrustes distance ρ, see e.g. [12] Chapter 9 for definitions. It turns out that this space is isometric to the complex projective space $\mathbb{C}P^{k-2}$ endowed with the Fubini-Study metric. In [13] and [12], the authors consider the case of an absolutely continuous probability measure μ with respect to the uniform law in (Σ_2^k, ρ). If there is a $p^* \in \Sigma_2^k$ such that the density of μ is a decreasing function of $\rho(p^*, .)$ (i.e μ is radially distributed probability measure around p^*) then p^* is the Fréchet mean of μ. See Theorem 9.2 and 9.3 of [12] for details and proofs. There exists a similar result for the special case of the circle $(\mathbb{S}^1, d_{\mathbb{S}^1})$ which is proved in [11]. Let $p^* \in \mathbb{S}^1$ and μ be an absolute continuous probability measure with density $f(d_{\mathbb{S}^1}(p^*, .))$ where $f : [0, \pi] \longrightarrow \mathbb{R}$ is a decreasing function. Then p^* is the Fréchet mean of μ. Note, that this decreasing radial distribution condition can be interpreted as a concentration condition around the Fréchet mean p^*.

3 Notations

In what follows, $\mathbb{1}_A$ denotes the indicator function of the set $A \subset \mathbb{R}$ and the notation $\int_a^b f(t)d\mu_p(t)$ stands for the Lebesgue integral $\int_{[a,b[} f(t)d\mu_p(t)$ if $a \leq b$ and $\int_{[b,a[} f(t)d\mu_p(t)$ if $b > a$.

3.1 The distance function

The one dimensional sphere $\mathbb{S}^1$ can be identified with the torus $\mathbb{R}/(2\pi\mathbb{Z})$, that is the real line $\mathbb{R}$ quotiented by the equivalent relation $\sim$ defined for all $\theta_1, \theta_2 \in \mathbb{R}$ by $\theta_1 \sim \theta_2$ if and only if $\theta_1 - \theta_2 = 2\pi k,\ k \in \mathbb{Z}$. The equivalence class of $\theta \in \mathbb{R}$ is denoted by $[\theta]$. By choosing an arbitrary $p_0 \in \mathbb{S}^1$, we can identify any $p \in \mathbb{S}^1$ with $\theta_p^{p_0}$, the angle between p_0 and p. The arclength distance defined in (1.1) reads now, for $\theta_1, \theta_2 \in \mathbb{R}$,

$$d_{\mathbb{S}_1}(\theta_1, \theta_2) = \min\{|\theta_1 - \theta_2 + 2\pi k|\,, k \in \mathbb{Z}\}.$$

The circle $\mathbb{S}^1$ is locally isometric to the real line $\mathbb{R}$ and we use the same notation for the spherical distance between points in $\mathbb{S}^1 \subset \mathbb{R}^2$ and points in $\mathbb{R}/(2\pi\mathbb{Z})$.

The cut locus of a point $p \in \mathbb{S}^1$ is denoted by $\tilde{p}$ which is equals to the opposite point of p, that is $\tilde{p} = -p$. In $\mathbb{R}/(2\pi\mathbb{Z})$, the cut locus of $[\theta]$ is $[\theta + \pi]$. We refer to [9] for details about cut loci.

3.2 Normal coordinates

To make explicit computation on $\mathbb{S}^1$, we use charts (also called coordinate systems), that is smooth one to one maps between $\mathbb{R}$ and $\mathbb{S}^1$. This is the terminology used in Riemannian geometry as the circle $\mathbb{S}^1$ is a flat sub-manifold of dimension 1. Even if the geometry of $\mathbb{S}^1$ is rather simple, this terminology allows us to connect the concepts used in this paper with more general situations. For each $p \in \mathbb{S}^1$, there is a canonical chart called the exponential map $e_p : \mathbb{R} \longrightarrow \mathbb{S}^1$. It is defined for all $\theta \in \mathbb{R}$ by $e_p(\theta) = R_\theta p$ where $R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ is the rotation matrix of angle θ that fixes 0 in $\mathbb{R}^2$. Equivalently, we could define $e_p : \mathbb{R} \longrightarrow \mathbb{R}/(2\pi\mathbb{Z})$ by $e_p(\theta) = [\theta]$ and the exponential map corresponds in this case to the quotient map $\mathbb{R} \longrightarrow \mathbb{R}/(2\pi\mathbb{Z})$.

For any $p \in \mathbb{S}^1$ the exponential map e_p, is onto but not one to one. In this paper, we choose to restrict the domain of definition of e_p to $[-\pi, \pi[$, for all $p \in \mathbb{S}^1$. Thus, there is now a unique $\theta_{p_2}^{p_1} \in [-\pi, \pi[$ satisfying $e_{p_1}(\theta_{p_2}^{p_1}) = p_2$ and we have for all $p \in \mathbb{S}^1$

$$e_p : [-\pi, \pi[\longrightarrow \mathbb{S}^1 \qquad \text{and} \qquad e_p^{-1} : \mathbb{S}^1 \longrightarrow [-\pi, \pi[.$$

Such parametrizations are called normal coordinate systems centered at p and $\theta_{p_2}^{p_1}$ is nothing else but the coordinate of p_2 read in a normal coordinate system centered in p_1. In these coordinate systems, the cut locus of a point p is $\theta_p - \pi$. To simplify notations, we will omit the exponent p_1 if no confusion is possible and we will write $\theta_{p_2}^{p_1} = \theta_{p_2}$.

In the case of the circle $\mathbb{S}^1$, a changing of coordinate system is particularly simple. It corresponds to translation in $\mathbb{R}/(2\pi\mathbb{Z})$. For any $p_1, p_2 \in \mathbb{S}^1 \simeq \mathbb{R}/(2\pi\mathbb{Z})$ we have $e_{p_2}^{-1} \circ e_{p_1}(\theta) = \theta - \theta_{p_1}^{p_2}$, where $e_{p_2}^{-1} \circ e_{p_1} : \mathbb{R} \longrightarrow \mathbb{R}$.

3.3 Probability measure on the circle

Unless specified, in what follows, μ denotes a general probability measure on $(\mathbb{S}^1, \mathcal{B}(\mathbb{S}^1))$ where $\mathcal{B}(\mathbb{S}^1)$ is the Borel set of $\mathbb{S}^1 \subset \mathbb{R}^2$. Given a $p \in \mathbb{S}^1$, μ_p is the image measure of μ through $e_p^{-1} : \mathbb{S}^1 \longrightarrow [-\pi, \pi[$. It defines a measure on $\mathbb{R}$ given by

$$\mu_p(A) = \mu \circ e_p(A \cap [-\pi, \pi[), \qquad \text{for all } A \in \mathcal{B}(\mathbb{R}), \tag{3.1}$$

where $\mathcal{B}(\mathbb{R})$ is the Borel set in $\mathbb{R}$. In words, μ_p is then nothing else but the measure μ read in the normal coordinate system centered at p. Note that we have $\mu_p([-\pi, \pi[) = 1$. For a measure μ_p on $\mathbb{R}$, let

$$m(\mu_p) = \int_{\mathbb{R}} t d\mu_p(t) \qquad \text{and} \qquad \mathrm{Var}(\mu_p) = \int_{\mathbb{R}} t^2 d\mu_p(\theta) - (m(\mu_p))^2$$

be the usual Euclidean mean/expectation and variance of μ. The Fréchet functional on $(\mathbb{R}, |\cdot|)$ endowed with a probability measure μ_p is $F_{\mu_p}(\theta) = \frac{1}{2}\int_{\mathbb{R}} |t-\theta|^2 d\mu_p(t)$, for all $\theta \in \mathbb{R}$. It attains its unique minimum at $m(\mu_p)$ and its value at this minimum is $\mathrm{Var}(\mu_p)$. Note that we have for all $p_0, p \in \mathbb{S}^1$

$$F_\mu(p) = F_\mu \circ e_{p_0}(\theta_p^{p_0}) = F_{\mu_{p_0}}(\theta_p^{p_0}).$$

Finally, following [6], a point $p \in \mathbb{S}^1$ satisfying $m(\mu_p) = 0$ is called an exponential barycenter.

4 The Fréchet functional on the Circle

Using the notations introduced in Section 3, we give an expression of the Fréchet functional F_μ in the normal coordinate system centered at $p_0 \in \mathbb{S}^1$. For all $\theta_p^{p_0} \in [-\pi, \pi[$,

$$F_\mu(p) = \frac{1}{2}\begin{cases} \int_{-\pi}^{\theta_p^{p_0}-\pi} (\theta + 2\pi - \theta_p^{p_0})^2 d\mu_{p_0}(\theta) + \int_{\theta_p^{p_0}-\pi}^{\pi} (\theta - \theta_p^{p_0})^2 d\mu_{p_0}(\theta), & \text{if } 0 \le \theta_p^{p_0} < \pi, \\ \int_{-\pi}^{\theta_p^{p_0}+\pi} (\theta - \theta_p^{p_0})^2 d\mu_{p_0}(\theta) + \int_{\theta_p^{p_0}+\pi}^{\pi} (\theta - 2\pi - \theta_p^{p_0})^2 d\mu_{p_0}(\theta), & \text{if } -\pi \le \theta_p^{p_0} < 0. \end{cases} \tag{4.1}$$

Then, the Fréchet functional is Lipschitz since by the triangle inequality we have $|d^2_{\mathbb{S}^1}(p_1, x) - d^2_{\mathbb{S}^1}(x, p_2)| \le 2\pi d_{\mathbb{S}^1}(p_1, p_2)$ for any $p_1, p_2, x \in \mathbb{S}^1$ which yields $|F_\mu(p_1) - F_\mu(p_2)| \le 2\pi d_{\mathbb{S}^1}(p_1, p_2)$. Then, F_μ is continuous everywhere on $\mathbb{S}^1$. Better, it is continuously differentiable everywhere except at the cut locus of points of strictly positive measure (the atoms of the measure).

4.1 The derivative of the Fréchet functional

A function $f : [-\pi, \pi[\longrightarrow \mathbb{R}$ is said left continuous on $[-\pi, \pi[$ if it is left continuous everywhere on $]-\pi, \pi[$ and with $\lim_{\varepsilon \to 0^-} f(\pi + \varepsilon) = f(-\pi)$. Similarly, f is said to be continuous on $[-\pi, \pi[$ if it is left and right continuous on $[-\pi, \pi[$. We provide an explicit expression of the derivative of F_μ,

Proposition 4.1. *Let μ be a probability measure on $(\mathbb{S}^1, d_{\mathbb{S}^1})$ and fix an arbitrary $p_0 \in \mathbb{S}^1$. Then, $F_\mu : \mathbb{S}^1 \longrightarrow \mathbb{R}$ is differentiable in following sense :*

1. *Let $p \in \mathbb{S}^1$ be a point with a cut locus of μ-measure 0, i.e $\mu(\{-p\}) = 0$. Then F_μ is continuously differentiable at p and we have*

$$\frac{d}{d\theta} F_{\mu_{p_0}}(\theta_p^{p_0}) = \begin{cases} \theta_p^{p_0} - 2\pi\mu_{p_0}([-\pi, -\pi + \theta_p^{p_0}[) - m(\mu_{p_0}), & \text{if } 0 \le \theta_p^{p_0} < \pi, \\ \theta_p^{p_0} + 2\pi\mu_{p_0}([\pi + \theta_p^{p_0}, \pi[) - m(\mu_{p_0}), & \text{if } -\pi \le \theta_p^{p_0} < 0. \end{cases} \tag{4.2}$$

2. *The function* $\frac{d}{d\theta}F_{\mu_{p_0}}$ *is left continuous on* $[-\pi,\pi[$. *Then we extend the definition of the derivative of* F_μ *by setting* for all $\theta \in [-\pi,\pi[$

$$\frac{d}{d\theta}F_{\mu_{p_0}}(\theta) := \lim_{\varepsilon\to 0^-}\frac{d}{d\theta}F_{\mu_{p_0}}(\theta+\varepsilon). \tag{4.3}$$

3. *Let* $p \in \mathbb{S}^1$ *be a point with a cut locus of positive measure, i.e* $\mu(\{-p\}) > 0$. *Then,* p *is a cusp point of* F_μ *in the sense that* $\lim_{\varepsilon\to 0^-}\frac{d}{d\theta}F_{\mu_{p_0}}(\theta_p^{p_0}+\varepsilon) - \lim_{\varepsilon\to 0^+}\frac{d}{d\theta}F_{\mu_{p_0}}(\theta_p^{p_0}+\varepsilon) = -\mu(\{-p\})$.

Note that the left-continuity comes from our convention on the exponential map which is defined on $[-\pi,\pi[$. If a measure μ is such that $\mu(\{p\}) = 0$ for all $p \in \mathbb{S}^1$ then F_μ is of class $\mathcal{C}^1$ on $[-\pi,\pi[$. Differentiability issues appear when the measure μ has atoms, i.e points p such that $\mu(\{p\}) > 0$. See Figure 1 where $F_{\mu_{p^*}}$ has three cusp points at $-\pi$, $-\frac{\pi}{3}$ and $\frac{\pi}{3}$ corresponding to cut loci of the three Dirac masses.

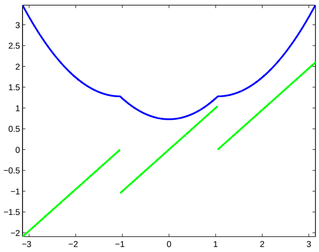

Figure 1: Let $\mu = \frac{1}{6}\delta_{p_1} + \frac{2}{3}\delta_{p^*} + \frac{1}{6}\delta_{p_2}$ with $p_1 = R_{\frac{2\pi}{3}}p^*$ and $p_2 = R_{-\frac{2\pi}{3}}p^*$. In blue: $F_{\mu_{p^*}}$. In green: $\frac{d}{d\theta}F_{\mu_{p^*}}$.

Proof. First of all, fix an arbitrary $p_0 \in \mathbb{S}^1$. For convenience we omit in this proof the superscript p_0 by writing $\theta_p = \theta_p^{p_0}$ for all $p \in \mathbb{S}^1$. Then, in the coordinate system centered at p_0 we have for all $\theta \in [-\pi,\pi[$

$$F_{\mu_{p_0}}(\theta) = \frac{1}{2}\int_{\mathbb{R}} t^2 d\mu_{p_0}(t) - \theta m(\mu_{p_0}) + \frac{1}{2}\theta^2 + 2\pi(g^+_{\mu_{p_0}}(\theta)\mathbb{1}_{[0,\pi[}(\theta) + g^-_{\mu_{p_0}}(\theta)\mathbb{1}_{[-\pi,0[},(\theta)) \tag{4.4}$$

where $g^+_{\mu_{p_0}}(\theta) = \int_{-\pi}^{-\pi+\theta}(\pi+t-\theta)d\mu_{p_0}(t)$ and $g^-_{\mu_{p_0}}(\theta) = \int_{\theta+\pi}^{\pi}(\pi-t+\theta)d\mu_{p_0}(t)$. Hence to prove the claim in Proposition 4.1, we just have to compute the derivative of $g^+_{\mu_{p_0}}$. The case of $g^-_{\mu_{p_0}}$ can be deduced in the same way. For all $\theta \in]0,\pi[$ and $\varepsilon \in \mathbb{R}$ such that $\theta+\varepsilon \in]0,\pi[$ we have,

$$\frac{1}{\varepsilon}(g^+_{\mu_{p_0}}(\theta_p+\varepsilon) - g^+_{\mu_{p_0}}(\theta_p)) = \frac{1}{\varepsilon}\int_{-\pi+\theta_p}^{-\pi+\theta_p+\varepsilon}(\pi+t-\theta_p)d\mu_{p_0}(t) - \int_{-\pi}^{-\pi+\theta_p+\varepsilon}d\mu_{p_0}(t). \tag{4.5}$$

Then, let $p \in \mathbb{S}^1$ with $\theta_p > 0$ and with a cut locus of μ-measure 0. This is equivalent to the fact that $\mu_{p_0}(\{\theta_p - \pi\}) = 0$. The first term in the right hand side of (4.5) tends to zero as $\varepsilon \to 0$ since the integrand is, in absolute value, less than $|\varepsilon|$ and if $\varepsilon > 0$ (resp. $\varepsilon < 0$), then $\mu_{p_0}([-\pi+\theta_p, -\pi+\theta_p+\varepsilon[)$ (resp. $\mu_{p_0}([-\pi+\theta_p+\varepsilon, -\pi+\theta_p[)$) tends to zero as $\varepsilon \to 0$ since $\mu_{p_0}(\{\theta_p-\pi\}) = 0$. Similarly, the second term tends to $\mu_{p_0}([-\pi,-\pi+\theta_p[)$ when $\varepsilon \to 0$. The same arguments can be applied for $g^-_{\mu_{p_0}}$ and we have, $\frac{d}{d\theta}g^+_{\mu_{p_0}}(\theta_p) = -\mu_{p_0}([-\pi,-\pi+\theta_p[)$, if $\theta_p \in]0,\pi[$ and $\frac{d}{d\theta}g^-_{\mu_{p_0}}(\theta_p) = \mu_{p_0}([\pi+\theta_p,\pi[)$, if $\theta_p \in]-\pi,0[$. If $\theta_p = 0$, (i.e $p_0 = p$) and $\mu_{p_0}(\{-\pi\}) = 0$ we can extend by continuity and let

$\frac{d}{d\theta} g^+_{\mu_{p_0}}(0) = \lim_{\varepsilon\to 0^-} \frac{d}{d\theta} g^-_{\mu_{p_0}}(0+\varepsilon) = 0$. Again, at $\theta_p = -\pi$ with $\mu_{p_0}(\{0\}) = 0$ we can extend by continuity and let $\frac{d}{d\theta} g^+_{\mu_{p_0}}(-\pi) = \lim_{\varepsilon\to 0^-} \frac{d}{d\theta} g^+_{\mu_{p_0}}(\pi+\varepsilon) = \mu^d_{p^*}([0,\pi[$.

Suppose now that $p \in \mathbb{S}^1$ is such that $\mu_{p_0}(\theta_p - \pi) > 0$. By the same arguments as above, we have if $\theta_p > 0$, $\lim_{\varepsilon\to 0^-} \frac{1}{\varepsilon}(g^+_{\mu_{p_0}}(\theta_p+\varepsilon) - g^+_{\mu_{p_0}}(\theta_p)) = -\mu_{p_0}([-\pi, -\pi+\theta_p[)$, and if $\theta_p < 0$, $\lim_{\varepsilon\to 0^-} \frac{1}{\varepsilon}(g^-_{\mu_{p_0}}(\theta_p + \varepsilon) - g^-_{\mu_{p_0}}(\theta_p)) = \mu_{p_0}([\pi+\theta_p, \pi[)$, and the formula can be extended to 0 and $-\pi$. Then, the function $\frac{d}{d\theta} F_\mu \mu_{p_0}$ is left continuous everywhere on $[-\pi,\pi[$, continuous at each point with a cut locus of μ-measure 0 and with a (negative) jump of size $-\lim_{\varepsilon\to 0^+} \mu_{p_0}([-\pi, -\pi+\theta_p+\varepsilon[) = -\mu_{p_0}(\{\pi+\theta_p\})$ at a point with a cut locus of positive measure. □

4.2 Local minimum of the Fréchet functional

In this section, we precise the link between critical points of F_μ (i.e points at which the derivative of F_μ, in the sens of Proposition 4.1, is 0), exponential barycenters of μ (i.e points p satisfying $m(\mu_p) = 0$) and local argmins of F_μ. Note that part of the results presented here are already known for general Riemannian manifold in [2] (in particular see Theorem 2.1), [19] Theorem 2 or [6]. The first result shows us that the critical points are exactly the exponential barycenter.

Corollary 4.1. *Let μ be a measure of probability on $\mathbb{S}^1$. We have*

$$\frac{d}{d\theta} F_{\mu_{p_0}}(\theta_p^{p_0}) = -m(\mu_p),$$

where the derivative is in the sens of equation (4.3).

Proof. We simply give an expression of $m(\mu_p) \in \mathbb{R}$ in normal coordinate centered at a $p_0 \in \mathbb{S}^1$,

$$\begin{aligned} m(\mu_p) = \int_{\mathbb{R}} t d\mu_p(t) &= \begin{cases} \int_{-\pi}^{\theta_p^{p_0}-\pi} (t - \theta_p^{p_0} + 2\pi) d\mu_{p_0}(t) + \int_{\theta_p^{p_0}-\pi}^{\pi} (t-\theta_p^{p_0}) d\mu_{p_0}(t) & \text{if } 0 < \theta_p^{p_0} < \pi, \\ \int_{-\pi}^{\theta_p^{p_0}+\pi} (t-\theta_p^{p_0}) d\mu_{p_0}(t) + \int_{\theta_p^{p_0}+\pi}^{\pi} (t - \theta_p^{p_0} - 2\pi) d\mu_{p_0}(t) & \text{if } -\pi \le \theta_p^{p_0} < 0 \end{cases} \\ &= -\frac{d}{d\theta} F_{\mu_{p_0}}(\theta_p^{p_0}). \end{aligned}$$

Where the last inequality is given by Proposition 4.1. □

When $\mu(\{p\}) = 0$ for all $p \in \mathbb{S}^1$, to be a critical point of F_μ is equivalent to be an extremum of F_μ. Unfortunately, when the measure μ has atoms, to be a critical point of F_μ is not a sufficient condition for p to be an extremum of F_μ. Consider example of Figure 1 wher the Fréchet mean is well defined and is located at p^*, the mass point of weight $\frac{2}{3}$. Then, $m(\mu_p) = 0$ for $p = p^*$ (that is $\theta_p^{p^*} = 0$) and $p = \tilde{p}_2 = -p_2$ (that is $\theta_p^{p^*} = -\frac{\pi}{3}$). Note that for $p = \tilde{p}_1 = -p_1$, we have $m(\mu_p) < 0$. This example is particular since the critical points are at the cut loci of the atoms of μ. Nevertheless, a local argmin of F_μ cannot be a cusp point as we have the following result :

Corollary 4.2. *Let μ be a probability measure on $\mathbb{S}^1$. The cut locus of a (local or global) minimum of F_μ is of μ-measure 0.*

Proof. Choose an arbitrary $p_0 \in \mathbb{S}^1$ and let $p_m \in \mathbb{S}^1$ be a minimum of F_μ satisfying $\mu(\{\tilde{p}_m\}) = \delta > 0$. Recall Statement 3 in Proposition 4.1, where it is shown that $\frac{d}{d\theta} F_{\mu_{p_0}}$ has a negative jump of size $-\delta$ at $\theta_{p_m}^{p_0}$. Then, the possible sign of $\left(\lim_{\varepsilon\to 0^-} \frac{d}{d\theta} F_{\mu_{p_0}}(\theta_{p_m}^{p_0}+\varepsilon), \lim_{\varepsilon\to 0^+} \frac{d}{d\theta} F_{\mu_{p_0}}(\theta_{p_m}^{p_0}+\varepsilon)\right)$ are $(+,+)$, $(+,-)$ and $(-,-)$. This means that $\theta_{p_m}^{p_0}$ cannot be a minimum of $F_{\mu_{p_0}}$ since it would correspond to the case $(-,+)$. □

Remark 4.1. *Note that assumptions of Corollary 1 in [19] and Theorem 1 in [2] contains a condition of the form $\mu\{\tilde{p}^*\} = 0$ to ensure the (classical) differentiability of the Fréchet functional at its minimum. In the case of the circle, Corollary 4.2 shows us that the Fréchet functional is* always *differentiable at its minimum.*

As an illustration, consider again Figure 1. The point $\theta_{pm}^{p^*} = -\frac{\pi}{6}$ corresponds to the case $(-,-)$, the point $\theta_{pm}^{p^*} = \frac{\pi}{6}$ to the case $(+,+)$ and the point $\theta_{pm}^{p^*} = -\pi$ to the case $(+,-)$. It can be shown that when μ is a purely atomic measure we have a better result: the 'regular' critical points (i.e critical point that are not cusp point) are local minima of F_μ.

We end this section with a result that allows us to compute efficiently the critical point of F_μ. The proof is omitted as it is an immediate consequence of Corollary 4.1.

Corollary 4.3. *Let μ be a measure of probability on $\mathbb{S}^1$, then the following propositions are equivalent*

1. *$p \in \mathbb{S}^1$ is a critical point of F_μ*
2. *$p \in \mathbb{S}^1$ satisfy $m(\mu_p) = 0$*
3. *for any $p_0 \in \mathbb{S}^1$,*

$$\begin{cases} \frac{1}{2\pi}(\theta_p^{p_0} - m(\mu_{p_0})) = \mu_{p_0}([-\pi, -\pi + \theta_p^{p_0}[), & \text{if } 0 \leq \theta_p^{p_0} < \pi, \\ \frac{1}{2\pi}(-\theta_p^{p_0} - m(\mu_{p_0})) = \mu_{p_0}([\pi + \theta_p^{p_0}, \pi[), & \text{if } -\pi \leq \theta_p^{p_0} < 0. \end{cases} \tag{4.6}$$

5 Necessary and sufficient condition for the existence of the Fréchet mean

5.1 Main result

In what follows, p^* denotes a critical point of F_μ, i.e a point at which the derivative of F_μ, in the sens of Proposition 4.1, is 0. Let us consider for all $p \in \mathbb{S}^1$, the functional

$$G_\mu(p) = F_\mu(p) - F_\mu(p^*),$$

which vanishes at $p = p^*$. We give here our main result,

Theorem 5.1. *Let μ be a general probability measure and $p^* \in \mathbb{S}^1$ be a critical point of F_μ. Then, the following propositions are equivalent,*

1. *p^* is a well defined Fréchet mean of $(\mathbb{S}^1, \mu)$.*
2. *For all $p \neq p^*$ $G(p) > 0$*
3. *For all $0 < \theta < \pi$*
$$\int_0^\theta \lambda([-\pi, -\pi + t[) - \mu_{p^*}([-\pi, -\pi + t[)dt > 0,$$
and for all $-\pi \leq \theta < 0$,
$$\int_\theta^0 \lambda(]\pi + t, \pi[) - \mu_{p^*}(]\pi + t, \pi[)dt > 0,$$
where λ is the uniform measure on $[-\pi, \pi[$ and μ_{p^*} is defined in (3.1).

Theorem 5.1 gives a necessary and sufficient condition for the existence of the Fréchet mean of a general measure μ on the circle $\mathbb{S}^1$. This condition is given in terms of comparison between the μ-measure and the uniform measure λ of balls centered at the cut locus of a global minimum. Note that the uniform measure λ has a density with respect to the Lebesgue measure equals to the constant function $\frac{1}{2\pi}$.

As μ is a probability measure, the functions $t \longmapsto \lambda([-\pi, -\pi + t[) - \mu_{p^*}([-\pi, -\pi + t[)$ and $t \longmapsto \lambda(]\pi - t, \pi[) - \mu_{p^*}(]\pi - t, \pi[)$ do not need to be always positive for $t \in [-\pi, 0[$ and $t \in [0, \pi[$ respectively. A situation where it is positive is studied in Proposition 5.2 of the next paragraph. The point is that the μ-measure of a (small) neighborhood of the cut locus of p^* cannot be larger than the uniform measure of this neighborhood.

5.2 Proof of Theorem 5.1

First of all, there exists at least one global argmin $p^* \in \mathbb{S}^1$ of F_μ since F_μ is a continuous function defined on the compact set $\mathbb{S}^1$. Then, Proposition 4.1 and Corollary 4.2 ensure that p^* is a critical point of F_μ. Note also that Statement 1. and 2. are equivalent as they are simple reformulations of the definition of a well defined Fréchet mean. To prove Theorem 5.1 we only need to show that Statement 2. is equivalent to Statement 3. The proof relies on the computation of the derivative of G_μ expressed in a well chosen coordinate system.

Corollary 4.1 ensures that p^* is an exponential barycenter, i.e satisfy $m(\mu_{p^*}) = 0$. Hence in the normal coordinate system centered at such a p^* the functional $G_{\mu_{p^*}} = F_{\mu_{p^*}}(\theta) - F_{\mu_{p^*}}(0)$ has a particularly simple expression, that is

$$G_{\mu_{p^*}}(\theta) = \frac{1}{2}\theta^2 + 2\pi \mathbb{1}_{[0,\pi[}(\theta) \int_{-\pi}^{\theta-\pi} (\pi + t - \theta) d\mu_{p^*}(t) + 2\pi \mathbb{1}_{[-\pi,0[}(\theta) \int_{\theta+\pi}^{\pi} (\pi - t + \theta) d\mu_{p^*}(t).$$

Using Proposition 4.1 we have the following result,

Lemma 5.1. *Let μ be a probability measure on $\mathbb{S}^1$ and $p^* \in \mathbb{S}^1$ be an argmin of F_μ. Then for any $\theta \in [-\pi, \pi[$*

$$G_{\mu_{p^*}}(\theta) = 2\pi \begin{cases} \int_0^\theta \frac{t}{2\pi} - \mu_{p^*}([-\pi, -\pi + t[) dt, & \text{if } 0 \le \theta < \pi, \\ \int_\theta^0 \frac{-t}{2\pi} - \mu_{p^*}([\pi + t, \pi[) dt, & \text{if } -\pi \le \theta < 0 \end{cases}$$

Proof. The probability measure μ can be decomposed as follow,

$$\mu = a\mu^d + (1-a)\mu^\delta, \qquad 0 \le a \le 1, \tag{5.1}$$

where μ^d is a probability measure such that $\mu_d(\{p\}) = 0$ for all $p \in \mathbb{S}^1$ and $\mu_\delta = \sum_{j=1}^{+\infty} \omega_j \delta_{p_j}$ where $\sum_{j=0}^{+\infty} \omega_j = 1$ and $p_1, \dots, p_n, \dots \in \mathbb{S}^1$. Hence, we consider the two cases separately : first, when the measure is non atomic, and then, when it is purely atomic. The general case follows immediately in view of equation (5.1).

First, assume that μ is an atomless measure of $\mathbb{S}^1$. Proposition 4.1 ensures that F_μ is continuously differentiable everywhere and the real function $F_{\mu_{p^*}}$ is of class $\mathcal{C}^1$ on $[-\pi, \pi[$. Formula (4.2) and the fundamental theorem of calculus gives for all $\theta \in [-\pi, \pi[$

$$g_{\mu_{p^*}}(\theta) = \int_0^\theta t dt - 2\pi \begin{cases} \int_0^\theta \mu_{p^*}([-\pi, -\pi + t[) dt, & \text{if } \quad 0 \le \theta < \pi, \\ \int_\theta^0 \mu_{p^*}([\pi + t, \pi[) dt, & \text{if } \quad -\pi \le \theta < 0. \end{cases} \tag{5.2}$$

Consider now the case where μ is a purely atomic measure. First, we treat the case where the number of mass of Dirac in the sum is finite, i.e $\mu = \sum_{j=1}^n \omega_j \delta_{p_j}$, $n \in \mathbb{N}$. Recall that $F_{\mu_{p^*}}$ is a Lipschitz function on $[-\pi, \pi[$. Proposition 4.1 ensures that the derivative is piecewise continuous and formula (4.2) holds for all $\theta \in [-\pi, \pi[\backslash \{\theta_{\tilde{p}_j}\}_{j=1}^n$, i.e points that have a cut locus of μ-measure 0. Hence for all $\theta \in [-\pi, \pi[$, equation (5.2) holds too.

To treat the case where $\mu = \sum_{j=1}^{+\infty} \omega_j \delta_{x_j}$ we proceed by approximation. Let $\phi(n) = \{j \in \mathbb{N} \mid \omega_j \ge \frac{1}{2^n}\}$ and remark that $\text{Card}(\phi(n)) < +\infty$ for all $n \in \mathbb{N}$ since $\sum_{j=1}^{+\infty} \omega_j = 1$. Then if $\nu_{p^*}^n = \frac{1}{c(n)} \sum_{j \in \phi(n)} \omega_j \delta_{x_j}$, where $c(n) = \sum_{j \in \phi(n)} \omega_j$ is a normalizing constant, we have for all $\theta \in [-\pi, \pi[$,

$$g_{\nu_{p^*}^n}(\theta) = \int_0^\theta t dt - 2\pi \begin{cases} \int_0^\theta \nu_{p^*}^n([-\pi, -\pi + t[) dt, & \text{if } \quad 0 \le \theta < \pi, \\ \int_\theta^0 \nu_{p^*}^n([\pi + t, \pi[) dt, & \text{if } \quad -\pi \le \theta < 0. \end{cases}$$

The sequence $(\nu_{p^*}^n)_{n \ge 1}$ converges to μ in total variation. By the dominated convergence Theorem for all $\theta_p \in [-\pi, \pi[$, $g_{\nu_{p^*}^n}(\theta)$ converge as $n \to \infty$ to (5.2). ☐

The proof of Theorem 5.1 is almost done. The uniform measure on $[-\pi, \pi[$ has a density of $\frac{1}{2\pi}$, that is $\lambda([-\pi, -\pi+t[) = \frac{t}{2\pi}$ for all $0 \le t < \pi$. Then Lemma 5.1 ensures that Statement 3. is equivalent to $G_{\mu_p}(\theta) > 0$ for $\theta \in [-\pi, 0[\cup]0, \pi[$ and thus Statement 2. □

5.3 Sufficient conditions of existence

In practice, it is convenient to derive simple sufficient conditions on probability measures to ensure the well definiteness of the Fréchet mean. The necessary and sufficient condition of Theorem 5.1 allows us to recover and extend the already known cases described Section 2.1. We also give a new criterion that does not restrict the support of the measure.

Bound on the support. A closed ball in $(\mathbb{S}^1, d_{\mathbb{S}^1})$ is the set $B(p_0, r) = \{p \in \mathbb{S}^1,\ d_{\mathbb{S}^1}(p, p_0) \le r\}$. The next result was proved in [3] for measure with finite support in $\mathbb{S}^n$, $n \in \mathbb{N}$. In $\mathbb{S}^1$, it is easy to extend the result to general measure.

Proposition 5.1. *Let μ be a probability measure on the circle with support included in a closed ball $B(\hat{p}, \frac{\pi}{2})$ centered at some $\hat{p} \in \mathbb{S}^1$ and of length $\frac{\pi}{2}$. If $\mu_{\hat{p}}(\{-\frac{\pi}{2}, \frac{\pi}{2}\}) < 1$ then, the Fréchet mean p^* of μ is well defined and belongs to $B(\hat{p}, \frac{\pi}{2})$.*

Proof. If p^* is a minimum of F_μ then p^* belongs to $B(\hat{p}, \frac{\pi}{2})$. To see this, we shortly present a reflection argument given in [3]. For any $p \in \mathbb{S}^1$ lying outside $B(\hat{p}, \frac{\pi}{2})$ consider $\bar{p}$, the symmetric of p with respect to the border of $B(\hat{p}, \frac{\pi}{2})$. In the normal coordinate system centered at $\hat{p}$, the point p satisfies $|\theta_p^{\hat{p}}| > \frac{\pi}{2}$ and $\theta_{\bar{p}}^{\hat{p}} = \pi - \theta_p^{\hat{p}}$ if $\theta_p^{\hat{p}}$ is positive or $-\pi - \theta_p^{\hat{p}}$ is $\theta_p^{\hat{p}}$ is negative. Assume now that $\frac{\pi}{2} < \theta_p^{\hat{p}} \le \pi$, the other case being similar. Since $\mu_{\hat{p}}(\{-\frac{\pi}{2}, \frac{\pi}{2}\}) < 1$, we have

$$\begin{aligned} F_\mu(p) &= \frac{1}{2}\int_{-\frac{\pi}{2}}^{\theta_p^{\hat{p}}} (\theta_p^{\hat{p}} - 2\pi - t)^2 d\mu_{\hat{p}}(t) + \frac{1}{2}\int_{\theta_p^{\hat{p}}}^{\frac{\pi}{2}} (\theta_p^{\hat{p}} - t)^2 d\mu_{\hat{p}}(t) \\ &> \frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\theta_p^{\hat{p}} - t)^2 d\mu_{\hat{p}}(t) = F_\mu(\bar{p}). \end{aligned}$$

Again with the condition $\mu_{\hat{p}}(\{-\frac{\pi}{2}, \frac{\pi}{2}\}) < 1$, we have $F_\mu(p) = F_\mu(\bar{p})$ if and only if $\theta_p^{\hat{p}} = \pm\frac{\pi}{2}$. Hence, we have proved that $|\theta_{p^*}^{\hat{p}}| < \frac{\pi}{2}$, i.e the argmin lies inside the open ball $\dot{B}(\hat{p}, \frac{\pi}{2})$.

Now remark that $F_{\mu_{\hat{p}}}$ is quadratic on $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and then achieves its unique minimum at $m(\mu_{\hat{p}})$. To conclude the proof, we take $p^* = e_{\hat{p}}(m(\mu_{\hat{p}}))$. □

It is possible to use the criterion of Theorem 5.1 to prove the preceding Proposition. Indeed, one can show that $G_{\mu_{p^*}}(\theta) > \frac{1}{2}(\theta + \pi - 2\theta_{\hat{p}})^2$ for $\theta \in [-\pi, \theta_{\hat{p}} - \frac{\pi}{2}[$, $G_{\mu_{p^*}} = \frac{\theta^2}{2}$ for $\theta \in [\theta_{\hat{p}} - \frac{\pi}{2}, \theta_{\hat{p}} + \frac{\pi}{2}[$ and $G_{\mu_{p^*}} > \frac{1}{2}(\theta - \pi - 2\theta_{\hat{p}})^2$ for all $\theta \in [\theta_{\hat{p}} + \frac{\pi}{2}, \pi[$. The case of equality corresponds to the distribution $\mu_{p^*} = (1-\varepsilon)\delta_{\theta_{\hat{p}}^{p^*} - \frac{\pi}{2}} + \varepsilon\delta_{\theta_{\hat{p}}^{p^*} + \frac{\pi}{2}}$ with $\varepsilon = \frac{\theta_{\hat{p}}^{p^*}}{\pi} + \frac{1}{2}$ and in this case, there are two global argmins at 0 and $2(\theta_{\hat{p}} \pm \frac{\pi}{2})$, see Figure 2.

Condition on the density. In this part we consider absolute continuous probability measure. First, we present a result for a subclass of radially distributed measures due to [11],

Proposition 5.2. *Let μ be a probability measure with density $f : \mathbb{S}^1 \longrightarrow \mathbb{R}$ which can be written as $f(p) = \rho(d(p, p^*))$ for some $p^* \in \mathbb{S}^1$ where $\rho : [0, \pi] \longrightarrow \mathbb{R}$ is a non constant decreasing function. Then p^* is the Fréchet mean of μ.*

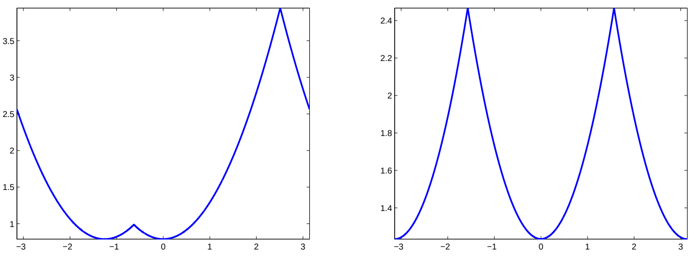

Figure 2: Let $\mu_\theta = (1-\varepsilon)\delta_{\theta-\frac{\pi}{2}} + \varepsilon\delta_{\theta+\frac{\pi}{2}}$ with $\varepsilon = \frac{\theta}{\pi} + \frac{1}{2}$. Left: F_{μ_θ} with $\theta = 0$. Right: F_{μ_θ} with $\theta = \frac{3\pi}{10}$

Proof. Consider the function $f_{p^*} = f \circ e_{p^*} : [-\pi, \pi[\longrightarrow \mathbb{R}$. By hypothesis, f_{p^*} is an even function decreasing on $[0, \pi[$ and increasing on $[-\pi, 0[)$. The point p^* is now clearly a critical point of F_μ as $m(\mu_{p^*}) = 0$, see Corollary 4.1. Let for all $\theta \in [-\pi, \pi[$

$$\begin{aligned} k_{\mu_{p^*}}(\theta) &= \mathbb{1}_{[0,\pi[}(\theta)\mu_{p^*}([-\pi, -\pi+\theta[) + \mathbb{1}_{[-\pi,0[}(\theta)\mu_{p^*}([\pi+\theta, \pi[) \\ &= \mathbb{1}_{[0,\pi[}(\theta)\int_0^\theta f_{p^*}(t-\pi)dt + \mathbb{1}_{[-\pi,0[}(\theta)\int_0^\theta -f_{p^*}(\pi+t)dt \end{aligned}$$

This function is even and convex as $t \mapsto f_{p^*}(t-\pi)$ and $t \mapsto -f_{p^*}(\pi+t)$ are increasing on $[0, \pi[$ and $[-\pi, 0[$ respectively. Now, the hypothesis on f implies that $k_{\mu_{p^*}}(-\pi) = \lim_{\theta\to\pi} k_{\mu_{p^*}}(\theta) = \frac{1}{2}$. Moreover, there is a $0 < t_0 < \pi$ such that $f(\theta) < \frac{1}{2\pi}$ for all $t_0 < \theta < \pi$ and $-\pi < \theta < -t_0$. This yields that $\frac{|t|}{2\pi} - k_{\mu_{p^*}}(t)$ is strictly positive for all $t \in]-\pi, 0[\cup]0, \pi[$ and vanishes at $t = 0$ and $t = -\pi$. Then, property 3 of Theorem 5.1 is satisfied and the claim is proved. ☐

A lot of classical probability distributions used in circular data analysis follow the hypothesis of Lemma 5.2: von Mises distribution, wrapped normal distribution, geodesic normal distribution [19] , etc... In the rest of this section, we present a concentration criterion that do not impose bounds on the support nor the symmetry of the density. Let us introduce the following definition :

Definition 5.1. *Let $f : \mathbb{S}^1 \longrightarrow \mathbb{R}^+$ be a probability density, $p \in \mathbb{S}^1$, $\alpha \in]0, 1]$ and $\varphi \in]0, \pi[$. We say that f satisfies the property $P(p, \alpha, \varphi)$ if for all $|\theta| \geq \varphi$*

$$\frac{1}{2\pi} - f_p(\theta) \geq \frac{\alpha}{2\pi}, \tag{5.3}$$

where $f_p = f \circ e_p : [-\pi, \pi[\longrightarrow \mathbb{R}^+$. Moreover, we say that $f \in P(\alpha, \varphi)$ if there is a $p \in \mathbb{S}^1$ such that f satisfies $P(p, \alpha, \varphi)$.

The parameters α and φ control the concentration of μ around p. The idea is to control the mass lying in the complementary of the ball $B(p, \varphi)$. Note that is equivalent to the fact that $f_p(\theta) \leq \frac{1-\alpha}{2\pi}$ for $|\theta| \geq \varphi$. We have the following properties :

Lemma 5.2. *Let $f : \mathbb{S}^1 \longrightarrow \mathbb{R}$ be a probability density on the circle. Then*

1. *$P(p, \alpha_1, \varphi_1) \Longrightarrow P(p, \alpha_2, \varphi_2)$ if $\alpha_1 \geq \alpha_2$ and $\varphi_1 \leq \varphi_2$.*
2. *Let $p_1, p_2 \in \mathbb{S}^1$ and $\varphi < \frac{\pi}{2}$. If $d_{\mathbb{S}^1}(p_1, p_2) < \pi - \varphi$ then $P(p_1, \alpha, \varphi) \Longrightarrow P(p_2, \alpha, \varphi + d_{\mathbb{S}^1}(p_1, p_2))$.*
3. *If f satisfies $P(p, \alpha, \varphi)$ then $|m(\mu_p)| \leq \varphi + \frac{1-\alpha}{4\pi}(\pi - \varphi)^2$.*

Proof. The first proposition is obvious in view of the Definition 5.1.

To prove the second claim suppose that $0 \leq \theta_{p_2}^{p_1} \leq \pi - \varphi$ (the other case is similar) and write

$$f_{p_1}(\theta) = f_{p_2}(\theta - \theta_{p_2}^{p_1})\mathbb{1}_{[\theta_{p_2}^{p_1}-\pi,\pi[}(\theta) + f_{p_2}(\theta - \theta_{p_2}^{p_1} + 2\pi)\mathbb{1}_{[-\pi,\theta_{p_2}^{p_1}-\pi[}(\theta).$$

In particular, it implies that $\theta_{p_2}^{p_1} - \pi \leq -\varphi$ and since $P(p_1, \alpha, \varphi)$ holds, we have $f_{p_2}(t) \leq \frac{1-\alpha}{2\pi}$, if $t \geq \varphi - \theta_{p_2}^{p_1}$ or $t \leq -\varphi - \theta_{p_2}^{p_1}$. This is equivalent to the fact that $P(p_2, \alpha, \min\{| - \varphi - \theta_{p_2}^{p_1}|, |\varphi - \theta_{p_2}^{p_1}|\}) = P(p_2, \alpha, \varphi + \theta_{p_2}^{p_1})$ holds. The case $\varphi - \pi \leq \theta_{p_2}^{p_1} \leq 0$ is similar and we have $P(p_2, \alpha, \varphi - \theta_{p_2}^{p_1})$. Finally recall that $|\theta_{p_2}^{p_1}| = d_{\mathbb{S}^1}(p_1, p_2)$ and the property is proved.

For the last proposition we show that the upper bound is attained for $\mu_p = (1 - \frac{\pi-\varphi}{2\pi})\delta_\varphi + (1 - \alpha)\mathbb{1}_{[\varphi,\pi[}d\lambda$, i.e 'push the mass as far as possible'. First, it is necessary to consider only the case where μ_p has its support on $[0, \pi[$. Indeed, $\mu_p = \omega\mu_p^- + (1 - \omega)\mu_p^+$ where $\mu_p^-([-\pi, 0[) = \mu_p^+([0, \pi[) = 1$ and $0 \leq \omega \leq 1$. It yields that $m(\mu_p) = \int_{\mathbb{R}} td(\omega\mu_p^- + (1 - \omega)\mu_p^+) = \int_{\mathbb{R}^+} td(-\omega\mu_p^- + (1 - \omega)\mu_p^+) \leq \int_{\mathbb{R}^+} td\mu_p^+$. Then, if the density f_p of μ_p has its support in $[0, \pi[$ we have

$$\begin{aligned} m(\mu_p) &\leq \varphi \left(1 - \int_\varphi^\pi f_p(t)dt\right) + \int_\varphi^\pi tf_p(t)dt \\ &= \varphi + \int_\varphi^\pi (t - \varphi)f_p(t)dt \leq \varphi + \frac{1-\alpha}{2\pi}\int_\varphi^\pi (t - \varphi)dt, \end{aligned}$$

which gives the result. □

If the density f is sufficiently concentrated around a critical point p^* of F_μ then, this point is the Fréchet mean of μ. More precisely we have the following,

Proposition 5.3. *Let μ be a probability measure with density $f : \mathbb{S}^1 \longrightarrow \mathbb{R}^+$. Consider a critical point p^* of F_μ, i.e. a point satisfying $m(\mu_{p^*}) = \int_{-\pi}^{\pi} tf_{p^*}(t)dt = 0$. If f satisfies $P(p^*, \alpha, \varphi)$ with $\alpha \in]0, 1]$ and $0 < \varphi < \varphi_\alpha = \pi\frac{\sqrt{\alpha}}{1+\sqrt{\alpha}}$ then, μ admits a well defined Fréchet mean at p^*.*

Before the proof we make two remarks about Proposition 5.3. For all $\alpha \in]0, 1]$ we have $\varphi_0 = 0 \leq \varphi_\alpha < \frac{\pi}{2} = \varphi_1$. Note that if $\alpha = 1$ then μ has its support included in the ball $B(p^*, \frac{\pi}{2})$. When $\alpha < 1$ the support of μ can be the entire circle $\mathbb{S}^1$. Now if α is small, the density f_{p^*} is allowed to approach 'from below' the uniform density on $[-\pi, \pi[\backslash[-\varphi_\alpha, \varphi_\alpha]$ and f_{p^*} can be greater than $\frac{1}{2\pi}$ only in the 'small' interval $[-\varphi_\alpha, \varphi_\alpha]$.

This result can also be used to generate absolutely continuous probability distribution on the circle with a given Frechet mean. Proceed as follows : fix a $p^* \in \mathbb{S}^1$ and choose a function $f_{p^*} : [-\pi, \pi[\longrightarrow \mathbb{R}^+$ satisfying $\int_{-\pi}^{\pi} f_{p^*}(t)dt = 1$, $\int_{-\pi}^{\pi} tf_{p^*}(t)dt = 0$ and such that equation (5.3) holds for $\alpha \in]0, 1]$ and $\varphi_\alpha = \pi\frac{\sqrt{\alpha}}{1+\sqrt{\alpha}}$. Then the probability measure μ on the circle with density $f = f_{p^*} \circ e_{p^*}^{-1}$ has a Fréchet mean at p^*.

Proof. As the measure μ admits a density, F_μ is twice differentiable and p^* is a critical point of F_μ since $m(\mu_{p^*}) = 0$, see Corollary 4.1. Moreover Lemma 5.1 ensures that the second derivative of $F_{\mu_{p^*}}$ is equal to $\frac{d^2}{d\theta^2}G_{\mu_{p^*}}(\theta) = 1 - 2\pi f(-\pi + \theta)$, if $0 \leq \theta < \pi$, and $\frac{d^2}{d\theta^2}G_{\mu_{p^*}}(\theta) = 1 - 2\pi f(\pi + \theta)$, if $-\pi \leq \theta < 0$. Thus by hypothesis, the function $F_{\mu_{p^*}}$ is convex on $[-\pi + \varphi, \pi - \varphi]$ and has a unique minimum at 0. Let us show that it is the only argmin of $F_{\mu_{p^*}}$. Let $\theta \in [\pi - \varphi, \pi[$, we have

$$G_{\mu_{p^*}}(\theta) = G_{\mu_{p^*}}(\pi - \varphi) + \int_{\pi-\varphi}^{\theta} t - 2\pi\mu_{p^*}([-\pi, -\pi + t[)dt$$

By hypothesis f satisfies $P(p^*, \alpha, \varphi)$ which implies that $G_{\mu_{p^*}}(\pi - \varphi) \geq \frac{\alpha}{2}(\pi - \varphi)^2$. The second term is bounded from below by $\int_{\pi-\varphi}^{\theta} t - 2\pi\nu([-\pi, -\pi + t[)dt$ where $\nu = \frac{1}{2}(\delta_{\pi-\varphi} + \delta_{\varphi-\pi})$, that is

$$\begin{aligned} G_{\mu_{p^*}}(\theta) &\geq \frac{\alpha}{2}(\pi - \varphi)^2 + \int_{\pi-\varphi}^{\theta} (t - \pi)dt \\ &\geq \frac{1}{2}((\alpha - 1)(\pi - \varphi)^2 + 2\pi(\pi - \varphi) - \pi^2) \end{aligned} \quad (5.4)$$

The right hand side of the preceding inequality is strictly positive if $\varphi < \varphi_\alpha = \pi\frac{\sqrt{\alpha}}{1+\sqrt{\alpha}}$. Similarly, for $\theta \in [-\pi, -\pi+\varphi[$ $G_{\mu_{p^*}}(\theta) > 0$. The result now follows by Theorem 5.1. □

To use the preceding Proposition, we need to know a critical point p^* of F_μ, that is we need to localize the Fréchet mean. This condition is not very realistic in practice and to relax it, one needs more concentration than in the preceding Proposition. Indeed, if there is a $p \in \mathbb{S}^1$ with f_p sufficiently concentrated, then there is a critical point in the neighborhood of p. Moreover, this critical point is the Fréchet mean of μ.

Theorem 5.2. *Let* $\delta \in]0, \frac{1}{2}[$ *and* α_δ *be the square of the root of* $(5-6\delta+\delta^2)X^3+(1-\delta^2)X^2-(2\delta+1)X-1$ *that lies in* $]0,1]$. *If* μ *is a probability measure with density* $f \in P(\alpha,\varphi)$ *(see Definition 5.1) with* $\alpha_\delta \le \alpha \le 1$ *and* $\varphi \le \delta\varphi_\alpha = \delta\pi\frac{\sqrt{\alpha}}{1+\sqrt{\alpha}}$ *then* μ *admits a well defined Fréchet mean.*

This result gives a functional class of densities that admit a well defined Fréchet mean. The parameter δ controls the concentration of f via the inequality $\alpha_\delta < \alpha$ and $\varphi \le \delta\varphi_\alpha$. There is a tradeoff between α and the possible value of φ. The smaller α is (i.e the less f is concentrated) the smaller φ must be (i.e we need to control the value of the density on a bigger interval). As a typical example, take $\delta = \frac{1}{3}$. In this case $\alpha_\delta = \alpha_{\frac{1}{3}} < 0.69$ and $\delta\varphi_{\alpha_\delta} = \frac{1}{3}\varphi_{\alpha_{\frac{1}{3}}} \le 0.47$. If $f \in P(\alpha,\varphi)$ with $0.69 < \alpha \le 1$ and $\varphi \le 0.48$ then there is a well defined Fréchet mean. In Tabular 1 we give some numerical values. Note that the column corresponding to $\delta = 0$ is given as a reference only as the set $P(\alpha_\delta, \delta\varphi_\alpha)$ is empty for this values of δ.

$\delta =$	0	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{1}{2}$
$\alpha_\delta \le$	0.39	0.46	0.54	0.69	1
$\delta\varphi_{\alpha_\delta} \ge$	0	0.12	0.26	0.47	$\frac{\pi}{4}$

Table 1: Some values of α_δ and $\delta\varphi_{\alpha_\delta}$ depending on $\delta \in]0, \frac{1}{2}[$.

Proof. If we show that under the hypothesis of the Theorem, there is a critical point p^* of F_μ satisfying $d_{\mathbb{S}^1}(p,p^*) \le (1-\delta)\varphi_\alpha$ where p is a point such that f satisfies $P(p,\alpha,\varphi)$, then, by Lemma 5.2, f will satisfy $P(p^*,\alpha,\delta\varphi_\alpha+(1-\delta)\varphi_\alpha) = P(p^*,\alpha,\varphi_\alpha)$ and Proposition 5.3 will ensure that p^* is the Fréchet mean of μ.

Hence the rest of the proof is devoted to show that there is a $p^* \in \mathbb{S}^1$ such that $\frac{d}{d\theta}F_{\mu_p}(\theta^p_{p^*}) = 0$ with $d_{\mathbb{S}^1}(p,p^*) \le (1-\delta)\varphi_\alpha$. Suppose that $m(\mu_p) \ge 0$ (the case $m(\mu_p) < 0$ is similar). We have

$$\frac{d}{d\theta}F_{\mu_p}(0) = -m(\mu_p) \le 0.$$

Then, remark that

$$\frac{d}{d\theta}F_{\mu_p}((1-\delta)\varphi_\alpha) = (1-\delta)\varphi_\alpha - 2\pi\mu_p([-\pi, -\pi+(1-\delta)\varphi_\alpha[) - m(\mu_p).$$

We have $-2\pi\mu_p([-\pi,-\pi+(1-\delta)\varphi_\alpha[) \ge (\alpha-1)(1-\delta)\varphi_\alpha$ since f satisfies the $P(p,\alpha,\delta\varphi_\alpha)$ and that $|-\pi+(1-\delta)\varphi_\alpha| \ge \delta\varphi_\alpha$. Moreover, $-m(\mu) \ge -|m(\mu_p)|$ which is controlled Lemma 5.2 Statement 3. It gives,

$$\begin{aligned}\frac{d}{d\theta}F_{\mu_p}((1-\delta)\varphi_\alpha) &\ge (1-\delta)\alpha\varphi_\alpha - \delta\varphi_\alpha - \frac{1-\alpha}{4\pi}(\pi-\delta\varphi_\alpha)^2 \\ &= \pi\frac{(5-6\delta+\delta^2)\alpha\sqrt{\alpha}+(1-\delta^2)\alpha-(2\delta+1)\sqrt{\alpha}-1}{1+\sqrt{\alpha}}\end{aligned}$$

This quantity is positive as soon as $1 \geq \alpha > \alpha_\delta$, where $\sqrt{\alpha_\delta}$ is the root of the polynomial $X \longmapsto (5-6\delta+\delta^2)X^3+(1-\delta^2)X^2-(2\delta+1)X-1$ that lies in $]0,1]$. Numerical experiment shows that the function $\delta \longmapsto \alpha_\delta$ takes its value in $]0.39,1[$ for $\delta \in]0,\frac{1}{2}[$. Nevertheless it is easy to see that $\alpha_{\frac{1}{2}} = 1$. Since the derivative of F_μ is continuous, the intermediate value Theorem ensures that there is a critical point such that $|\theta^p_{p^*}| \leq (1-\delta)\varphi_\alpha$. This is what we need to complete the proof. □

6 Fréchet mean of an empirical measure

6.1 Consistency of the empirical Fréchet mean

Let $X_1,\dots,X_n$ be independent and identically distributed random variables with value in $(\mathbb{S}^1, d_{\mathbb{S}^1})$ and of probability distribution μ. The empirical measure is defined as usual by $\mu^n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$. Following [2], we call empirical Fréchet mean set the set of argmins of

$$p \longmapsto F_{\mu^n}(p) = \frac{1}{2n}\sum_{i=1}^n d^2_{\mathbb{S}^1}(p, X_i), \qquad p \in \mathbb{S}^1.$$

If the argmin of F_{μ^n} is unique it is called the empirical Fréchet mean, and will be denoted by p^*_n. In [20] a strong law of large number is given for the empirical Fréchet mean in a semi metric space which is the case of $(\mathbb{S}^1, d_{\mathbb{S}^1})$. If μ admits a well defined Fréchet mean, any measurable choice in the empirical Fréchet mean set of μ^n is a consistent estimator of p^*. In particular if p^*_n exists for each $n \in \mathbb{N}$, the empirical Fréchet mean is a consistent estimator of the Fréchet mean. Nevertheless, the empirical Fréchet mean is well defined almost surely for a wide class of probability measures. The following fact is from [2] Remark 2.6 :

Lemma 6.1. *Let μ be a non atomic probability measure on the circle, i.e satisfying $\mu(\{p\}) = 0$ for all $p \in \mathbb{S}^1$. Then for all $n \in \mathbb{N}$ the empirical Fréchet mean exists almost surely.*

Lemma 6.1 shows that the empirical Fréchet mean p^*_n of a probability measure can be computed even if this measure does not possess a well defined Fréchet mean. If the Fréchet mean p^* of μ is well defined, we study the rate of convergence of the empirical Fréchet mean p^*_n to p^*. To this end, we derive a concentration inequality that shows the consistency of the empirical Fréchet mean.

Proposition 6.1. *Let μ be a measure with density $f : \mathbb{S}^1 \longrightarrow \mathbb{R}$ that admits a well defined Fréchet mean p^*. Then there exists a strictly increasing function $\rho : [0,\pi[\longrightarrow \mathbb{R}^+$ such that for all $p \in \mathbb{S}^1$*

$$F_\mu(p) \geq \rho(d_{\mathbb{S}^1}(p, p^*)).$$

If p^*_n denotes the empirical Fréchet mean, we have for all $x > 0$

$$\mathbb{P}\left(\rho(d_{\mathbb{S}^1}(p^*_n, p^*)) \geq C(s)\sqrt{\frac{x}{n}}\right) \leq 2e^{-x}.$$

where $s = \max\{|x-y|\,,\ x,y \in \mathrm{support}(\mu)\}$ and $C(s) = (4\pi^2 + 4\pi^2 s + 2s) \leq 4\pi(2\pi^2+\pi+1)$.

The function ρ in the statement of the preceding Proposition 6.1 determines the rate of convergence of p^*_n to p^*. Indeed, the rate of convergence of p^*_n to p^* will depend on how fast $\rho^{-1}(t)$ is going to 0 when $t \to 0$. For example, when the support of μ is strictly included in an hemisphere, then ρ is a polynomial of order 2. Below, we will see that under the assumptions of Theorem 5.2 we keep similar asymptotic rates of convergence but without this bound on the support.

Proof. The first claim about the lower bound ρ is a direct consequence of the following Lemma:

Lemma 6.2. *Let $f : [-\pi, \pi[\longrightarrow \mathbb{R}^+$ be a continuous function on $[-\pi, \pi[$ and which satisfies $\lim_{\theta \to \pi^-} f(\theta) = \lim_{\theta \to -\pi^+} f(\theta) = f(-\pi)$. Suppose that f vanishes at a unique point $\theta_0 \in [-\pi, \pi[$. Then, there exists a strictly increasing function $\rho : [0, \pi[\longrightarrow \mathbb{R}^*$ such that for all $\theta \in [-\pi, \pi[$,*

$$f(\theta) \geq \rho(d_{\mathbb{S}^1}(\theta_0, \theta)).$$

Proof. As the function f is periodic we can assume, without loss of generality, that $\theta_0 = 0$. Define the function $g : [-\pi, \pi[\longrightarrow \mathbb{R}^+$ by

$$g(\theta) = \min \left\{ f(-\theta), f(\theta), \min_{|t| > \theta} f(t) \right\}.$$

Remark that the considered minima are attained as f is a continuous function on $\mathbb{R}/(2\pi\mathbb{Z})$ which is compact. We have $g \leq f$. The function g is even, is increasing on $[0, \pi[$ and vanishes only in 0. We now have to construct an even, positive, strictly increasing function that bounds g from below. Consider for all $\theta \in [-\pi, \pi[$,

$$G(\theta) = \frac{1}{|\theta|} \int_0^\theta g(t) dt.$$

We have $G(\theta) \leq g(\theta)$ for all $\theta \in [0, \pi[$ since g is an increasing function on $[0, \pi[$. Moreover G is even and we have $G \leq g$ on $[-\pi, \pi[$. The function G is strictly increasing since its derivative and $g - G$ have the same sign. Now $g(\theta) - G(\theta) = 0$ if and only if $g(t) = g(0)$ for all $t \in]0, \theta]$ which is impossible by the construction of g. We conclude the proof of Lemma 6.2 by setting $G(\theta) = \rho(|\theta|)$. □

We now focus on the concentration inequality of Proposition 6.1. The proof is divided in two steps. First we show the uniform convergence in probability of $F_{\mu_p^n}$ to F_{μ_p}. Then, we deduce the convergence of their argmins by using the lower bound given by ρ. Recall the notations, $p^* = \operatorname{argmin}_{p \in \mathbb{S}^1} F_\mu(p)$ and $p_n^* = \operatorname{argmin}_{p \in \mathbb{S}^1} F_{\mu^n}(p)$. We fix an arbitrary $p \in \mathbb{S}^1$ and in normal coordinate centered at p we have for all $\theta \in [-\pi, \pi[$

$$F_{\mu_p}(\theta) = \int_{\mathbb{R}} d^2_{\mathbb{S}^1}(\theta, t) d\mu_p(t) \qquad \text{and} \qquad F_{\mu_p^n}(\theta) = \int_{\mathbb{R}} d^2_{\mathbb{S}^1}(\theta, t) d\mu_p^n(t).$$

Let $\theta_{p^*} = \theta^p_{p^*} = \operatorname{argmin}_{\theta \in \mathbb{R}} F_{\mu_p}(\theta)$ and $\theta_{p_n^*} = \theta^p_{p_n^*} = \operatorname{argmin}_{\theta \in \mathbb{R}} F_{\mu_{pn}}(\theta)$.

For all $\theta \in [-\pi, \pi[$, let $H(\theta) = \mu_p([-\pi, \theta[)$ be the cumulative distribution function (c.d.f) of μ_p and $H_n(\theta) = \mu^n([-\pi, \theta[)$ be and the empirical c.d.f. Recall Proposition 4.1 where we have shown that $\frac{d}{d\theta} F_{\mu_p}(\theta) = \theta - 2\pi H(-\pi + \theta) - m(\mu_p)$ and $\frac{d}{d\theta} F_{\mu_p^n}(\theta) = \theta - 2\pi H^n(-\pi + \theta) - m(\mu_p^n)$ if $\theta \in [0, \pi[$ and $\frac{d}{d\theta} F_{\mu_p}(\theta) = \theta + 2\pi - 2\pi H(\pi + \theta) - m(\mu_p)$ and $\frac{d}{d\theta} F_{\mu_p^n}(\theta) = \theta + 2\pi - 2\pi H^n(\pi + \theta) - m(\mu_p^n)$ if $\theta \in [-\pi, 0[$. Then,

$$\begin{aligned} 2 \sup_{\theta \in \mathbb{R}} \left| F_{\mu_p^n}(\theta) - F_{\mu_p}(\theta) \right| &\leq 2\pi \sup_{\theta \in \mathbb{R}} \left| \frac{d}{d\theta} F_{\mu_p^n}(\theta) - \frac{d}{d\theta} F_{\mu_p}(\theta) \right| + 2 \left| F_{\mu_p^n}(0) - F_{\mu_p}(0) \right| \\ &\leq 4\pi^2 \sup_{\theta \in [-\pi, \pi[} |H(\theta) - H^n(\theta)| + 4\pi^2 \left| m(\mu_p) - m(\mu_p^n) \right| \qquad (6.1) \\ &\quad + 2 \left| m_2(\mu_p) - m_2(\mu_p^n) \right| \qquad (6.2) \end{aligned}$$

where $m_2(\nu) = \int_{\mathbb{R}} t^2 d\nu(t)$, for a measure ν on $\mathbb{R}$. The first term of the preceding upper bound can be controlled in probability using the Dvoretzky-Kiefer-Wolfowitz inequality (see e.g [17]), we have for all $x > 0$,

$$\mathbb{P} \left(4\pi^2 \sup_{\theta \in [-\pi, \pi[} |H(\theta) - H^n(\theta)| \geq 4\pi^2 \sqrt{\frac{x}{n}} \right) \leq 2e^{-x}.$$

For the second and third term which involve the first and second moment of μ_p and μ_p^n, we use an Hoeffding type inequality which gives for all $x > 0$, $\mathbb{P}\Big(4\pi^2 \left| m(\mu_p) - m(\mu_p^n) \right| + 2 \left| m_2(\mu_p) - m_2(\mu_p^n) \right| \geq$

$s(4\pi^2+2s)\sqrt{\frac{x}{n}}\Big) \leq 2e^{-x}$, where $s=\max\{|x-y|\,,\ x,y\in \text{support }\mu\}$ is the diameter of the support of μ. Combining the two concentration inequalities and (6.2), it yields for all $x>0$,

$$\mathbb{P}\left(2\sup_{\theta\in\mathbb{R}}\left|F_{\mu_p}(\theta)-F_{\mu_p^n}(\theta)\right| \geq (4\pi^2+4\pi^2 s+2s)\sqrt{\frac{x}{n}}\right) \leq 2e^{-x}. \tag{6.3}$$

Now that the uniform convergence in probability is shown, we use a classical inequality in M-estimation, $F_{\mu_p}(\theta_{p_n^*})-F_{\mu_p}(\theta_{p^*}) \leq 2\sup_{\theta\in\mathbb{R}}\left|F_{\mu_p^n}(\theta)-F_{\mu_p}(\theta)\right|$. By Lemma 6.2, there exists an increasing function $\rho:\mathbb{R}^+\longrightarrow\mathbb{R}^+$ such that $F(\theta_{p_n^*})-F(\theta_{p^*})\geq \rho(d_{\mathbb{S}^1}(\theta_{p_n^*},\theta_{p^*}))$. Plugging this in equation (6.3) we have,

$$\mathbb{P}\left(\rho(d_{\mathbb{S}^1}(\theta_{p_n^*},\theta_{p^*}))\geq (4\pi^2+4\pi^2 s+2s)\sqrt{\frac{x}{n}}\right)\leq 2e^{-x},$$

and the proof of Proposition 6.1 is completed. □

Function ρ that appears in the statement of Proposition 6.1 can be explicitly computed if the density $f:\mathbb{S}^1\longrightarrow\mathbb{R}^+$ satisfies property $P(p,\alpha,\varphi)$ for some $p\in\mathbb{S}^1$, see Definition 5.1. The parameter $\alpha\in]0,1]$ can be interpreted as a measure of the convexity of F_{μ_p} on the interval $[-\varphi,\varphi]$. For example, if $\alpha=1$ and $\varphi=\varphi_\alpha=\frac{\pi}{2}$, then μ has its support contained in $[-\frac{\pi}{2},\frac{\pi}{2}]$ and F_{μ_p} is quadratic on $[-\frac{\pi}{2},\frac{\pi}{2}]$ with a second derivative equals to 1.

Proposition 6.2. *Let μ be a probability measure with density $f:\mathbb{S}^1\longrightarrow\mathbb{R}^+$ satisfying the hypothesis of Theorem 5.2. Then, for all $x>0$ we have*

$$\mathbb{P}\left(d_{\mathbb{S}^1}(p_n^*,p^*)\geq\sqrt{B(\alpha,\varphi)}\left(\frac{x}{n}\right)^{\frac{1}{4}}\right)\leq 2e^{-x},$$

where $B(\alpha,\varphi)=C\max\left\{\frac{\pi^2}{\gamma(\alpha,\varphi)},\frac{2}{\alpha}\right\}$ with $\gamma(\alpha,\varphi)=\frac{1}{2}((\alpha-1)(\pi-\varphi)^2+2\pi(\pi-\varphi)-\pi^2)$ and $C=4\pi(2\pi^2+\pi+1)$.

Proof. This result follows from Proposition 6.1 and we only have to find a strictly increasing function $\rho:[0,\pi]\longrightarrow\mathbb{R}^+$ satisfying for all $\theta\in[-\pi,\pi[$, $F_\mu(p)-F_\mu(p^*)\geq\rho(d_{\mathbb{S}^1}(p,p^*))$. By Theorem 5.2, the Fréchet mean p^* of μ exists and Lemma 6.2 ensures that there is a stricly increasing function ρ which satisfies,

$$F_{\mu_{p^*}}(\theta)-F_{\mu_{p^*}}(0)\geq\rho(|\theta|). \tag{6.4}$$

for all $\theta\in[-\pi,\pi[$. As μ_{p^*} admits a density f_{p^*}, the Fréchet functional $F_{\mu_{p^*}}$ is twice differentiable. Moreover f satisfies $P(p^*,\alpha,\varphi_\alpha)$, see proof of Theorem 5.2. For all $\theta\in[-\pi+\varphi_\alpha,\varphi_\alpha-\pi]$, a second order Taylor expansion of $F_{\mu_{p^*}}$ at 0 ensures that for some $\tilde{\theta}\in[-\pi+\varphi_\alpha,\pi-\varphi_\alpha]$,

$$F_{\mu_{p^*}}(\theta)-F_{\mu_{p^*}}(0)=\frac{1}{2}\theta^2\frac{d^2}{d\theta^2}F_{\mu_{p^*}}(\tilde{\theta})\geq\frac{\alpha}{2}\theta^2.$$

The last inequality is a direct consequence of property $P(p^*,\alpha,\varphi_\alpha)$ as we have $\frac{d^2}{d\theta^2}F_{\mu_p}(\theta)=1-2\pi f(-\pi+\theta)$, if $0\leq\theta<\pi$ and $\frac{d^2}{d\theta^2}F_{\mu_p}(\theta)=1-2\pi f(\pi+\theta)$, if $-\pi\leq\theta<0$. For all $\theta\in[-\pi,-\varphi[\cup]\varphi,\pi[$ we have by inequality (5.4),

$$F_{\mu_{p^*}}(\theta)-F_{\mu_{p^*}}(0)\geq\frac{1}{2}((\alpha-1)(\pi-\varphi)^2+2\pi(\pi-\varphi)-\pi^2)=\gamma(\alpha,\varphi)>0.$$

Then, let $\rho(t)=t^2\min\{\frac{\gamma(\alpha,\varphi)}{\pi^2},\frac{\alpha}{2}\}$ and the proof is completed.

□

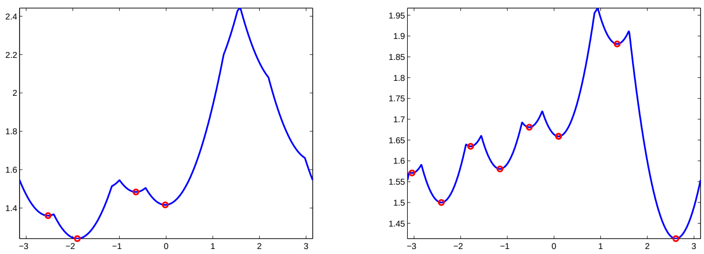

Figure 3: Plots of F_{μ^n} where $n = 10$ and $\mu^n = \sum_{i=1}^n \delta_{X_i}$ where the X_i i.i.d of uniform law λ. The red points are the local minima computed with the described algorithm.

6.2 Computation of the Empirical Fréchet mean

Computation of the Fréchet mean of a general probability measure may not be an easy task as it is a *global* optimization problem. In practice the Fréchet functional is not a convex function and a gradient descent algorithm will only give a *local* minimum which depends on the initialization point chosen.

In the following we will use the results of section 4.2 to derive an algorithm to compute the empirical Fréchet mean. Recall that the regular critical points (i.e no cusp point) of F_μ are the local minima of F_μ. Moreover, Corollary 4.3 gives us a simple mean to compute them by solving at most n affine relations, see equation (4.6). In a coordinate system centered at some $p \in \mathbb{S}^1$, it amounts to compute the cumulative distribution function of μ^n which is, here, piecewise constant with jumps of size $\frac{1}{n}$. Indeed, we have,

$$\mu_p^n([-\pi, t[) = \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{[-\pi,t[}(\theta_{X_i}^p) = \frac{1}{n}\mathrm{Card}\{\theta_{X_i}^p < t\}.$$

Note, that in practice, there are less than n solutions, see e.g. Figure 3.

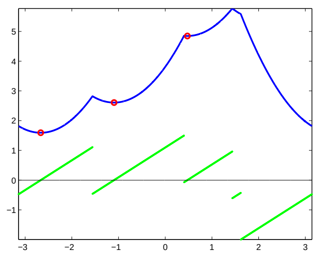

Figure 4: In blue: plot of F_{μ^n} where $n = 4$. In green: the derivative of F_{μ^n}. The critical points are given by the intersection between the green curve and the x-axis in black.

The following algorithm takes as input the values $\{X_i\}_{i=1}^n$ and returns the Fréchet mean of $\mu^n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$. See also Figure 4 for an illustration.

Initialization Step : Choose an arbitrarily point $p \in \mathbb{S}^1$.
Compute the coordinates $\{\theta_{X_i}^p\}_{i=1}^n$ and reorder them is increasing order. We denote $\tau_0^- = -\pi \le \tau_1^- \le \tau_2^- \le \ldots \le \tau_{n_1}^- < 0 = \tau_{n_1+1}^-$ the n_1 negative sorted terms and $\tau_0^+ = \pi > \tau_1^+ \ge \ldots \ge \tau_{n_2}^+ \ge 0 = \tau_{n_2+1}^+$ the $n_2 = n - n_1$ positive sorted terms.
Compute the mean $m(\mu_p^n) = \frac{1}{n}(\tau_1^- + \ldots + \tau_{n_1}^- + \tau_1^+ + \ldots + \tau_{n_2}^+)$ and initialize θ_{p*}^p to 0, says.

`# The first step compares all the local minima in` $[0, \pi[$

Step 1: For i from 0 to n_1 do
`#` $\theta^p_{p^*,new}$ `is the candidate to be a critical point between` τ_i^- `and` τ_{i+1}^-
Let $\theta^p_{p^*,new} = 2\pi\frac{i}{n} + m(\mu_p^n)$
`# verify if` $\theta^p_{p^*,new}$ `is a critical point and then test its value. If it is better, keep it.`
if $\tau_i^- + \pi \leq \theta^p_{p^*,new} \leq \tau_{i+1}^- + \pi$ and $F_{\mu_p^n}(\theta^p_{p^*,new}) \leq F_{\mu_p^n}(\theta^p_{p^*})$ then $\theta^p_{p^*} := \theta^p_{p^*,new}$ end if
end for.
`# The Step 2 is the same as Step 1 but for local minima in` $[-\pi, 0[$
Step 2: For i = 0 to n_2 do
Let $\theta^p_{p^*,new} = -2\pi\frac{i}{n} + m(\mu_p^n)$
if $\tau_{i+1}^+ - \pi \leq \theta^p_{p^*,new} \leq \tau_i^+ - \pi$ and $F_{\mu_p^n}(\theta^p_{p^*,new}) \leq F_{\mu_p^n}(\theta^p_{p^*})$ then $\theta^p_{p^*} := \theta^p_{p^*,new}$ end if
end for.
`# The value of` $\theta^p_{p^*}$ `is the best argmin`
Output Return $p^* = e_p(\theta^p_{p^*})$.

This algorithm can be extended to more general measures that the empirical one. The approach will be the same: find the critical points of the Fréchet functional with formula of Corollary 4.3. Unfortunately, there may be some computational issues as general cumulative distribution function will be not piecewise constant anymore.

7 Conclusion

It is not straightforward to extend criterion such as the one given in Theorem 5.1 to more general spaces, e.g. for the n dimensional sphere $\mathbb{S}^n$. Recall that the circle $\mathbb{S}^1$ is a flat space in the sense that it is locally isometric to the Euclidean space $\mathbb{R}$. Then, the only phenomenon that induces uniqueness issues of the Fréchet mean is the presence of a cut locus. The criterion presented in this note relies on an explicit formula for the gradient of the Fréchet mean. Curvature has an extra effect on the metric and makes difficult to derive exact computation on the Fréchet functional and its gradient. Moreover, it is not clear if the role played by the uniform measure as a benchmark in the well definiteness of the Fréchet mean in $\mathbb{S}^1$ can be extended to n-spheres or non flat manifolds.

Acknowledgements - The author would like to thank Dominique Bakry for help and encouragement during the writing of this paper and Jérémie Bigot for his careful reading of the manuscript. The author acknowledges the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-JCJC-SIMI1 DEMOS.

References

[1] Bijan Afsari. Riemannian L^p center of mass: existence, uniqueness, and convexity. *Proc. Amer. Math. Soc.*, 139(2):655–673, 2011.

[2] Rabi Bhattacharya and Vic Patrangenaru. Large sample theory of intrinsic and extrinsic sample means on manifolds. i. *The Annals of Statistics*, 31(1):pp. 1–29, 2003.

[3] Samuel R. Buss and Jay P. Fillmore. Spherical averages and applications to spherical splines and interpolation. *ACM Trans. Graph.*, 20:95–126, April 2001.

[4] José Manuel Corcuera and Wilfrid S. Kendall. Riemannian barycentres and geodesic convexity. *Math. Proc. Cambridge Philos. Soc.*, 127(2):253–269, 1999.

[5] I. L. Dryden and K. V. Mardia. *Statistical shape analysis.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1998.

[6] Michel Émery and G. Mokobodzki. Sur le barycentre d'une probabilité dans une variété. In *Séminaire de Probabilités, XXV*, volume 1485 of *Lecture Notes in Math.*, pages 220–233. Springer, Berlin, 1991.

[7] N. I. Fisher. *Statistical analysis of circular data*. Cambridge University Press, Cambridge, 1993.

[8] Maurice Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H. Poincaré*, 10:215–310, 1948.

[9] Sigurdur Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 34 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Corrected reprint of the 1978 original.

[10] H. Karcher. Riemannian center of mass and mollifier smoothing. *Comm. Pure Appl. Math.*, 30(5):509–541, 1977.

[11] David Kaziska and Anuj Srivastava. The karcher mean of a class of symmetric distributions on the circle. *Statistics & Probability Letters*, 78(11):1314 – 1316, 2008.

[12] D. G. Kendall, D. Barden, T. K. Carne, and H. Le. *Shape and shape theory*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1999.

[13] Huiling Le. On the consistency of procrustean mean shapes. *Adv. in Appl. Probab.*, 30(1):53–63, 1998.

[14] Huiling Le. Locating Fréchet means with application to shape spaces. *Adv. in Appl. Probab.*, 33(2):324–338, 2001.

[15] Huiling Le. Estimation of Riemannian barycentres. *LMS J. Comput. Math.*, 7:193–200 (electronic), 2004.

[16] Kanti V. Mardia and Peter E. Jupp. *Directional statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 2000. Revised reprint of ıt Statistics of directional data by Mardia [MR0336854 (49 #1627)].

[17] P. Massart. The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *Ann. Probab.*, 18(3):1269–1283, 1990.

[18] J. M. Oller and J. M. Corcuera. Intrinsic analysis of statistical estimation. *Ann. Statist.*, 23(5):1562–1581, 1995.

[19] Xavier Pennec. Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements. *J. Math. Imaging Vision*, 25(1):127–154, 2006.

[20] Herbert Ziezold. On expected figures and a strong law of large numbers for random elements in quasi-metric spaces. In *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and of the Eighth European Meeting of Statisticians (Tech. Univ. Prague, Prague, 1974), Vol. A*, pages 591–602. Reidel, Dordrecht, 1977.