

Article

# A Simple Formula for the Hilbert Metric with Respect to a Sub-Gaussian Cone

Stéphane Chrétien <sup>1,\*</sup> and Juan-Pablo Ortega <sup>2</sup><sup>1</sup> National Physical Laboratory, Hampton Road, Teddinton TW11 0LW, UK<sup>2</sup> Faculty of Mathematics and Statistics, University of St. Gallen, CH-9000 St. Gallen, Switzerland; juan-pablo.ortega@unisg.ch

\* Correspondence: stephane.chretien@npl.co.uk

Received: 11 January 2018; Accepted: 19 February 2018; Published: 2 March 2018

**Abstract:** The Hilbert metric is a widely used tool for analysing the convergence of Markov processes and the ergodic properties of deterministic dynamical systems. A useful representation formula for the Hilbert metric was given by Liverani. The goal of the present paper is to extend this formula to the non-compact and multidimensional setting with a different cone, tailored for sub-Gaussian tails.

**Keywords:** dynamical systems; Hilbert metric; Liverani's formula

## 1. Introduction

Let  $\mathbb{V}$  be a topological vector space and  $\mathcal{C}$  a closed convex cone inside  $\mathbb{V}$  enjoying the property that  $\mathcal{C} \cap -\mathcal{C} = \emptyset$ . Then,  $\mathcal{C}$  defines a partial ordering “ $\preceq$ ” by setting

$$f \preceq g \iff g - f \in \mathcal{C} \cup \{0\}. \quad (1)$$

Using this ordering, the Hilbert semi-metric  $\Theta$  on  $\mathcal{C}$  is given by

$$\Theta(f, g) = \log \left[ \frac{\beta(f, g)}{\alpha(f, g)} \right], \quad (2)$$

where

$$\alpha(f, g) = \sup\{\lambda \in \mathbb{R}^+ \mid \lambda f \preceq g\}, \quad (3)$$

and

$$\beta(f, g) = \inf\{\mu \in \mathbb{R}^+ \mid g \preceq \mu f\} \quad (4)$$

and  $\alpha = 0$  and  $\beta = +\infty$  if the corresponding sets are empty. In the sequel, we will focus on the case where  $\mathbb{V}$  is the space of continuous integrable functions on  $\mathbb{R}^n$  with Euclidean norm denoted by  $\|\cdot\|$ .

The Hilbert metric, originally introduced in [1], has been very useful in ergodic theory for deterministic dynamical systems [2–4], for the analysis and application of Markov Chains to control systems statistics and information theory [5–7]. Robust routing problems have also been studied using the Hilbert metric view point [8].

The first application of the Hilbert metric to the field of dynamical systems is Birkhoff's approach to the Perron-Frobenius theorem [9]; see also [10]. The ground-breaking result of Birkhoff is the following.

**Theorem 1.** Let  $\mathcal{C}$  be a cone in a vector space  $\mathbb{V}$  and  $\mathcal{K}$  be a cone in a vector space  $\mathbb{W}$ . If  $L : \mathbb{V} \mapsto \mathbb{W}$  is a linear mapping with  $L(\mathcal{C}) \subset \mathcal{K}$  and projective diameter  $\Delta(L) = \sup_{x, y \in \mathcal{C}, L(x) \sim_{\mathcal{K}} L(y)} d(Lx, Ly)$ . Then

$$\inf_{x \sim_{\mathcal{C}} y \in \mathcal{C}} d(Lx, Ly) / d(x, y) \leq \tanh \left( \frac{1}{4} \Delta(L) \right) \quad (5)$$

where  $\tanh(\infty) = 1$ .

Birkhoff’s theorem provides an elegant way to prove that certain maps between cones are contracts and therefore obtain existence and uniqueness for certain problems such as in Perron-Frobenius theory for positive operators.

The goal of this short note is to extend to the noncompact and multidimensional setting a useful formula for the Hilbert semi-metric which was previously given by Liverani [2] in the case in which  $\mathbb{V}$  is the space  $L^1([0, 1])$  of integrable functions of the interval  $[0, 1]$ . The cone  $\mathcal{C}$  in that result, which will be denoted in the sequel by  $\mathcal{C}_a, a > 0$ , is given by

$$\mathcal{C}_a([0, 1]) = \left\{ g \in C^0([0, 1]) \cap L_1([0, 1]) \mid \forall x, y \in [0, 1] \quad g > 0 \text{ and } \frac{g(x)}{g(y)} \leq e^{a|x-y|} \right\}. \tag{6}$$

In that setup, the Hilbert semi-metric is given by the following expression:

$$\Theta(f, g) = \ln \left[ \sup_{x, y, u, v \in [0, 1]} \frac{(e^{a|x-y|^2} g(y) - g(x))(e^{a|u-v|^2} f(v) - f(u))}{(e^{a|x-y|^2} f(y) - f(x))(e^{a|u-v|^2} g(v) - g(u))} \right]. \tag{7}$$

### 2. Main Result

The cone we use in the sequel is different from the cone chosen by Liverani in [1].

$$\mathcal{C}_a(\mathbb{R}^n) = \left\{ g \in C^0(\mathbb{R}^n) \cap L_1(\mathbb{R}^n) \mid \forall x, y \in \mathbb{R}^n \quad g > 0 \text{ and } \frac{g(x)}{g(y)} \leq e^{a\|x-y\|^2} \right\} \tag{8}$$

This cone is chosen to contain densities with subgaussian tails on  $\mathbb{R}^n$ . Our main result is the following theorem.

**Theorem 2.** *Let  $\Theta$  be the Hilbert semi-metric associated to  $\mathcal{C}_a, a > 0$ ; when  $f, g \in \mathcal{C}_a$ ,*

$$\Theta(f, g) = \ln \left[ \sup_{\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{v} \in E} \frac{(e^{a\|\mathbf{x}-\mathbf{y}\|^2} g(\mathbf{y}) - g(\mathbf{x}))(e^{a\|\mathbf{u}-\mathbf{v}\|^2} f(\mathbf{v}) - f(\mathbf{u}))}{(e^{a\|\mathbf{x}-\mathbf{y}\|^2} f(\mathbf{y}) - f(\mathbf{x}))(e^{a\|\mathbf{u}-\mathbf{v}\|^2} g(\mathbf{v}) - g(\mathbf{u}))} \right]. \tag{9}$$

**Proof.** We have to compute  $\alpha(f, g)$  and  $\beta(f, g)$ . Just as in the proof p. 247 of Lemma 2.2 of [2], we obtain

$$\alpha(f, g) = \min \left\{ \inf_{\mathbf{x} \in \mathbb{R}^n} \frac{g(\mathbf{x})}{f(\mathbf{x})} ; \inf_{\mathbf{x}, \mathbf{y} \in \mathbb{R}^n} \frac{e^{a\|\mathbf{x}-\mathbf{y}\|^2} g(\mathbf{y}) - g(\mathbf{x})}{e^{a\|\mathbf{x}-\mathbf{y}\|^2} f(\mathbf{y}) - f(\mathbf{x})} \right\}. \tag{10}$$

We now prove that

$$\inf_{\mathbf{x} \in \mathbb{R}^n} \frac{g(\mathbf{x})}{f(\mathbf{x})} \geq \inf_{\mathbf{x}, \mathbf{y} \in \mathbb{R}^n} \frac{e^{a\|\mathbf{x}-\mathbf{y}\|^2} g(\mathbf{y}) - g(\mathbf{x})}{e^{a\|\mathbf{x}-\mathbf{y}\|^2} f(\mathbf{y}) - f(\mathbf{x})} \tag{11}$$

Let  $(\mathbf{x}_n)_{n \in \mathbb{N}} \subset \mathbb{R}^n$  be a minimizing sequence for the left hand side of Equation (11). We have to split the analysis into two cases.

*First case.* Assume that  $(\mathbf{x}_n)_{n \in \mathbb{N}}$  has a bounded subsequence, denoted by  $(\mathbf{x}_{\sigma(n)})_{n \in \mathbb{N}}$ . Moreover, by the Bolzano-Weierstrass Theorem we may assume without loss of generality that this subsequence converges to some limit point  $\mathbf{x}^*$ . Fix  $\epsilon > 0$  and let  $N \in \mathbb{N}$  be such that

$$\frac{g(\mathbf{x}_{\sigma(n)})}{f(\mathbf{x}_{\sigma(n)})} \leq \inf_{\mathbf{x} \in \mathbb{R}^n} \frac{g(\mathbf{x})}{f(\mathbf{x})} + \epsilon \tag{12}$$

for all  $n \geq N$ . Now take  $\mathbf{x}$  such that

$$\left| \sup_{n \in \mathbb{N}} e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)}) - f(\mathbf{x}) \right| < +\infty. \tag{13}$$

Then, we get

$$\begin{aligned} \frac{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} g(\mathbf{x}_{\sigma(n)}) - g(\mathbf{x})}{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)}) - f(\mathbf{x})} &= \frac{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} \frac{g(\mathbf{x}_{\sigma(n)})}{f(\mathbf{x}_{\sigma(n)})} f(\mathbf{x}_{\sigma(n)}) - \frac{g(\mathbf{x})}{f(\mathbf{x})} f(\mathbf{x})}{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)}) - f(\mathbf{x})} \\ &\leq \frac{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} \left( \frac{g(\mathbf{x})}{f(\mathbf{x})} f(\mathbf{x}_{\sigma(n)}) + \epsilon \right) - \frac{g(\mathbf{x})}{f(\mathbf{x})} f(\mathbf{x})}{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)}) - f(\mathbf{x})} \\ &= \frac{g(\mathbf{x})}{f(\mathbf{x})} + \epsilon \frac{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)})}{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)}) - f(\mathbf{x})}. \end{aligned} \tag{14}$$

Let  $n$  tend towards  $+\infty$ , and obtain

$$\frac{e^{a\|\mathbf{x}-\mathbf{x}^*\|^2} g(\mathbf{x}^*) - g(\mathbf{x})}{e^{a\|\mathbf{x}-\mathbf{x}^*\|^2} f(\mathbf{x}^*) - f(\mathbf{x})} \leq \frac{g(\mathbf{x})}{f(\mathbf{x})} + \epsilon \frac{e^{a\|\mathbf{x}-\mathbf{x}^*\|^2} f(\mathbf{x}^*)}{e^{a\|\mathbf{x}-\mathbf{x}^*\|^2} f(\mathbf{x}^*) - f(\mathbf{x})} \tag{15}$$

and the result follows by taking  $\epsilon \rightarrow 0$  due to the assumption of Equation (13) on  $\mathbf{x}$ . The result is then easily seen to hold for all  $\mathbf{x}$  by continuity.

*Second case.* In this case,  $\|\mathbf{x}_n - \mathbf{x}\| \rightarrow +\infty$ . Consider the unit vector  $d_n$  given by

$$d_n = \frac{\mathbf{x}_n - \mathbf{x}}{\|\mathbf{x}_n - \mathbf{x}\|} \tag{16}$$

and extract a convergent subsequence  $(d_{\sigma(n)})_{n \in \mathbb{N}}$  denoting its limit by  $d^*$ . Now take  $\mathbf{z}_\delta = \mathbf{x} - \delta d^*$ . Let us prove that

$$\lim_{n \rightarrow +\infty} \frac{e^{a\|\mathbf{z}_\delta - \mathbf{x}_{\sigma(n)}\|^2} g(\mathbf{x}_{\sigma(n)}) - g(\mathbf{z}_\delta)}{e^{a\|\mathbf{z}_\delta - \mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)}) - f(\mathbf{z}_\delta)} \leq \frac{g(\mathbf{z}_\delta)}{f(\mathbf{z}_\delta)} \tag{17}$$

Letting  $\delta$  tend towards zero will prove the desired result using continuity. As in the first case, one easily finds that for each  $\epsilon > 0$  there exists  $N \in \mathbb{N}$  such that for all  $n \geq N$ ,

$$\begin{aligned} \frac{e^{a\|\mathbf{z}_\delta - \mathbf{x}_{\sigma(n)}\|^2} g(\mathbf{x}_{\sigma(n)}) - g(\mathbf{z}_\delta)}{e^{a\|\mathbf{z}_\delta - \mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)}) - f(\mathbf{z}_\delta)} &\leq \frac{g(\mathbf{z}_\delta)}{f(\mathbf{z}_\delta)} + \epsilon \frac{1}{1 - \frac{f(\mathbf{z}_\delta)}{e^{a\|\mathbf{z}_\delta - \mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)})}} \\ &= \frac{g(\mathbf{z}_\delta)}{f(\mathbf{z}_\delta)} + \epsilon \frac{1}{1 - \frac{f(\mathbf{x})}{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)})} \frac{f(\mathbf{z}_\delta) e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2}}{e^{a\|\mathbf{z}_\delta - \mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)})}}. \end{aligned} \tag{18}$$

Looking at this expression, we obviously wonder about the asymptotic behavior of

$$c_n = \frac{f(\mathbf{x})}{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)})} \frac{f(\mathbf{z}_\delta) e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2}}{e^{a\|\mathbf{z}_\delta - \mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)})}. \tag{19}$$

For this purpose, first notice that since  $f$  belongs to  $\mathcal{C}_a$ , we have

$$\frac{f(\mathbf{x})}{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2} f(\mathbf{x}_{\sigma(n)})} \leq 1. \tag{20}$$

and

$$\frac{f(\mathbf{z}_\delta)e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2}}{e^{a\|\mathbf{z}_\delta-\mathbf{x}_{\sigma(n)}\|^2}f(\mathbf{x}_{\sigma(n)})} \leq \frac{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2}e^{a\|\mathbf{x}-\mathbf{z}_\delta\|^2}}{e^{a\|\mathbf{z}_\delta-\mathbf{x}_{\sigma(n)}\|^2}} \tag{21}$$

but

$$e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2-a\|\mathbf{z}_\delta-\mathbf{x}_{\sigma(n)}\|^2+a\|\mathbf{x}-\mathbf{z}_\delta\|^2} = e^{-2a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|\|\mathbf{x}-\mathbf{z}_\delta\|}, \tag{22}$$

and since  $\|\mathbf{x}_n - x\| \rightarrow +\infty$ , we obtain that  $c_n \rightarrow 0$  as  $n \rightarrow +\infty$ , implying that the value one is not an accumulation point. Using this, we may deduce from Equation (18) that

$$\frac{e^{a\|\mathbf{z}_\delta-\mathbf{x}_{\sigma(n)}\|^2}g(\mathbf{x}_{\sigma(n)})-g(\mathbf{z}_\delta)}{e^{a\|\mathbf{z}_\delta-\mathbf{x}_{\sigma(n)}\|^2}f(\mathbf{x}_{\sigma(n)})-f(\mathbf{z}_\delta)} \leq \frac{g(\mathbf{z}_\delta)}{f(\mathbf{z}_\delta)} + \epsilon. \tag{23}$$

Letting  $\epsilon$  tend towards zero we obtain

$$\inf_{n \in \mathbb{N}} \frac{e^{a\|\mathbf{z}_\delta-\mathbf{x}_{\sigma(n)}\|^2}g(\mathbf{x}_{\sigma(n)})-g(\mathbf{z}_\delta)}{e^{a\|\mathbf{z}_\delta-\mathbf{x}_{\sigma(n)}\|^2}f(\mathbf{x}_{\sigma(n)})-f(\mathbf{z}_\delta)} \leq \frac{g(\mathbf{z}_\delta)}{f(\mathbf{z}_\delta)}. \tag{24}$$

Letting  $\delta$  tend towards zero and using continuity, we finally conclude that

$$\inf_{n \in \mathbb{N}} \frac{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2}g(\mathbf{x}_{\sigma(n)})-g(\mathbf{x})}{e^{a\|\mathbf{x}-\mathbf{x}_{\sigma(n)}\|^2}f(\mathbf{x}_{\sigma(n)})-f(\mathbf{x})} \leq \frac{g(\mathbf{x})}{f(\mathbf{x})}, \tag{25}$$

implying the desired result. □

### 3. Conclusions

In this short note, we presented an extension of a formula for the Hilbert metric obtained in [2] to the multidimensional and non-compact case for a different reference cone. Using the presented formula and Theorem 1, applications to the study of Markov Chain Monte Carlo methods and the analysis of the Sinkhorn algorithms for Optimal Transportation will be undertaken in future work.

**Author Contributions:** Stéphane Chrétien and Juan-Pablo Ortega worked equally on the content and on the writing of the paper.

**Conflicts of Interest:** The authors declare no conflicts of interest.

### References

1. Hilbert, D. Über die gerade Linie als kürzeste Verbindung zweier Punkte. *Math. Ann.* **1895**, *46*, 91–96.
2. Liverani, C. Decay of correlations. *Ann. Math.* **1995**, *142*, 239–301.
3. Baladi, V. *Positive Transfer Operators and Decay of Correlations*; Advanced Series in Nonlinear Dynamics, 16; World Scientific Publishing Co., Inc.: River Edge, NJ, USA, 2000.
4. Naud, F. Birkhoff cones, symbolic dynamics and spectrum of transfer operators. *Discret. Contin. Dyn. Syst.* **2004**, *11*, 581–598.
5. Bodnart, R.; Stettner, L. Asymptotics of controlled finite memory filters. *Syst. Control Lett.* **2002**, *47*, 181–190.
6. Di Masi, G.B.; Stettner, L. Ergodicity of hidden Markov models. *Math. Control Signals Syst.* **2005**, *17*, 269–296.
7. Le Gland, F.; Oudjane, N. Stability and uniform approximation of nonlinear filters using the Hilbert metric and application to particle filters. *Ann. Appl. Probab.* **2004**, *14*, 144–187.
8. Chen, Y.; Georgiou, T.; Pavon, M.; Tannenbaum, A. Relaxed Schroedinger bridges and robust network routing. *arXiv* **2018**, arXiv:1801.07852.

9. Birkhoff, G. Extensions of Jentzsch's theorem. *Trans. Am. Math. Soc.* **1957**, *85*, 219–227.
10. Kohlberg, E.; Pratt, J. The contraction mapping approach to the Perron-Frobenius theory: Why Hilbert's metric? *Math. Oper. Res.* **1982**, *7*, 198–210.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).