www.thalesgroup.com

Geometry of Structured Matrices based on Information, Koszul and Contact Geometries

F. Barbaresco

THALES

Outlines

- **General Framework of Matrix Geomety: Information/Contact & Koszul Geometries**
- **Probability in Metric Spaces by Fréchet**
- **Metric Geometry of Hermitian Positive Definites Matrices (HPD)**
  - View point of geometers: Cartan-Siegel Symmetric Bounded Domains
  - View point of probabilists: Information Geometry and Fisher Information metric
- **Karcher Flows for HPD matrices**
  - Basic Karcher flow to compute « Mean » Barycenter of N HPD matrices
  - Weiszfeld-Karcher flow to compute « Median » Barycenter of N HPD matrices
- **Karcher Flows for Toeplitz HPD matrices (THPD)**
  - Trench/Verblunsky Theorems: Partial Iwasawa Decomposition and CAR model
  - Karcher Flow for Hessian metric/Information Geometry metric of Complex Autoregressive Model
- **New « Ordered Statistic High Doppler Resolution CFAR » (OS-HDR-CFAR)**
  - OS-HDR-CFAR Processing Chain & results on real recorder data
- **Karcher Flows for Toeplitz-Block-Toeplitz HPD matrices (TBTHPD)**
  - Matrix Extension of Trench/Verblunsky Theorem: Multivariate CAR parametrization
  - Karcher Flow in Siegel Disk with Mostow Decomposition/Berger Fibration
- **New « Ordered Statistic Space-Time Adaptive Processing » (OS-STAP)**
- **Miscellaneous: Quantization of Complex Symmetric Spaces, Analysis on Symmetric Cone, Homogeneous Hyperbolic Affine Hyperspheres**

www.thalesgroup.com

Introduction

THALES

Information Theory
Nicolas .L. BRILLOUIN
Claude. SHANNON

Homogeneous Convex Cones G.
E.B. VINBERG
Jean-Louis KOSZUL

Geometric Science of Information

Probability/G. on structures
Y. OLLIVIER
M. GROMOV

Affine G.
Takeshi SASAKI
W. BLASCHKE
Eugenio CALABI

KOSZUL-VINBERG METRIC (KOSZUL-VINBERG CHARACTERISTIC FUNCTION)

Information G.
Calyampudi R. RAO
Nikolai N. CHENTSOV

FISHER METRIC

Probability in Metric Space
Maurice R. FRECHET

Probability on Riemannian Manifold
Michel EMERY
Marc ARNAUDON

Hessian G.
Hirohiko SHIMA
Jean-Louis KOSZUL

Sympl. G.
Von Thomas FRIEDRICH
Y. SHISHIDO

Contact G.
Vladimir ARNOLD

Homogeneous Symmetric Bounded Domains G.
Elie CARTAN
Carl Ludwig SIEGEL

THALES

Probability on Riemannian Manifold

$$\varphi(x) \leq E[\varphi(x)] \text{ , } \varphi \text{ continuous convex function}$$

$$b = E[g(x)] \Rightarrow \int_M \exp_b^{-1}(g(x))P(dx) = 0$$

Michel Emery

IRMA Lab, Strasbourg University

Probability on Riemannian Manifold
Stochastc Calculus on Manifolds

M. Émery, G. Mokobodzki, « **Sur le barycentre d'une probabilité dans une variété** », Séminaire de probabilités de Strasbourg, 25 , p. 220-233, 1991 http://archive.numdam.org/article/SPS_1991__25__220_0.pdf

M. Émery, P.A. Meyer, « Stochastic calculus in manifolds », Springer 1989

M. Émery, W. Zheng, « Fonctions convexes et semimartingales dans une variété », Séminaire de Probabilités XVIII, Lecture Notes in Mathematics 1059, Springer 1984

M. Fréchet, « L'intégrale abstraite d'une fonction abstraite et son application à la moyenne d'un élément aléatoire de nature quelconque », Revue Scientifique, 483-512, 1944

M. Fréchet, « **Les éléments aléatoires de nature quelconque dans un espace distancié** ». Annales IHP, 10 no. 4, p. 215-310, 1948 http://archive.numdam.org/article/AIHP_1948__10_4_215_0.pdf

Marc Arnaudon

PhD with M. Emery, Bordeaux University

P-Means Computation on Riemannian Manifold
Stochastic Flow on Riemannian Manifold

M. Arnaudon, L. Miclo, "**Means in complete manifolds: uniqueness and approximation**", http://arxiv.org/abs/1207.3232

M. Arnaudon, C. Dombry, A. Phan, Le Yang, « Stochastic algorithms for computing means of probability measures. », Stochastic Processes and their Applications122, pp. 1437-1455, 2012

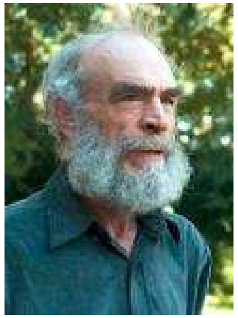
M. Arnaudon, A. Thalmaier, "Brownian motion and negative curvature", Boundaries and Spectra of Random Walks, Progress in Probability, Vol. 64, 145--163, Springer Basel ,2011

M. Arnaudon, F. Barbaresco, Le Yang, "**Medians and means in Riemannian geometry: existence, uniqueness and computation**" Matrix Information Geometry, Nielsen, Frank; Bhatia, Rajendra (Eds.), Springer, http://arxiv.org/pdf/1111.3120v1

Le Yang, « Médianes de mesures de probabilité dans les variétés riemanniennes et applications à la détection de cibles radar », PhD with advisors M. arnaudon & F. Barbaresco http://tel.archives-ouvertes.fr/docs/00/66/41/88/PDF/Dissertation-Le_YANG.pdf , THALES PhD Award 2012

THALES

Probability on structures: Gromov/Ollivier work on Fisher Metric



$$g_{ij} = -E\left[\frac{\partial^2 \log p(x,\theta)}{\partial\theta_i \partial\theta_j}\right] = E\left[\frac{\partial \log p(x,\theta)}{\partial\theta_i}\frac{\partial \log p(x,\theta)}{\partial\theta_j}\right]$$

$$ds^2 = K\left[p(x,\theta), p(x,\theta+d\theta)\right] = \sum_{i,j} g_{ij} d\theta_i d\theta_j$$

Misha Gromov, IHES

Abel Prize 2009

Mathematics about "interesting structures"

Category-theoretic approach of Fisher Metric

M. Gromov, « In a Search for a Structure, Part 1: On Entropy », preprint, July 2012
http://www.ihes.fr/~gromov/PDF/structre-serch-entropy-july5-2012.pdf

Chap 2. « **Fisher Metric** and Von Neumann Entropy»

Such a rescaling, being a non-trivial symmetry, is a significant structure in its own right; for example, the group of families of such "rescalings" leads the amazing orthogonal symmetry of the Fisher metric

Let us ponder over Boltzmann's function $e(p) = \sum_i p_i \log p_i$. All our inequalities for the entropy were reflections of the convexity of this $e(p)$, $p = \{p_i\}$, $i \in I$, on the unit simplex $\triangle(I)$, $\sum_i p_i = 1$, in the positive cone $\mathbb{R}^I_+ \subset \mathbb{R}^I$.

Convexity translates to the language of calculus as positive definiteness of Hessian $h = Hess(e)$ on $\triangle(I)$; following Fisher (1925) let us regard h as a Riemannian metric on $\triangle(I)$.

M. Gromov, « **Convex sets and Kähler manifolds**», in Advances in J. Differential Geom., F. Tricerri ed., World Sci., Singapore, pp. 1-38, 1990
http://www.ihes.fr/~gromov/PDF/%5B68%5D.pdf

Yann Ollivier, Paris-Sud University, LRI Dept.

CNRS Bronze Medal 2011

Introduction of probabilistic models on structured objects

Interplay between probability and differential geometry

Natural Gradient by Fisher Information Matrix (IGO)

Y. Ollivier, « Probabilités sur les espaces de configuration d'origine géométrique », PhD with advisers M. Gromov & P. Pansu
http://www.yann-ollivier.org/rech/publs/these.pdf

Y. Ollivier, « Le Hasard et la Courbure (Randomness and Curvature) », habilitation:
http://www.yann-ollivier.org/rech/publs/hdr_intro.pdf

Y. Ollivier & Youhei Akimoto, « Objective improvement in **information-geometric optimization** », FOGA 2013, preprint 2012
http://www.yann-ollivier.org/rech/publs/deeptrain.pdf

Yann Ollivier, Ludovic Arnold, Anne Auger, and Nikolaus Hansen, « **Information-geometric optimization**: A unifying picture via invariance principles », arXiv:1106.3708v1, 2011
http://www.yann-ollivier.org/rech/publs/quantile_igo.pdf

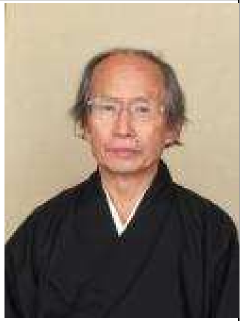
Video « seminaire Brillouin »: http://archiprod-externe.ircam.fr/video/VI02023900-226.mp4

Koszul forms, characteristic function & metric: Hessian structure

$x \in \Omega$, Ω and Ω^* dual cones, D flat connection

$$\psi(x) = \int_{\Omega^*} e^{-\langle x, x^* \rangle} dx^*, \quad g = Dd \log \psi > 0$$

Koszul-Vinberg Characteristic Function & metric of regular convex cone Ω



$$d \log(\psi \circ s) = d(\log \psi - \log \det s) = d \log \psi$$

Jean-Louis Koszul, French Sciences Academy
PhD student of Henri Cartan, Bourbaki member
Introduction of Koszul forms, Koszul-Vinberg characteristic function & metric

J.L. Koszul, « Sur la forme hermitienne canonique des espaces homogènes », complexes, Canad. J. Math. 7, pp. 562-576., 1955
J.L. Koszul, « Domaines bornées homogènes et orbites de groupes de transformations affines », Bull. Soc. Math. France 89, pp. 515-533., 1961
J.L. Koszul, « Ouverts convexes homogènes des espaces affines », Math. Z. 79, pp. 254-259., 1962
J.L. Koszul, « Variétés localement plates et convexité », Osaka J. Maht. 2, pp. 285-290., 1965
J.L. Koszul, « Déformations des variétés localement plates », .Ann Inst Fourier, 18 , 103-114., 1968
See: M. N. Boyom, « Convexité locale dans l'espace des connexions symétriques. Critère de comparaison des modèles statistiques »,, March 2012, IHP, Paris
http://www.ceremade.dauphine.fr/~peyre/mspc/mspc-thales-12/
« Les connexions symétriques est un sous-ensemble convexe contenant le sous-ensemble des connexions localement plates »

Hirohiko Shima, Emeritus Professor of Yamaguchi Univ.
PhD from Osaka University
Interplay between the Geometry of Hessian Structures and Information Geometry

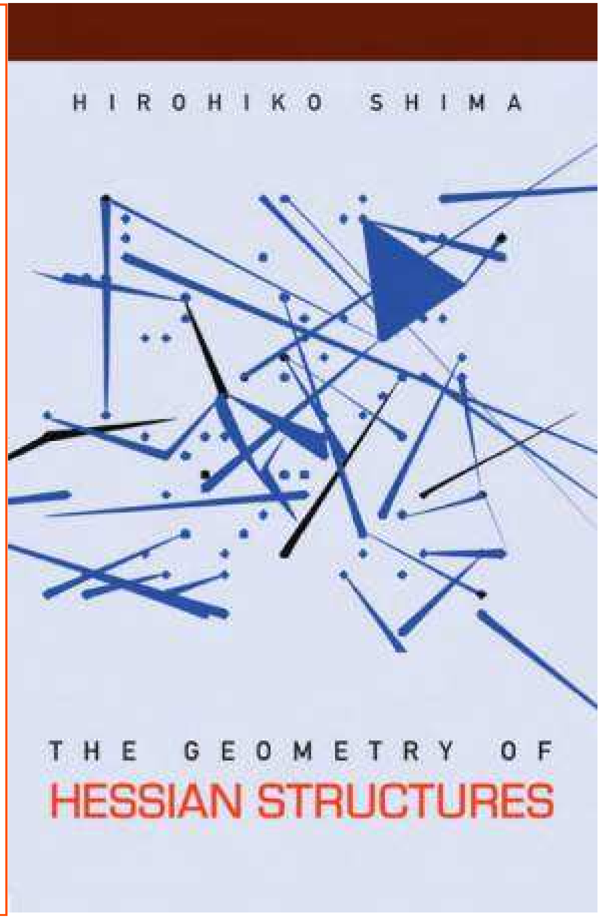
H. Shima, "The Geometry of Hessian Structures", World Scientific, 2007
http://www.worldscientific.com/worldscibooks/10.1142/6241
dedicated to Prof. Jean-Louis Koszul (« the content of the present book finds their origin in his studies »)
H. Shima,Symmetric spaces with invariant locally Hessian structures, J. Math. Soc. Japan,, pp. 581-589., 1977
H. Shima, « Homogeneous Hessian manifolds », Ann. Inst. Fourier, Grenoble, pp. 91-128., 1980
H. Shima, « Vanishing theorems for compact Hessian manifolds », Ann. Inst. Fourier, Grenoble, pp.183-205., 1986
H. Shima, « Harmonicity of gradient mappings of level surfaces in a real affine space », Geometriae Dedicata, pp. 177-184., 1995
H. Shima, « Hessian manifolds of constant Hessian sectional curvature », J. Math. Soc. Japan, pp. 735-753., 1995
H. Shima, « Homogeneous spaces with invariant projectively flat affine connections », Trans. Amer. Math. Soc., pp. 4713-4726, 1999

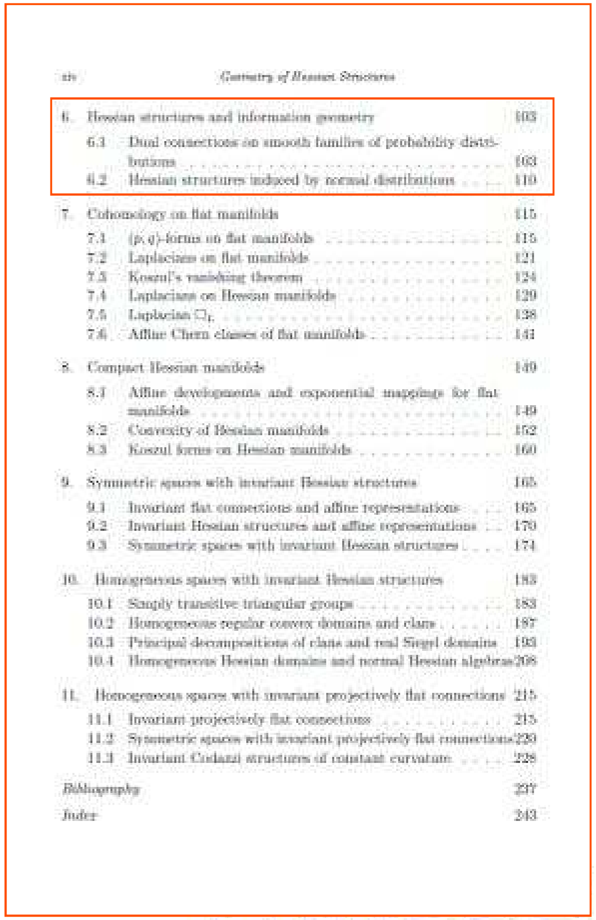
Hessian Geometry and J.L. Koszul Works

- Hirohiko Shima Book, « Geometry of Hessian Structures », world Scientific Publishing 2007, dedicated to Jean-Louis Koszul

Contents

xiii





xiv Geometry of Hessian Structures

The Geometry of Hessian Structures, Hirohiko Shima

H. Shima, "The Geometry of Hessian Structures", World Scientific, 2007
http://www.worldscientific.com/worldscibooks/10.1142/6241
dedicated to Prof. Jean-Louis Koszul (« the content of the present book finds their origin in his studies »)

Siegel domain	$\longleftrightarrow$	*Regular convex cone*
Holomorphic coordinate system $\{z^1, \cdots, z^n\}$	$\longleftrightarrow$	*Affine coordinate system* $\{x^1, \cdots, x^n\}$
Bergman kernel function $K(z, w)$	$\longleftrightarrow$	*Characteristic function* $\psi(x)$
Bergman metric $\sum_{i,j} \frac{\partial^2 \log K(z, \bar{z})}{\partial z^i \partial \bar{z}^j} dz^i d\bar{z}^j$	$\longleftrightarrow$	*Canonical metric* $\sum_{i,j} \frac{\partial^2 \log \psi}{\partial x^i \partial x^j} dx^i dx^j$

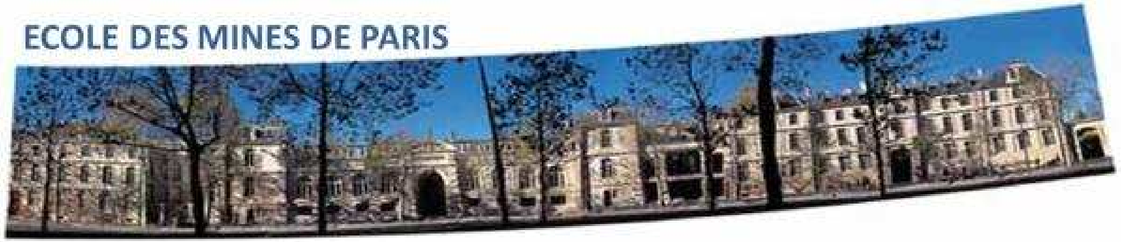
A pair *(D; g)* of a flat connection *D* and a Hessian metric *g* is called a Hessian structure.
J.L. Koszul studied a flat manifold endowed with a closed 1-form α such that $D\alpha$ is positive definite, whereupon $D\alpha$ is a Hessian metric. This is the ultimate origin of the notion of Hessian structures

A Hessian structure *(D; g)* is said to be of Koszul type, if there exists a closed 1-form α such that $g = D\alpha$

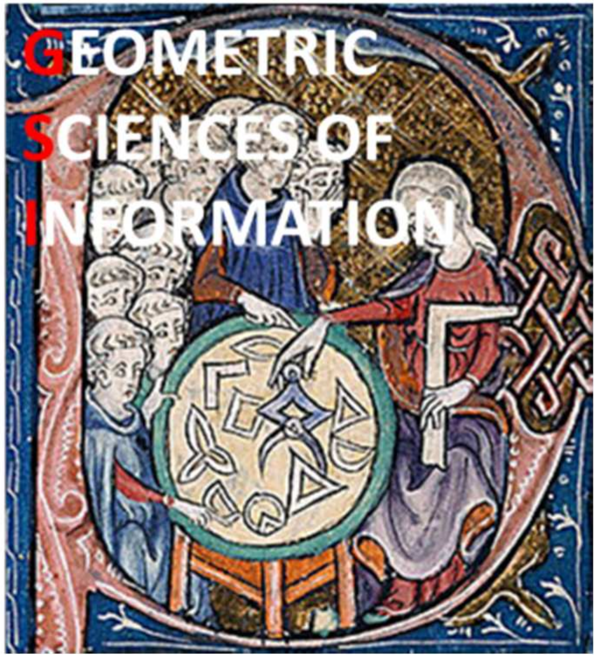
The second Koszul form β plays an important role similar to the Ricci tensor for a Kählerian metric

THALES

GSI'13 Geometric Sciences of Information, Ecole des Mines de Paris



Special sessions



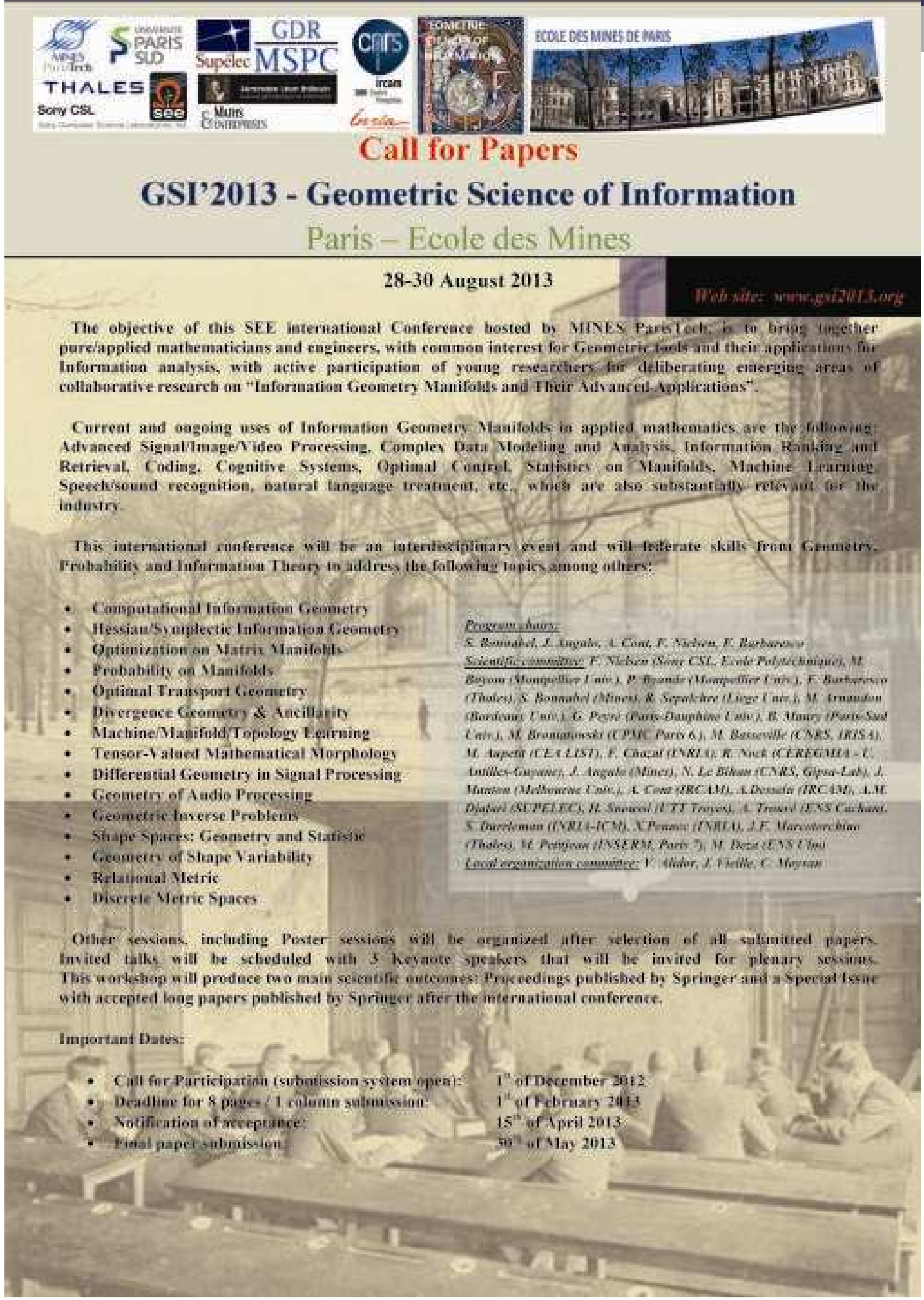
- Computational Information Geometry
- Hessian Information Geometry
- Symplectic Information Geometry
- Optimization on Matrix Manifolds
- Probability on Manifolds
- Divergence Geometry & Ancillarity
- Machine/Manifold/Topology Learning
- Optimal Transport Geometry
- Mathematical Morphology for Tensor/Matrix-Valued Images
- Geometry of Audio Processing
- Shape Spaces: Geometry and Statistic
- Geometry of Shape Variability
- Geometric Inverse Problems: Filtering/Estimation on Manifolds
- Differential Geometry in Signal Processing
- Relational Metric
- Finite Metric Spaces

28-30th August 2013 Paris

http://www.gsi2013.com

THALES

GSI'13 Geometric Sciences of Information, Ecole des Mines de Paris



Call for Papers

GSI'2013 - Geometric Science of Information

Paris – Ecole des Mines

28-30 August 2013

Web site: www.gsi2013.org

The objective of this SEE international Conference hosted by MINES ParisTech, is to bring together pure/applied mathematicians and engineers, with common interest for Geometric tools and their applications for Information analysis, with active participation of young researchers for deliberating emerging areas of collaborative research on "Information Geometry Manifolds and Their Advanced Applications".

Current and ongoing uses of Information Geometry Manifolds in applied mathematics are the following: Advanced Signal/Image/Video Processing, Complex Data Modeling and Analysis, Information Ranking and Retrieval, Coding, Cognitive Systems, Optimal Control, Statistics on Manifolds, Machine Learning, Speech/sound recognition, natural language treatment, etc., which are also substantially relevant for the industry.

This international conference will be an interdisciplinary event and will federate skills from Geometry, Probability and Information Theory to address the following topics among others:

- Computational Information Geometry
- Hessian/Symplectic Information Geometry
- Optimization on Matrix Manifolds
- Probability on Manifolds
- Optimal Transport Geometry
- Divergence Geometry & Ancillarity
- Machine/Manifold/Topology Learning
- Tensor-Valued Mathematical Morphology
- Differential Geometry in Signal Processing
- Geometry of Audio Processing
- Geometric Inverse Problems
- Shape Spaces: Geometry and Statistic
- Geometry of Shape Variability
- Relational Metric
- Discrete Metric Spaces

Program chairs: S. Bonnabel, J. Angulo, A. Cont, F. Nielsen, F. Barbaresco
Scientific committee: F. Nielsen (Sony CSL, Ecole Polytechnique), M. Boyom (Montpellier Univ.), P. Byande (Montpellier Univ.), F. Barbaresco (Thales), S. Bonnabel (Mines), R. Sepulchre (Liege Univ.), M. Arnaudon (Bordeaux Univ.), G. Peyré (Paris-Dauphine Univ.), B. Maury (Paris-Sud Univ.), M. Broniatowski (UPMC Paris 6), M. Basseville (CNRS, IRISA), M. Aupetit (CEA LIST), F. Chazal (INRIA), R. Nock (CEREGMIA - U. Antilles-Guyane), J. Angulo (Mines), N. Le Bihan (CNRS, Gipsa-Lab), J. Manton (Melbourne Univ.), A. Cont (IRCAM), A. Dessein (IRCAM), A.M. Djafari (SUPELEC), H. Snoussi (UTT Troyes), A. Trouvé (ENS Cachan), S. Durrleman (INRIA-ICM), X. Pennec (INRIA), J.F. Marcotorchino (Thales), M. Petitjean (INSERM, Paris 7), M. Deza (ENS Ulm)
Local organization committee: V. Alidor, J. Vieille, C. Moysan

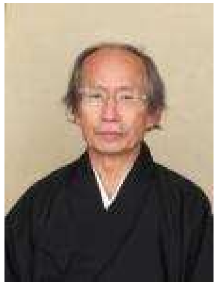
Other sessions, including Poster sessions will be organized after selection of all submitted papers. Invited talks will be scheduled with 3 Keynote speakers that will be invited for plenary sessions. This workshop will produce two main scientific outcomes: Proceedings published by Springer and a Special Issue with accepted long papers published by Springer after the international conference.

Important Dates:

- Call for Participation (submission system open): 1st of December 2012
- Deadline for 8 pages / 1 column submission: 1st of February 2013
- Notification of acceptance: 15th of April 2013
- Final paper submission: 30th of May 2013

- **Yann OLLIVIER** (Paris-Sud University, France): « Information-geometric optimization: The interest of information theory for discrete and continuous optimization »

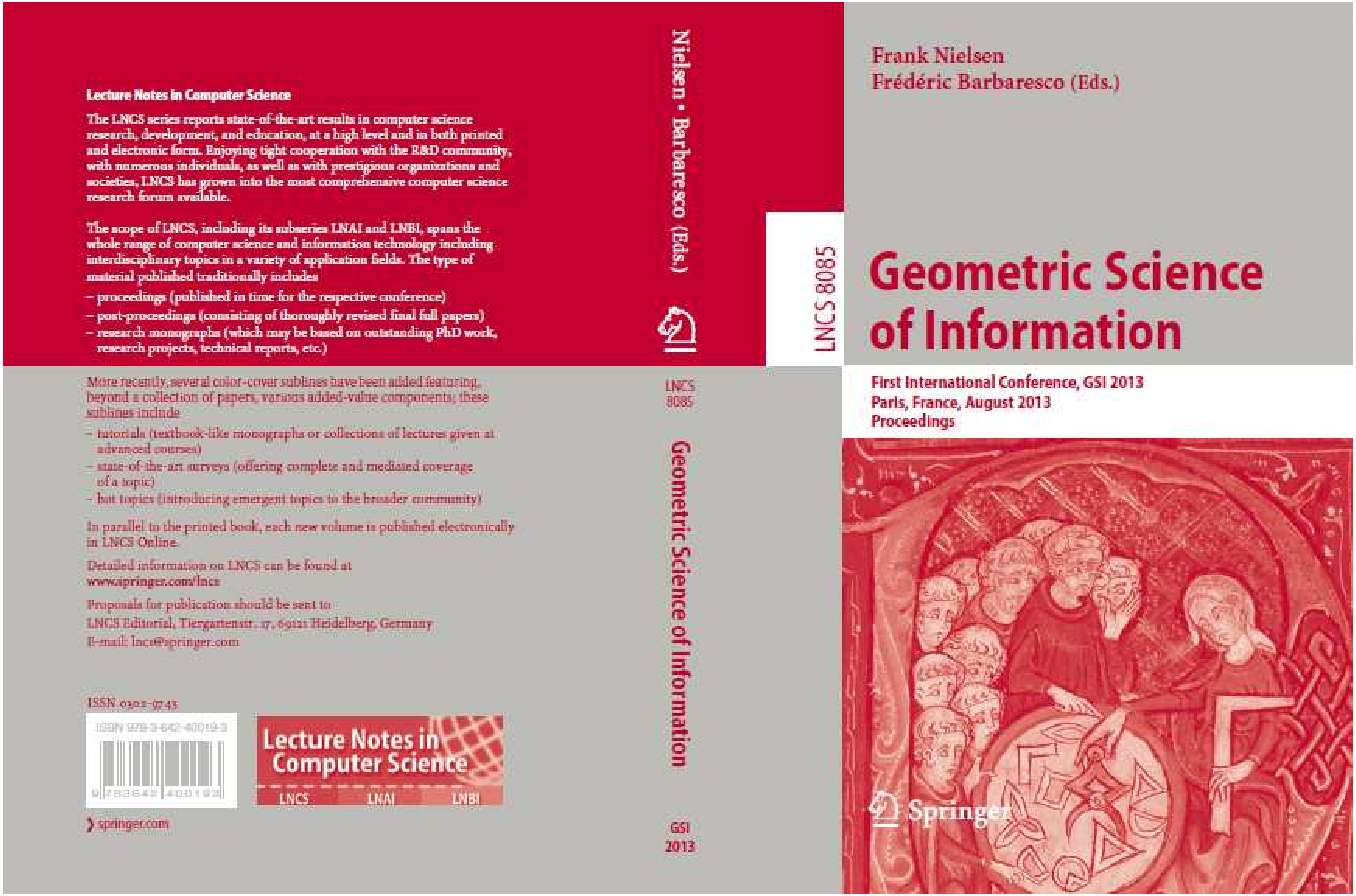
- **Hirohiko SHIMA** (Yamaguchi University, Japan): « Geometry of Hessian Structures »



- **Giovanni PISTONE** (Collegio Carlo Alberto, Italy): « Non-Parametric Information Geometry »

THALES

GSI'13 Geometric Sciences of Information, Ecole des Mines de Paris



THALES

Séminaire Léon Nicolas Brillouin

http://repmus.ircam.fr/brillouin/home

Séminaire Léon Brillouin

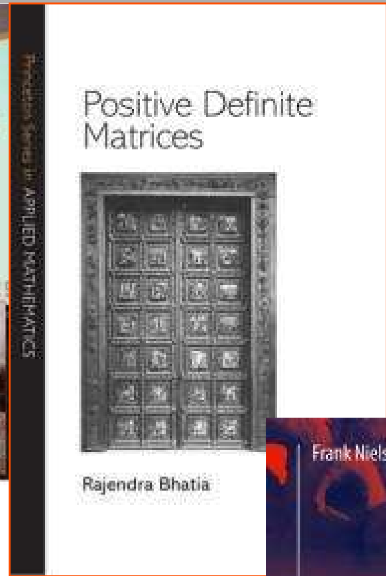
Sciences géométriques de l'information

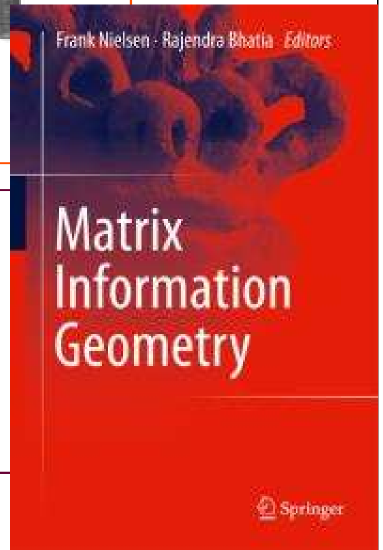
Corpus thématique : « Science géométrique de l'information »

- **Laboratoire d'accueil :**
  - IRCAM, Salle Igor Stravinsky
  - Projet INRIA/CNRS/Ircam **MuSync** (Arshia Cont)
- **Animation : Arshia Cont (IRCAM & INRIA), Frank Nielsen (Sony Research), F. Barbaresco (Thales)**
- **Les laboratoires**
  - IRCAM (Arshia Cont, Gerard Assayag, Arnaud Dessein)
  - Polytechnique (Olivier Schwanger,F.Nielsen)
  - Mines ParisTech (Silvere Bonnabel, Jesus Angulo)
  - UTT Troyes (Hichem Snoussi)
  - Univ. Poitiers (Marc Arnaudon, Le Yang)
  - Univ. Montpellier (Michel Boyom, Paul Bryand)
  - INRIA (Xavier Pennec, Arshia cont)
  - Thales (Frédéric Barbaresco, Alexis Decurninge)
  - Sony Resarch (Frank Nielsen)
- **Site web:** http://repmus.ircam.fr/brillouin/past-events

THALES

French-Indian CEFIPRA « Matrix Information Geometry» Thales Research & Technology and Ecole Polytechnique





With Prof. Rajendra Bhatia (New Dehli Indian Institut of Statistics)

MATRIX INEQUALITIES

Some general principles

Marvelous fact

In $(\mathbb{P}, \delta_2)$ the natural parametric equation for the geodesic joining A, B is

$$\gamma(t) = A^{1/2} (A^{-1/2} B A^{-1/2})^t A^{1/2} =: A \#_t B \quad 0 \le t \le 1$$

The geometric mean $A \#_{1/2} B$ is the midpoint of this!

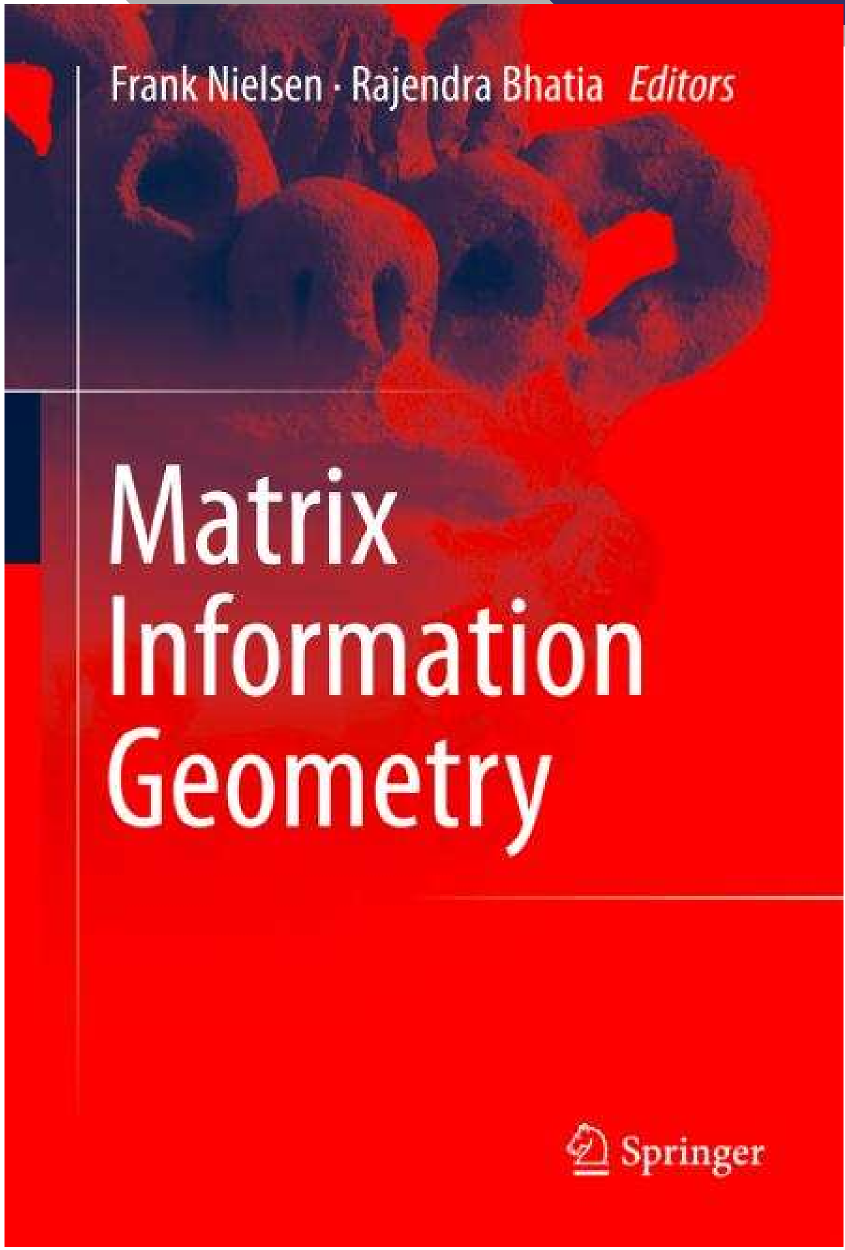
GEOMETRY & GEOMETRIC MEAN

$$\delta_2(A,B) = \delta_2(I, A^{-1/2} B A^{-1/2}) = \| \log A^{-1/2} B A^{-1/2} \|_2$$

(EMI equality)

$$= \left[\Sigma \log^2 \lambda_i (A^{-1/2} B A^{-1/2}) \right]^{1/2} = \left[\Sigma \log^2 \lambda_i (A^{-1} B) \right]^{1/2}$$

SPRINGER BOOK: « MATRIX INFORMATION GEOMETRY » (440 pages)

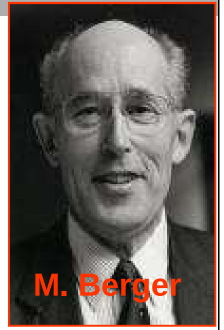


http://www.informationgeometry.org/MIG/MIGBOOKWEB/

THALES

« La Messe est dite » by Marcel Berger

Marcel Berger (IHES), « 150 ans de Géométrie Riemannienne », Géométrie au 20ième siècle, Histoire et horizons, Hermann Éditeur, 2005



- **Premier Miracle :**

  La théorie des espaces symétriques peut être considérée comme le premier miracle de la géométrie riemannienne, en fait comme un nœud de forte densité dans l'arbre de toutes les mathématiques … On doit à Elie Cartan dans les années 1926 d'avoir découvert que ces géométries sont , dans une dimension donnée, en nombre fini, et en outre toutes classées.

- **Second Miracle :**

  Entre les variétés localement symétriques et les variétés riemanniennes générales, il existe une catégorie intermédiaire, celle des variétés kählériennes. … On a alors affaire pour décrire le panorama des métriques kählériennes sur notre variété, non pas à un espace de formes différentielles quadratiques, très lourd, mais à un espace vectoriel de fonctions numériques [le potentiel de Kähler]. … La richesse Kählérienne fait dire à certains que la géométrie kählérienne est plus importante que la géométrie riemannienne.

- **Pas d'espoir d'autre miracle :**

  Ne cherchez pas d'autres miracles du genre des espaces (localement) symétriques et des variétés kählériennes. En effet, c'est un fait depuis 1953 que les seules variétés riemanniennes irréductibles qui admettent un invariant par transport parallèle autre que g elle-même (et sa forme volume) sont les espaces localement symétriques, les variétés kählériennes, les variétés kählérienne de Calabi-Yau, et les variétés hyperkählériennes.

Berger, Marcel, « Les espaces symétriques noncompacts ». Annales scientifiques de l'École Normale Supérieure, Sér. 3, 74 no. 2 (1957), p. 85-177

Read More : Marcel Berger, « A Panoramic View of Riemannian Geometry », Springer 2003

www.thalesgroup.com

General Framework of Matrix Geometry: Information/Contact Geometries & Koszul Geometry

THALES

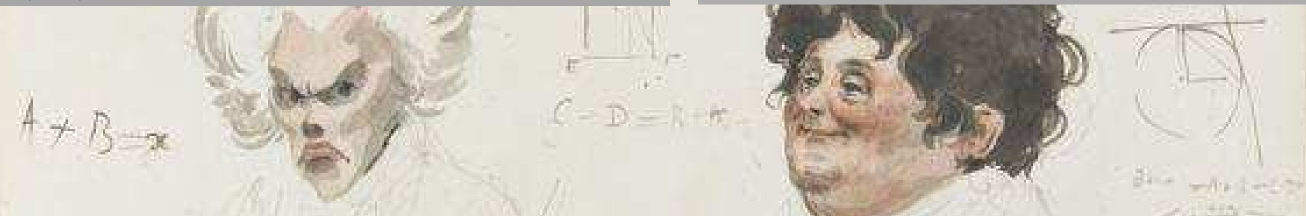
DUALITY Concept in Mathematic/Mechanic/Thermodynamic

INFORMATION GEOMETRY METRIC

$$g = -d^2 \log \Phi = d^2 \Psi \qquad g^* = d^2 \Psi^* = d^2 S$$

$$k\rho . \frac{\partial \phi(\rho)}{\partial (k\rho)} - \phi(\rho) = S \ \text{ with } \ \phi(\rho) = E / T - S$$

$$H(p,q,t) = \underset{\dot{q}}{Sup}\left(p.\dot{q} - L(q,\dot{q},t) \right)$$



(projective) LEGENDRE DUALITY

LEGENDRE TRANSFORM
(between Dual Space)

CONTACT/SYMPLECTIC GEOMETRIES
(Legendre mapping, fibration,…)

(Analytic) FOURIER DUALITY

FOURIER TRANSFORM
(Time-/Frequency Dual Spaces)

LINEAR ALGEBRA
(Linear Signal Processing)

$$L_\Phi(y) = \Psi^*(y) = \underset{y}{Arg\,Max} \left\{ \langle x, y \rangle - \Psi(x) \right\}$$

$$\Psi(x) = -\log \Phi(x) = -\log \int_{\Omega^*} e^{-\langle x,y \rangle} dy$$

Duality of Lagrangian/Hamiltonian in Physics

Massieu-Duhem Dual Potentials in Thermodynamics

Dual Potentials in Information Geometry

$$\Psi^*(x^*) = \langle x, x^* \rangle - \Psi(x)$$

$$\Psi^* = -\int_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi$$

Legendre Transform of minus logarithm of characteristic function (Fourier transform) = ENTROPY !!!

$$p_x(\xi) = e^{-\langle \xi, x \rangle} / \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi = e^{-\langle x, \xi \rangle + \Phi(x)}$$

$$x^* = \int_{\Omega^*} \xi . p_x(\xi) d\xi$$

Koszul-Vinberg Characteristic Function/Metric of convex cone

- J.L. Koszul [and E. Vinberg have introduced an affinely invariant Hessian metric on a sharp convex cone through its characteristic function.
- Ω is a sharp open convex cone in a vector space E of finite R dimension on (a convex cone is sharp if it does not contain any full straight line).
- Ω^* is the dual cone of Ω and is a sharp open convex cone.
- Let $d\xi$ the Lebesgue measure on E^* dual space of E, the following integral:

$$\psi_\Omega(x) = \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \quad \forall x \in \Omega$$

is called the **Koszul-Vinberg characteristic function**

Koszul-Vinberg Characteristic Function/Metric of convex cone

$$\psi_\Omega(x) = \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \quad \forall x \in \Omega$$

- ψ_Ω is analytic function defined on the interior of Ω and $\psi_\Omega(x) \to +\infty$ as $x \to \partial\Omega$
- If $g \in Aut(\Omega)$ then $\psi_\Omega(gx) = |\det g|^{-1} \psi_\Omega(x)$
- ψ_Ω is logarithmically strictly convex, the function $\phi_\Omega(x) = \log(\psi_\Omega(x))$ is strictly convex
- **Koszul 1-form α**: The differential 1-form $\alpha = d\phi_\Omega = d\log\psi_\Omega = d\psi_\Omega / \psi_\Omega$ is invariant by $G = Aut(\Omega)$. If $x \in \Omega$ and $u \in E$ then $\langle \alpha_x, u \rangle = -\int_{\Omega^*} \langle \xi, u \rangle . e^{-\langle \xi, x \rangle} d\xi$ and $\alpha_x \in -\Omega^*$
- **Koszul 2-form β** : The symmetric differential 2-form $\beta = D\alpha = d^2 \log \psi_\Omega$ is a positive definite symmetric bilinear form on E invariant under $G = Aut(\Omega)$ (from Schwarz inequality and $d^2 \log\psi_\Omega(u,v) = \int_{\Omega^*} \langle \xi, u \rangle \langle \xi, v \rangle e^{-\langle \xi, u \rangle} d\xi$)
- Koszul-Vinberg Metric: $D\alpha$ defines a Riemanian structure invariant by $Aut(\Omega)$ and then the Riemanian metric $g = d^2 \log \psi_\Omega$

THALES

Koszul-Vinberg Characteristic Function/Metric of convex cone

- **Koszul-Vinberg Metric :** $g = d^2 \log \psi_\Omega$

$$d^2 \log \psi(x) = d^2 \left[\log \int \psi_u du \right] = \frac{\int \psi_u d^2 \log \psi_u du}{\int \psi_u du} + \frac{1}{2} \frac{\iint \psi_u \psi_v (d \log \psi_u - d \log \psi_v)^2 dudv}{\iint \psi_u \psi_v dudv}$$

- **We can define a diffeomorphism by:** $x^* = -\alpha_x = -d \log \psi_\Omega(x)$

with $\langle df(x), u \rangle = D_u f(x) = \left. \frac{d}{dt} \right|_{t=0} f(x+tu)$

- **When the cone Ω is symmetric, the map $x^* = -\alpha_x$ is a bijection and an isometry with a unique fixed point (the manifold is a Riemannian Symmetric Space given by this isometry):**

$(x^*)^* = x$, $\langle x, x^* \rangle = n$ **and** $\psi_\Omega(x) \psi_{\Omega^*}(x^*) = cste$

- x^* **is characterized by** $x^* = \arg\min \left\{ \psi(y) / y \in \Omega^*, \langle x, y \rangle = n \right\}$

- x^* **is the center of gravity of the cross section** $\left\{ y \in \Omega^*, \langle x, y \rangle = n \right\}$ **of** Ω^* **:**

$$x^* = \int_{\Omega^*} \xi . e^{-\langle \xi, x \rangle} d\xi \Big/ \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$$

Koszul Entropy via Legendre Transform

- we can deduce "Koszul Entropy" defined as Legendre Transform of minus logarithm of Koszul-Vinberg characteristic function $\Phi(x) = -\log \psi_\Omega(x)$:

$\Phi^*(x^*) = \left\langle x, x^* \right\rangle - \Phi(x)$ with $x^* = D_x \Phi$ and $x = D_{x^*} \Phi^*$ where

- Demonstration: we set $\psi_\Omega(x) = \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \quad \forall x \in \Omega$

Using $x^* = \int_{\Omega^*} \xi . e^{-\langle \xi, x \rangle} d\xi / \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$

and $\left\langle -x^*, h \right\rangle = d_h \log \psi_\Omega(x) = -\int_{\Omega^*} \langle \xi, h \rangle e^{-\langle \xi, x \rangle} d\xi / \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$

we can write: $-\left\langle x^*, x \right\rangle = \int_{\Omega^*} \log e^{-\langle \xi, x \rangle} . e^{-\langle \xi, x \rangle} d\xi / \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$

and

$$\Phi^*(x^*) = -\int_{\Omega^*} \log e^{-\langle \xi, x \rangle} . e^{-\langle \xi, x \rangle} d\xi / \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi + \log \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$$

$$\Phi^*(x^*) = \left[\left(\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \right) . \log \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi - \int_{\Omega^*} \log e^{-\langle \xi, x \rangle} . e^{-\langle \xi, x \rangle} d\xi \right] / \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$$

Koszul Entropy via Legendre Transform

$$\Phi^*(x^*) = \left[\left(\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \right) . \log \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi - \int_{\Omega^*} \log e^{-\langle \xi, x \rangle} . e^{-\langle \xi, x \rangle} d\xi \right] / \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$$

- **We can then consider this Legendre transform as an entropy, that we could named "Koszul Entropy":**

$$\Phi^* = -\int_{\Omega^*} \frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} \log \frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} d\xi = -\int_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi$$

With $p_x(\xi) = e^{-\langle \xi, x \rangle} / \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi = e^{-\langle x, \xi \rangle - \log \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} = e^{-\langle x, \xi \rangle + \Phi(x)}$

and $x^* = D_x \Phi = \int_{\Omega^*} \xi . p_x(\xi) d\xi = \int_{\Omega^*} \xi . e^{-\langle x, \xi \rangle + \Phi(x)} d\xi = \int_{\Omega^*} \xi . e^{-\Phi^*(\xi)} d\xi$

- **If we denote by S_c the level surface of ψ_Ω : $S_c = \{\psi_\Omega(x) = c\}$ which is a non-compact submanifold in Ω , and by ω_c the induced metric of $d^2 \log \psi_\Omega$ on S_c , then assuming that the cone Ω is homogeneous under $G(\Omega)$, Sasaki proved that is a homogeneous hyperbolic affine hypersphere and every such hyperspheres can be obtained in this way .**
- **Sasaki also remarks that ω_c is identified with the affine metric and S_c is a global Riemannian symmetric space when Ω is a self-dual cone.**
- **Let Ω be a regular convex cone and let $g = d^2 \log \psi_\Omega$ be the canonical Hessian metric, then each level surface of the characteristic function ψ_Ω is a minimal surface of the Riemannian manifold (Ω, g) .**

THALES

Koszul-Vinberg Characteristic Function/Metric of convex cone

- Let v be the volume element of g. We define a closed 1-form α and β a symmetric bilinear form by:

$$D_X v = \alpha(X)v \quad \text{and} \quad \beta = D\alpha$$

- The forms α and β are called the first Koszul form and the second Koszul form for a Hessian structure *(D; g)* :

$$v = \left(\det\left[g_{ij}\right]\right)^{1/2} dx^1 \wedge ..dx^n \Rightarrow \begin{cases} \alpha_i = \dfrac{\partial}{\partial x^i} \log\left(\det\left[g_{ij}\right]\right)^{\frac{1}{2}} v \\ \beta_{ij} = \dfrac{\partial \alpha_i}{\partial x^j} = \dfrac{1}{2} \dfrac{\partial^2 \log \det\left[g_{kl}\right]}{\partial x^i \partial x^j} \end{cases}$$

- A pair *(D; g)* of a flat connection *D* and a Hessian metric *g* is called a Hessian structure.
- J.L. Koszul studied a flat manifold endowed with a closed 1-form α such that $D\alpha$ is positive definite, whereupon $D\alpha$ is a Hessian metric.
- This is the ultimate origin of the notion of Hessian structures. A Hessian structure *(D; g)* is said to be of Koszul type, if there exists a closed 1-form α such that $g = D\alpha$

Koszul-Vinberg Characteristic Function/Metric of convex cone

- We can apply this Koszul theory for Symmetric Positive Definite Matrices.
- Let the inner product $\langle x, y\rangle = Tr(xy), \forall x, y \in Sym_n(R)$
- Ω be the set of symmetric positive definite matrices is an open convex cone and is self-dual $\Omega^* = \Omega$:

$$\psi_\Omega(x) = \int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi \underset{\substack{\langle x,y\rangle = Tr(xy) \\ \Omega^*=\Omega \text{ self-dual}}}{=} \det x^{-\frac{n+1}{2}} \psi(I_n)$$

$$g = d^2 \log \psi_\Omega = -\frac{n+1}{2} d^2 \log \det x$$

$$x^* = -d \log \psi_\Omega = \frac{n+1}{2} d \log \det x = \frac{n+1}{2} x^{-1}$$

- Let be the regular convex cone consisting of all positive definite symmetric matrices of degree n. Then $(D, Dd \log \det x)$ is a Hessian structure on Ω, and each level surface of $\det x$ is a minimal surface of the Riemannian manifold $(\Omega, g = -Dd \log \det x)$

THALES

Koszul-Vey Theorem

- **J.L. Koszul and J. Vey have proved the following theorem:**
  - Koszul J.L., Variétés localement plates et convexité, Osaka J. Math. , n°2, p.285-290, 1965
  - Vey J., Sur les automorphismes affines des ouverts convexes saillants, Annali della Scuola Normale Superiore di Pisa, Classe di Science, 3e série, tome 24,n°4, p.641-665, 1970
- **Koszul-Vey Theorem:**

Let M be a connected Hessian manifold with Hessian metric g .

Suppose that admits a closed 1-form α such that $D\alpha = g$ and

there exists a group G of affine automorphisms of M preserving α :

- If M/G is quasi-compact, then the universal covering manifold of M is affinely isomorphic to a convex domain Ω real affine space not containing any full straight line.

- If M/G is compact, then Ω is a sharp convex cone.

[] Koszul J.L., Variétés localement plates et convexité, Osaka J. Math. , n°2, p.285-290, 1965
[] Vey J., Sur les automorphismes affines des ouverts convexes saillants, Annali della Scuola Normale Superiore di Pisa, Classe di Science, 3e série, tome 24,n°4, p.641-665, 1970

Characteristic function for probability measure

- Let μ be a positive Borel Measure on euclidean space *V*. Assume that the following integral is finite for all *x* in an open set $\Omega \subset V$:

$$x \in \Omega \qquad \psi_x(y) = \int e^{-\langle y,x \rangle} d\mu(x) \qquad p(y,dx) = \frac{1}{\psi_x(y)} e^{-\langle y,x \rangle} d\mu(x)$$

$$m(y) = \int x p(y,dx) = -\nabla \log \psi_x(y)$$

$$\langle V(y)u,v \rangle = \int \langle x - m(y),u \rangle \langle x - m(y),v \rangle p(y,dx) = D_u D_v \log \psi_x(y)$$

- P. Levy has made a systematic use of characteristic function, and he claimed that he was influenced by 1912, 2nd edition of Poincaré's book "Calcul des probabilités" where Poincaré designated for the first time the term "fonction caractéristique". Poincaré used characteristic function, for discuting exceptions to the Gaussian law. My opinion is that Poincaré used this name because of Massieu work used by him.

- For the Koszul-Vinberg characteristic function, if we replace the Lebesgue measure by Borel measure, then we recover the classical definition of characteristic function in Probability, and the previous KVCF could be compared by analogy with:

$$\psi_X(z) = \int_{-\infty}^{+\infty} e^{izx} dF(x) = \int_{-\infty}^{+\infty} e^{izx} p(x).dx = E\left[e^{izx}\right]$$

- **Large deviation principle**: Let $\{\Sigma_n\}$ be a sequence of random variables indexed by the positive integer n, and let $P(\Sigma_n \in d\varsigma) = P(\Sigma_n \in [\varsigma, \varsigma + d\varsigma])$ denote the probability measure associated with these random variables. We say that Σ_n or $P(\Sigma_n \in d\varsigma)$ satisfies a large deviation principle if the limit

$$I(\varsigma) = \underset{n\to\infty}{Lim} - \frac{1}{n} \log P(\Sigma_n \in d\varsigma)$$ exists.

The function $I(\varsigma)$ defined by this limit is called the rate function; the parameter *n* of decay is called in large deviation theory the speed. The existence of a large deviation principle for Σ_n means concretely that the dominant behavior of $P(\Sigma_n \in d\varsigma)$ is a decaying exponential with *n*, with rate $I(\varsigma)$:

$$P(\Sigma_n \in d\varsigma) \approx e^{-n.I(\varsigma)} d\varsigma$$ or $$p(\Sigma_n = \varsigma) \approx e^{-n.I(\varsigma)}$$

Laplace Principle of Large Deviation

- If we write the Legendre-Fenchel transform of a function $h(x)$ defined by $g(k) = \sup_x \{k.x - h(x)\}$, we can express the following principle and Laplace's method:
- **Laplace Principle** : Let $\{(\Omega_n, F_n, P_n), n \in N\}$ be a sequence of probability spaces, Θ a complete separable metric space, $\{Y_n, n \in N\}$ a sequence of random variables such that Y_n maps Ω_n into Θ , and I a rate function on Θ . Then, Y_n satisfies the Laplace principle on Θ with rate function I if for all bounded, continuous functions f mapping into R:

$$\lim_{n \to \infty} \frac{1}{n} \log E_{P_n}\left[e^{n.f(Y_n)}\right] = \sup_{x \in \Theta} \{f(x) - I(x)\}$$

THALES

- **Laplace Principle :**

$$\lim_{n\to\infty}\frac{1}{n}\log E_{P_n}\left[e^{n.f(Y_n)}\right]=\sup_{x\in\Theta}\{f(x)-I(x)\}$$

- **Proof:**

If Y_n satisfies the large deviation principle on Θ with rate function I , then $P_n(Y_n\in dx)\approx e^{-n.I(x)}dx$:

$$\lim_{n\to\infty}\frac{1}{n}\log E_{P_n}\left[e^{n.f(Y_n)}\right]=\lim_{n\to\infty}\frac{1}{n}\log\int_{\Omega_n}e^{nf(Y_n)}.dP_n=\lim_{n\to\infty}\frac{1}{n}\log\int_{\Theta}e^{n.f(x)}P_n(Y_n\in dx)$$

$$\approx\lim_{n\to\infty}\frac{1}{n}\log\int_{\Theta}e^{n.f(x)}e^{-n.I(x)}.dx=\lim_{n\to\infty}\frac{1}{n}\log\int_{\Theta}e^{n.[f(x)-I(x)]}.dx$$

- **The asymptotic behavior of the last integral is determined by the largest value of the integrand :**

$$\lim_{n\to\infty}\frac{1}{n}\log E_{P_n}\left[e^{n.f(Y_n)}\right]=\frac{1}{n}\log\left[\sup_{x\in\Theta}\left\{e^{n.[f(x)-I(x)]}\right\}\right]=\sup_{x\in\Theta}\{f(x)-I(x)\}$$

Laplace Principle of Large Deviation

$$\underset{n\to\infty}{Lim}\frac{1}{n}\log E_{P_n}\left[e^{n.f(Y_n)}\right]=\underset{x\in\Theta}{Sup}\{f(x)-I(x)\}$$

- The generating function of Σ_n is defined as:

$$\psi_n(x)=E\left[e^{nx\Sigma_n}\right]=\int e^{nx\varsigma}P(\Sigma_n\in d\varsigma)$$

In terms of the density $p(\Sigma_n)$, we have instead

$$\psi_n(x)=\int e^{nx\varsigma}p(\Sigma_n=\varsigma)d\varsigma$$

In both expressions, the integral is over the domain of Σ_n .

The function $\phi(x)$ defined by the limit:

$$\phi(x)=\underset{n\to\infty}{Lim}\frac{1}{n}\log\psi_n(x)$$

is called the scaled cumulant generating function or the log-generating function of Σ_n .

The existence of this limit is equivalent to writing $\psi_n(x)\approx e^{n.\phi(x)}$

THALES

Laplace Principle of Large Deviation

- **Gärtner-Ellis Theorem:**

  If $\phi(x)$ is differentiable, then Σ_n satisfies a large deviation principle with rate function $I(\varsigma)$ by the Legendre-Fenchel transform of $\phi(x)$:

$$I(\varsigma) = \underset{x}{Sup}\{x.\varsigma - \phi(x)\}$$

- **Varadhan's Theorem:**

  If Σ_n satisfies a large deviation principle with rate function $I(\varsigma)$, then its scaled cumulant generating function $\phi(x)$ is the Legendre-Fenchel transform of $I(\varsigma)$:

$$\phi(x) = \underset{\varsigma}{Sup}\{x.\varsigma - I(\varsigma)\}$$

- Properties of $\phi(x)$: statistical moments of Σ_n are given by derivatives of $\phi(x)$:

$$\left.\frac{d\phi(x)}{dx}\right|_{x=0} = \lim_{n\to\infty} E[\Sigma_n] \qquad \left.\frac{d^2\phi(x)}{dx^2}\right|_{x=0} = \lim_{n\to\infty} Var[\Sigma_n]$$

- In 1869, François Jacques Dominique MASSIEU, French Engineer from Corps des Mines, has presented two papers to French Science Academy on « Characteristic function » in Thermodynamic.
- Massieu has demonstrated that some mechanical and thermal properties of physical and chemical systems could be derived from two potentials called "characteristic functions".
- The infinitesimal amount of heat dQ received by a body produces external work of dilatation, internal work, and an increase of body sensible heat. The last two effects could not be identified separately and are noted dE (function E accounted for the sum of mechanical and thermal effects by equivalence between heat and work).
  The external work $P.dV$ is thermally equivalent to $A.P.dV$ (with A the conversion factor between mechanical and thermal measures). The first principle provides $dQ = dE + A.P.dV$. For a closed reversible cycle (Joule/Carnot principles) $\int dQ / T = 0$ that is the complete differential dS of a function S of $dS = dQ / T$.

[] Massieu F., Thermodynamique: Mémoire sur les fonctions caractéristiques des divers fluides et sur la théorie des vapeurs, 92-pgs, Académie des Sciences, 1876
[] Duhem P., Sur les équations générales de la Thermodynamique, Annales Scientifiques de l'Ecole Normale Supérieure, 3e série, tome VIII, p. 231, 1891
[] Arnold V.I., Contact geometry: the geometrical method of Gibbs's thermodynamics, Pro-ceedings of the Gibbs Symposium, p.163–179, Amer. Math. Soc., Providence, RI, 1990

Legendre Duality in Thermodynamic: Massieu Potentials

- **If we select volume *V* and temperature *T* as independent variables:**

$$T.dS = dQ \Rightarrow T.dS - dE = A.P.dV \Rightarrow d\left(TS\right) - dE = S.dT + A.P.dV$$

If we set $H = TS - E$ **, then we have**

$$dH = S.dT + A.P.dV = \frac{\partial H}{\partial T}.dT + \frac{\partial H}{\partial V}.dV$$

Massieu has called H **the "characteristic function" because all body characteristics could be deduced of this function:**

$$S = \frac{\partial H}{\partial T} \quad , \quad P = \frac{1}{A}\frac{\partial H}{\partial V} \quad \text{and} \quad E = TS - H = T\frac{\partial H}{\partial T} - H$$

- **If we select pression P and temperature T as independent variables:**

Massieu characteristic function is then given by $H' = H - AP.V$

$$dH' = dH - AP.dV - AV.dP = S.dT - AV.dP = \frac{\partial H'}{\partial T}.dT + \frac{\partial H'}{\partial P}.dP$$

and we can deduce: $S = \dfrac{\partial H'}{\partial T}$ **and** $V = -\dfrac{1}{A}\dfrac{\partial H'}{\partial P}$

and inner energy:

$$E = TS - H = TS - H' - AP.V \Rightarrow E = T\frac{\partial H'}{\partial T} - H' + P.\frac{\partial H'}{\partial P}$$

Legendre Duality in Thermodynamic: Massieu Potentials

- **Deriving all body properties dealing with thermodynamics from Massieu characteristic function and its derivatives :**
  - *« je montre, dans ce mémoire, que toutes les propriétés d'un corps peuvent se déduire d'une fonction unique, que j'appelle la fonction caractéristique de ce corps»*
- **In thermodynamics, the Massieu potential is the Legendre transform of the Entropy, and depends on :** $\rho = 1/kT$

$$\phi(\rho) = -k\rho.F = S - E/T$$

where F **is the Free Energy** $F = E - TS$ **and** E **inner energy**

$$-\frac{\partial\phi(\rho)}{\partial(k\rho)} = \frac{\partial(k\rho F)}{\partial(k\rho)} = F + k\rho\frac{\partial F}{\partial(k\rho)} = F + k\rho\left(\frac{\partial F}{\partial T}\right)_V\left(\frac{\partial T}{\partial(k\rho)}\right)$$

$$-\frac{\partial\phi(\rho)}{\partial(k\rho)} = F + k\rho(-S)\left(-\frac{1}{(k\rho)^2}\right) = F + \frac{S}{k\rho} = (E - TS) + TS = E$$

The Legendre transform of the Massieu potential gives Entropy S

$$L(\phi) = k\rho.\frac{\partial\phi(\rho)}{\partial(k\rho)} - \phi(\rho) = k\rho.(-E) - k\rho.F = k\rho(F - E) = -S$$

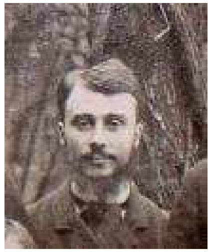
THALES

Legendre Duality in Thermodynamic: Duhem-Massieu Potentials

Pierre Duhem Thermodynamic Potentials

$$\Omega = G(E - TS) + W$$

- **Duhem P., « Sur les équations générales de la Thermodynamique », Annales Scientifiques de l'Ecole Normale Supérieure, 3e série, tome VIII, p. 231, 1891**



  - "Nous avons fait de la Dynamique un cas particulier de la Thermodynamique, une Science qui embrasse dans des principes communs tous les changements d'état des corps, aussi bien les changements de lieu que les changements de qualités physiques "

- **four scientists were credited by Duhem with having carried out "the most important researches on that subject":**
  - **F. Massieu** had managed to derive Thermodynamics from a "characteristic function and its partial derivatives"
  - **J.W. Gibbs** had shown that Massieu's functions "could play the role of potentials in the determination of the states of equilibrium" in a given system.
  - **H. von Helmholtz** had put forward "similar ideas"
  - **A. von Oettingen** had given "an exposition of Thermodynamics of remarkable generality" based on general duality concept in **"Die thermodynamischen Beziehungen antithetisch entwickelt", St. Petersburg 1885**

THALES

Duality of Information Geometry: Multivariate Gaussian Law

Dual Coordinates systems & Potential functions

- Potential Functions are Dual and related by **Legendre transformation** :

$$\text{Dual coordinates} \begin{cases} \tilde{\Theta} = (\theta, \Theta) = \left(\Sigma^{-1}m, (2\Sigma)^{-1}\right) \\ \tilde{H} = (\eta, H) = \left(m, -\Sigma + mm^T\right) \end{cases}$$

$$\Rightarrow \begin{cases} \tilde{\Psi}(\tilde{\Theta}) = 2^{-2}Tr\left(\Theta^{-1}\theta\theta^T\right) - 2^{-1}\log(\det\Theta) + 2^{-1}n\log(\pi) \\ \tilde{\Phi}(\tilde{H}) = -2^{-1}\log\left(1+\eta^T H^{-1}\eta\right) - 2^{-1}\log(\det(-H)) - 2^{-1}n\log(2\pi e) \end{cases}$$

$$\begin{cases} \dfrac{\partial\tilde{\Psi}}{\partial\theta} = \eta \\ \dfrac{\partial\tilde{\Psi}}{\partial\Theta} = H \end{cases} \text{ and } \begin{cases} \dfrac{\partial\tilde{\Phi}}{\partial\eta} = \theta \\ \dfrac{\partial\tilde{\Phi}}{\partial H} = \Theta \end{cases}$$

$$\tilde{\Phi} \equiv \left\langle \tilde{\Theta}, \tilde{H} \right\rangle - \tilde{\Psi}$$

$$\text{with } \left\langle \tilde{\Theta}, \tilde{H} \right\rangle = Tr\left(\theta\eta^T + \Theta H^T\right)$$

$$\tilde{\Phi}(\tilde{H}) = E[\log p]$$ ← **Entropy**

- Hessians are convexe and define Riemannian metrics :

$$g_{ij} = \frac{\partial^2\tilde{\Psi}}{\partial\Theta_i\partial\Theta_j} \text{ and } g_{ij}^* \equiv \frac{\partial^2\tilde{\Phi}}{\partial H_i \partial H_j}$$

THALES

Duality of Information Geometry: Multivariate Gaussian Law

Link with Kullback Divergence

- For each Dual Geometry, we can built a divergence that is directly related to Kulback-Leibler Divergence :

$$Div[N(m_2,\Sigma_2),N(m_1,\Sigma_1)] \equiv \tilde{\Psi}(\tilde{\Theta}_2)+\tilde{\Phi}(\tilde{H}_1)-\langle\tilde{\Theta}_2,\tilde{H}_1\rangle \geq 0$$

$$Div^*[N(m_2,\Sigma_2),N(m_1,\Sigma_1)] \equiv \tilde{\Psi}(\tilde{\Theta}_1)+\tilde{\Phi}(\tilde{H}_2)-\langle\tilde{\Theta}_1,\tilde{H}_2\rangle \geq 0$$

$$Div[N(m_2,\Sigma_2),N(m_1,\Sigma_1)] = \int p(x/m_1,\Sigma_1)\log\frac{p(x/m_1,\Sigma_1)}{p(x/m_2,\Sigma_2)}dx$$

$$Div(N(m_2,\Sigma_2),N(m_1,\Sigma_1)) = \frac{1}{2}\begin{bmatrix}-\log(\det(\Sigma_1\Sigma_2^{-1}))+Tr(\Sigma_1(\Sigma_2^{-1}-\Sigma_1^{-1})) \\ +Tr(\Sigma_2^{-1}(m_1-m_2)(m_1-m_2)^T)\end{bmatrix}$$

Riemannian Metric

- As Potential are convexe, their Hessians define Riemannian Metrics :

$$ds^2 = \frac{1}{2}g_{ij}d\Theta_i d\Theta_j + O(|d\Theta_i|^3) = \frac{1}{2}g^{ij}dH_i dH_j + O(|dH_i|^3)$$

THALES

Legendre Duality in Mechanics

- **In Mechanics, Legendre Duality gives the relation between:**
  - the variational Euler-Lagrange
  - the symplectic Hamilton-Jacobi formulations

  of the equations of motion

- **As described by Vladimir Arnold, in the general case, we can define the Hamiltonian *H* as the fiberwise Legendre transformation of the Lagrangian L:** $$H(p,q,t) = \underset{\dot{q}}{Sup}\left(p.\dot{q} - L(q,\dot{q},t)\right)$$

- **Due to strict convexity,** $H(p,q,t) = p.\dot{q} - L(q,\dot{q},t)$

  supremum is reached in a unique point $\dot{q}$ **such that :**

  $p = \partial_{\dot{q}} L(q,\dot{q},t)$ **and** $\dot{q} = \partial_p H(p,q,t)$

- **Young-Fenchel inequality. For all** $q,t,\dot{q},p$ **,the following holds :**

  $$p.\dot{q} \le L(q,\dot{q},t) + H(p,q,t)$$

  with equality if and only if $p = \partial_{\dot{q}} L(q,\dot{q},t)$

Legendre Duality in Mechanics

- **If we consider total differential of Hamiltonian:**

$$\begin{cases} dH = \dot{q}dp + pd\dot{q} - \partial_q L dq - \partial_{\dot{q}} L d\dot{q} - \partial_t L dt = \dot{q}dp - \partial_q L dq - \partial_t L dt \\ dH = \partial_p H dp + \partial_q H dq + \partial_t H dt \end{cases}$$

$$\Rightarrow \begin{cases} \dot{q} = \partial_p H \\ -\partial_q L = \partial_q H \end{cases}$$

- **Euler-Lagrange equation** $\partial_t \partial_{\dot{q}} L - \partial_q L = 0$

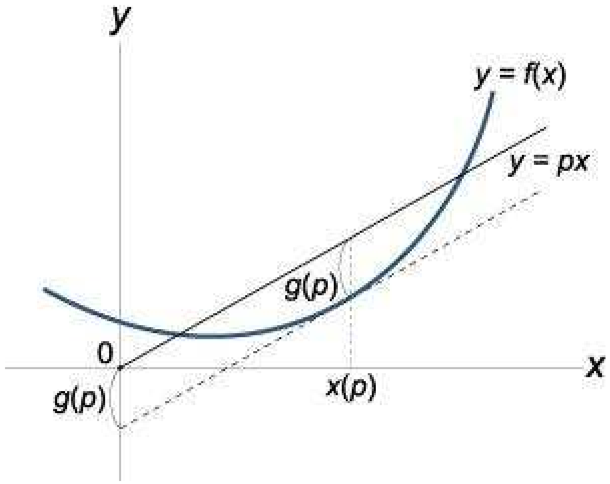
  with $p = \partial_{\dot{q}} L$ **and** $-\partial_q L = \partial_q H$

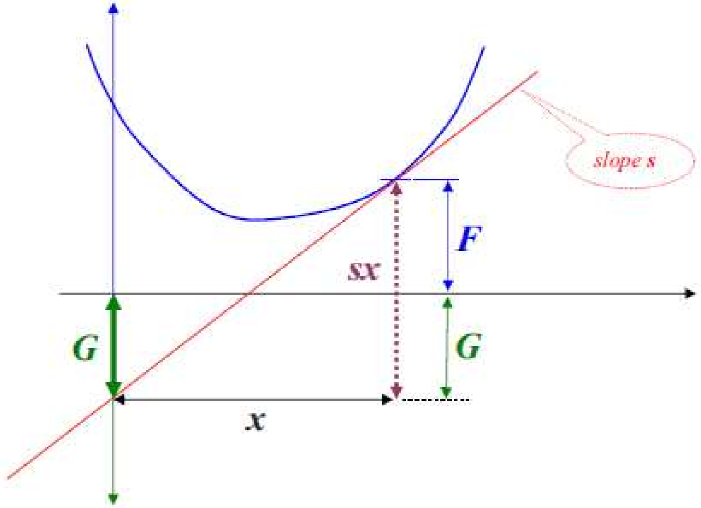
  provides the 2nd Hamilton equation $\dot{p} = -\partial_q H$

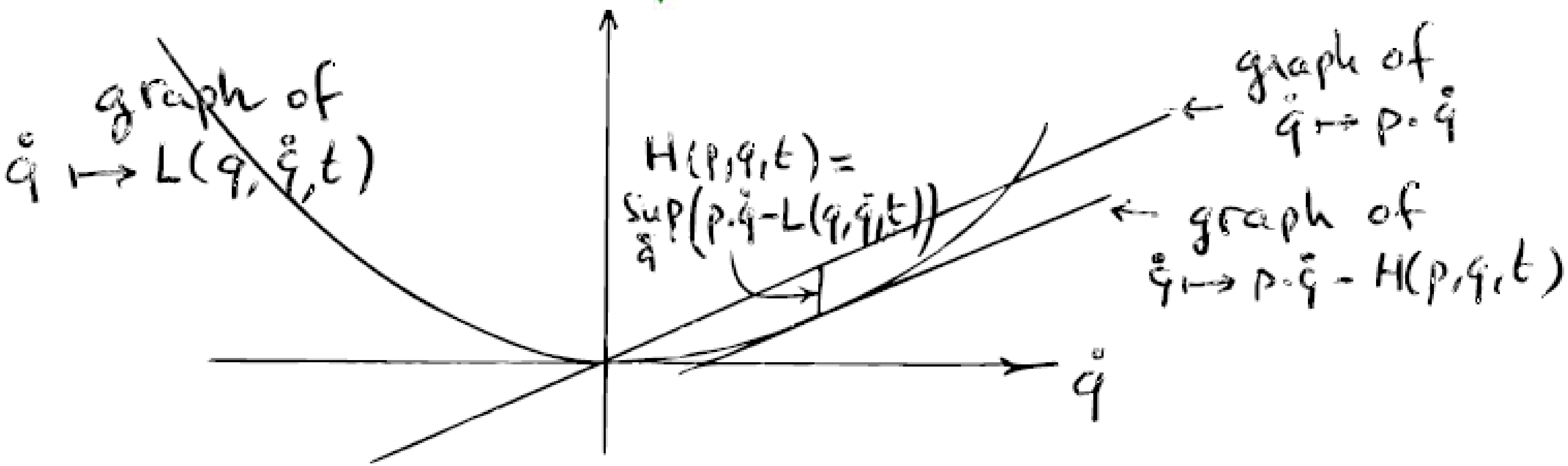
  with $\dot{q} = \partial_p H$

  in Darboux coordinates.

Legendre Duality in Mechanics







$$H(p,q,t) = \underset{\dot{q}}{Sup}\left(p.\dot{q} - L(q,\dot{q},t)\right)$$

$$p = \partial_{\dot{q}} L(q,\dot{q},t) \qquad \dot{q} = \partial_{p} H(p,q,t)$$

THALES

Pfaffian Form and Poincaré-Cartan Integral Invariant

- **Considering Pfaffian form** $\omega = p.dq - H.dt$

  related to Poincaré-Cartan integral invariant, based on:

  $p = \partial_{\dot{q}} L$ **and** $H = p.\dot{q} - L$

  we can deduce:

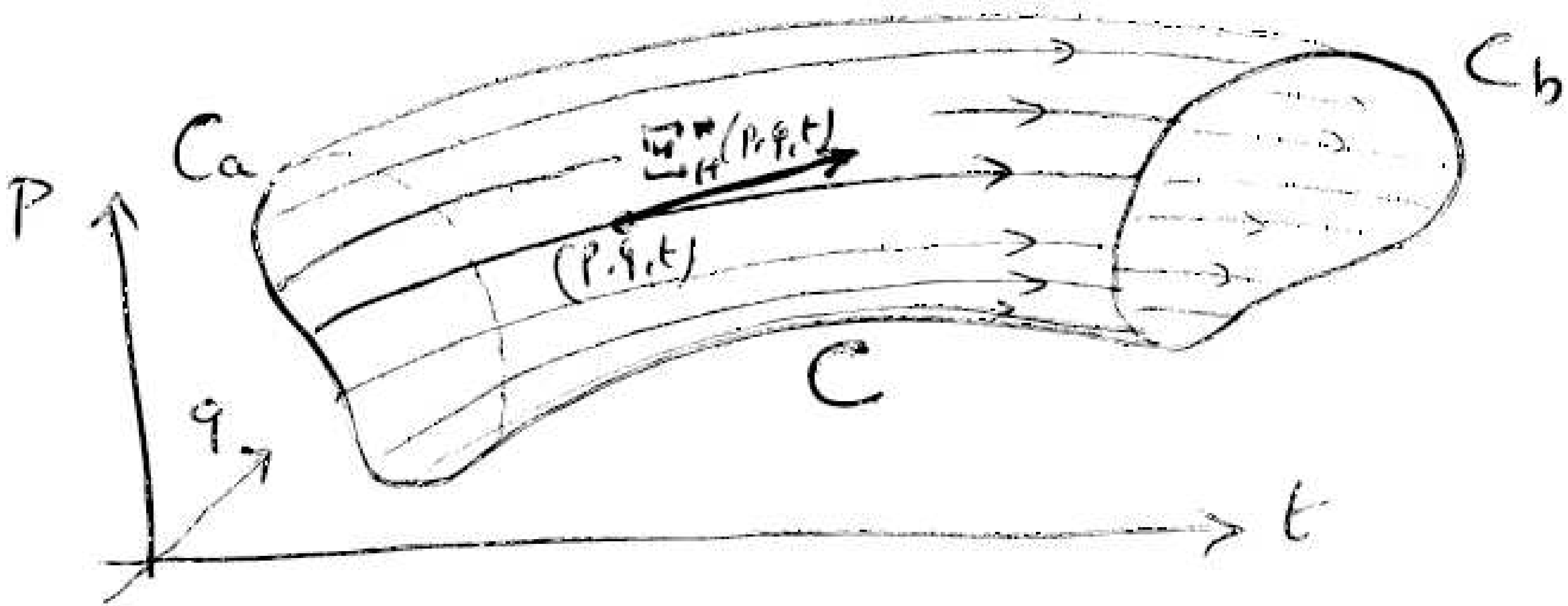
  $$\omega = \partial_{\dot{q}} L.dq - \left(\partial_{\dot{q}} L.\dot{q} - L\right).dt = L.dt + \partial_{\dot{q}} L \varpi$$

  with $\varpi = dq - \dot{q}.dt$

- **P. Dedecker has observed, that the property that among all forms** $\theta \equiv L.dt \bmod \varpi$ **the form** $\omega = p.dq - H.dt$ **is the only one satisfying** $d\theta \equiv 0 \bmod \varpi$ **, is a particular case of more general T. Lepage congruence related to transversality condition.**

[] Cartan E., Leçons sur les invariants intégraux, Hermann, Paris, 1922
[] Dedecker P., A property of differential forms in the calculus of variations, Pacific J. Math. Volume 7, Number 4,p. 1545-1549, 1957
[] Lepage T., Sur les champs géodésiques du calcul des variations, Bull. Acad. Roy. Belg., CL. Sci.27, p.716-729, pp. 1036-1046, 1936

$$\int_{C_a} p \cdot dq - H(p,q)dt = \int_{C_b} p \cdot dq - H(p,q)dt\,.$$



- Legendre transform and contact geometry where used in Mechanic and in Thermodynamic.
- Integral submanifolds of dimension n in 2n+1 dimensional contact manifold are called Legendre submanifolds.
- A smooth fibration of a contact manifold, all of whose are Legendre, is called a Legendre Fibration.
- In the neighbourhood of each point of the total space of a Legendre Fibration there exist contact Darboux coordinates *(z, q, p)* in which the fibration is given by the projection *(z, q, p) =>(z, q).*
- Indeed, the fibres *(z, q)* = *cst* are Legendre subspaces of the standard contact space.
- A Legendre mapping is a diagram consisting of an embedding of a smooth manifold as a Legendre submanifold in the total space of a Legendre fibration, and the projection of the total space of the Legendre fibration onto the base.

- Let us consider the two Legendre fibrations of the standard contact space R^{2n+1} of 1 -jets of functions on R^n :

  $(u, p, q) \mapsto (u, q)$

  and

  $(u, p, q) \mapsto (p.q - u, p)$

- the projection of the 1-graph of a function $u = S(q)$ onto the base of the second fibration gives a Legendre mapping:

  $$q \mapsto \left(q\frac{\partial S}{\partial q} - S(q), \frac{\partial S}{\partial q} \right)$$

- If S is convex, the front of this mapping is the graph of a convex function, the Legendre transform of the function S:

  $\left(S^*(p), p\right)$

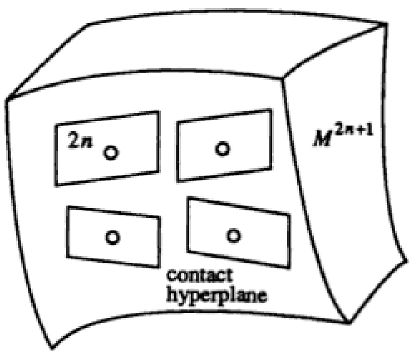
Legendre Duality & Contact Geometry

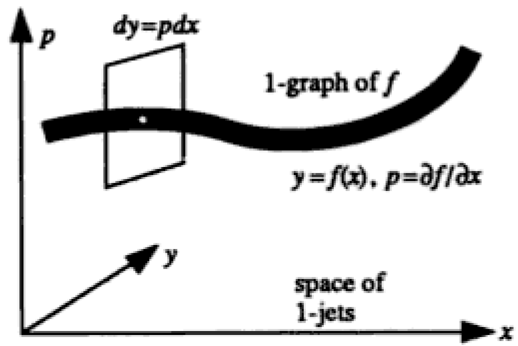
- Symplectic geometry of even-dimensional phase spaces has an odd-dimensional twin: contact geometry.
- The relation between contact geometry and symplectic geometry is similar to the relation between linear algebra and projective geometry. Any fact in symplectic geometry can be formulated as a contact geometry fact and vice versa. The calculations are simpler in the symplectic setting, but their geometric content is better seen in the contact version.
- The functions and vector fields of symplectic geometry are replaced by hypersurfaces and line fields in contact geometry.
- Each contact manifold has a symplectization, which is a symplectic manifold whose dimension exceeds that of the contact manifold by one.
- Symplectic manifolds have contactizations whose dimensions exceed their own dimensions by one.
- If a manifold has a serious reason to be odd dimensional it usually carries a natural contact structure.
- one might now say ‘‘symplectic geometry is all geometry,’’ but I prefer to formulate it in a more geometrical form: contact geometry is all geometry.

[] Arnold, V.I., Givental, A.G., Symplectic geometry, Encyclopedia of mathematical science, vol. 4., Springer Verlag (translated from Russian)

Legendre Duality & Contact Geometry

- **Contact structures and Legendre submanifolds:**
  - A contact structure on an odd-dimensional manifold M2n+1 is a field of hyperplanes(of linear subspaces of codimension 1) in the tangent spaces to M at all its points.
  - All the generic fields of hyperplanes of a manifold of a fixed dimension are locally equivalent. They define the (local) contact structures.
- **Example: A 1-jet of a function $y = f(x_1, x_2, ..., x_n)$ at point *x* of manifold V^n is defined by the point $(x, y, p) \in R^{2n+1}$ where $p_i = \partial f / \partial x_i$ The natural contact structure of this space is defined by the following condition: the 1-graphs $\{x, y = f(x), p = \partial f / \partial x\} \subset J^1(V^n, R)$ of all the functions on *V* should be the tangent structure hyperplane at every point. In coordinates, this conditions means that the 1-form $dy - p.dx$ should vanish on the hyperplanes of the contact field.**





Gibbs contact structure In Thermodynamics

$$dE = TdS - pdV$$

THALES

Projective duality and Legendre transformation

- **A contact structure on a manifold is a nondegenerate field of tangent hyperplanes**
- **The manifold of contact elements in projective space coincides with the manifold of contact elements of the dual projective space**
  - A contact element in projective space is a pair, consisting of a point of the space and of a hyperplane containing this point. The hyperplane is a point of the dual projective space and the point of the original space defines a contact element of the dual space.
- **The manifold of contact elements of the projective space has two natural contact structures:**
  - The first is the natural contact structure of the manifold of contact elements of the original projective space.
  - The second is the natural contact structure of the manifold of contact elements of the dual projective space
- **The dual of the dual hypersurface is the initial hypersurface (at least if both are smooth for instance for the boundaries of convex bodies)**
- **The affine or coordinate version of the projective duality is called the Legendre transformation. Thus contact geometry is the geometrical base of the theory of Legendre transformation.**

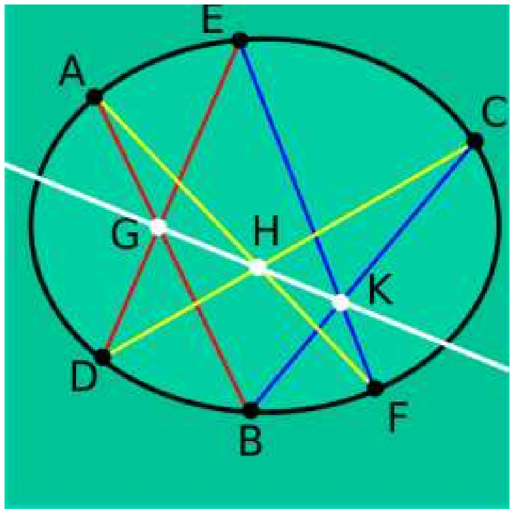
THALES

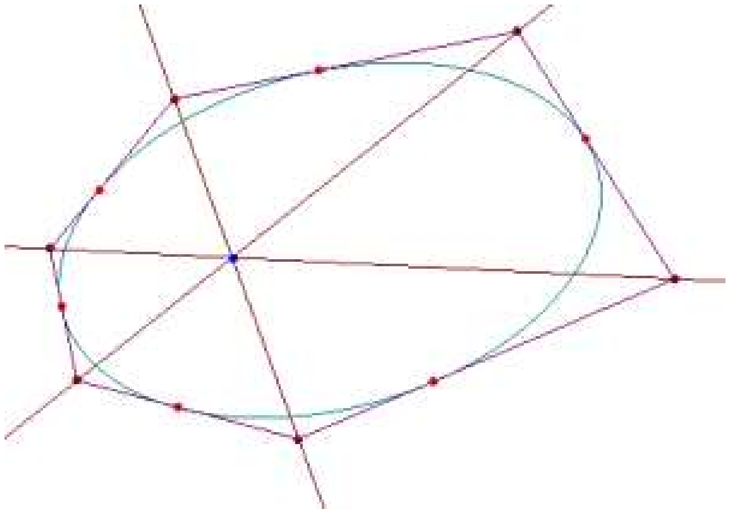
Duality in Projective Geometry: Pascal's Mystic Hexagram

- **In projective geometry, Pascal's theorem (the Hexagrammum Mysticum Theorem) states that:**

  - if an arbitrary six points are chosen on a conic (i.e., ellipse, parabola or hyperbola) and joined by line segments in any order to form a hexagon, then the three pairs of opposite sides of the hexagon (extended if necessary) meet in three points which lie on a straight line, called the Pascal line of the hexagon

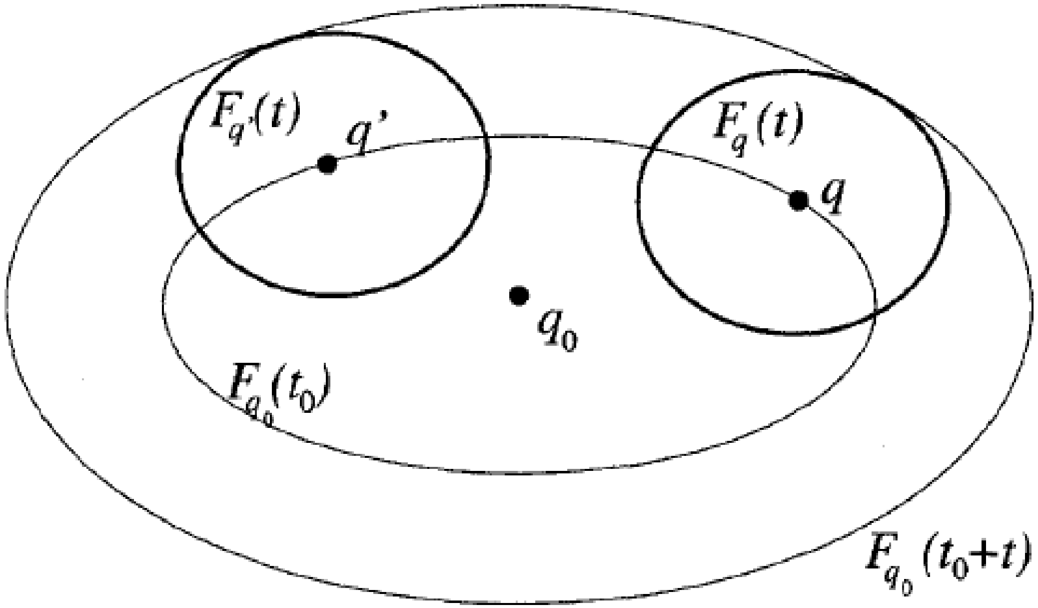
- **The dual of Pascal's Theorem is known as Brianchon's Theorem:**
  - **Pascal's Theorem:** If the vertices of a simple hexagon are points of a point conic, then its diagonal points are collinear.
  - **Brianchon's Theorem:** If the sides of a simple hexagon are lines of a line conic, then the diagonal lines are concurrent.





THALES

- **Huygens Principle**: the wave front $F_{qo}(to + t)$ is the envelope of the fronts $F_q(t)$



- 1872, Sophus Lie: the notion of contact transformation (Beriihrungstransformation) as a geometric tool to study systems of differential equations.
- 19th century and the beginning of the 20th century, F. Engel, H. Poincaré, E. Goursat and E. Cartan

www.thalesgroup.com

Probability in Metric Spaces by Fréchet

THALES

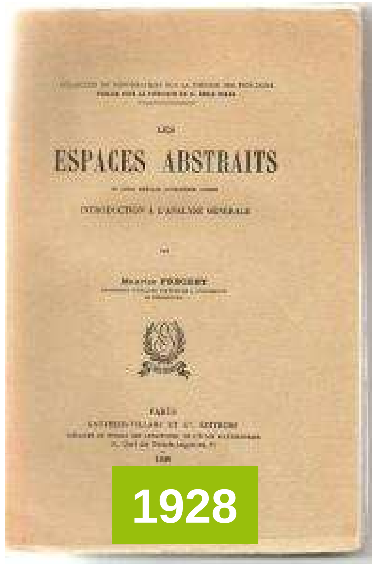
Maurice René Fréchet: Probability in Metric Spaces

Jacques Hadamard

PhD Supervisor

Maurice René Fréchet

1928 Book from Fréchet PhD
Abstract Spaces
Invention of Metric Space

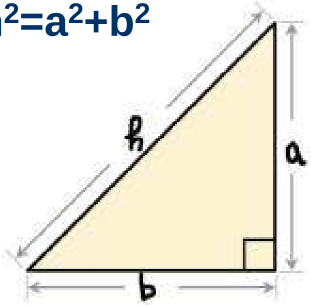


1948 (Annales de l'IHP)
Les éléments aléatoires de nature quelconque
dans un espace distancié
Extension of Probability/Statistic in abstract/Metric space

« Mean » of structured data: Fréchet Barycenter in Metric Space

Right Triangle {a,b,h}

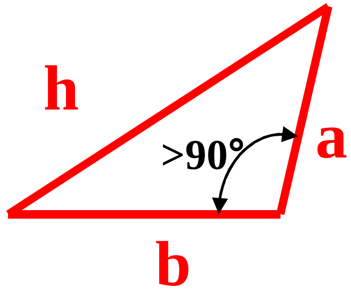
$h^2=a^2+b^2$

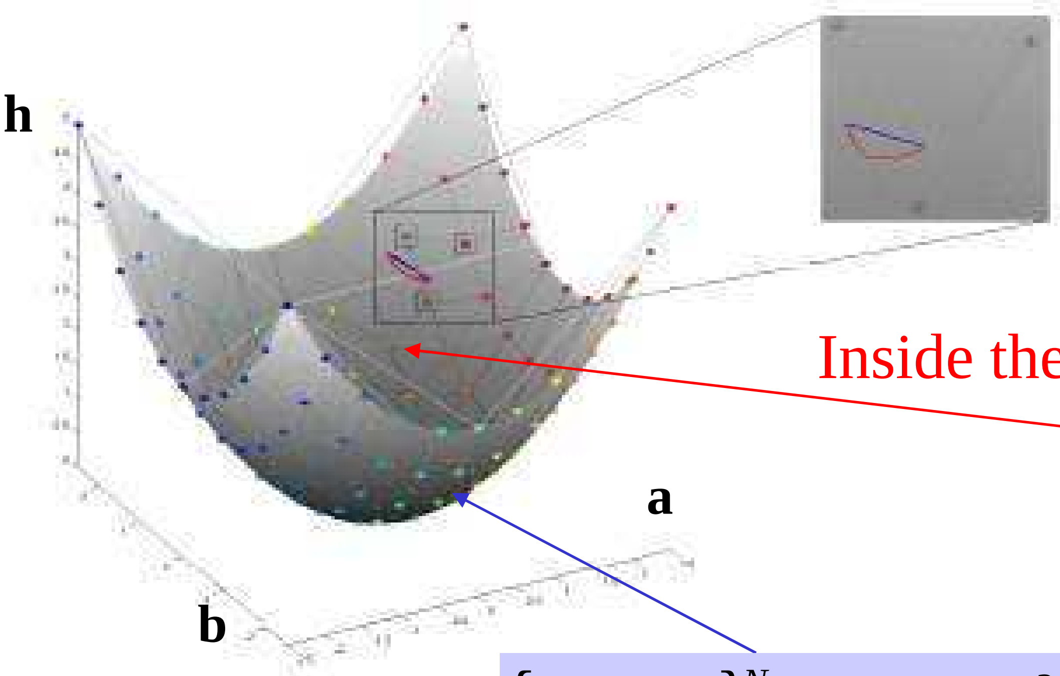


$$\{A,B,H\} = \arg \underset{\{A,B,H\}}{Min} \sum_{i=1}^{N} d^{\,p}_{geodesique}\left(\{a_i,b_i,h_i\},\{A,B,H\}\right)$$

p=2 : Mean, p=1 Median

Ambligone Triangle





Inside the **cone** (**bounded domain**)

$$h^2 > a^2 + b^2$$

$$\{a_i,b_i,h_i\}_{i=1}^{N} \quad \text{with} \quad h_i^2 = a_i^2 + b_i^2$$

THALES

Toeplitz Hermitian PD Matrices: Simple case n=2

$$\Omega = \begin{bmatrix} h & a-ib \\ a+ib & h \end{bmatrix} > 0$$

$$\det \Omega > 0 \Leftrightarrow h^2 > a^2 + b^2$$

Toeplitz Hermitian Positive Definite

Scale Parameter

Shape Parameter

$$\{\log(h), \mu\} \quad \text{with} \quad \mu = \frac{a+ib}{h}$$

$$\log(h) \in R$$

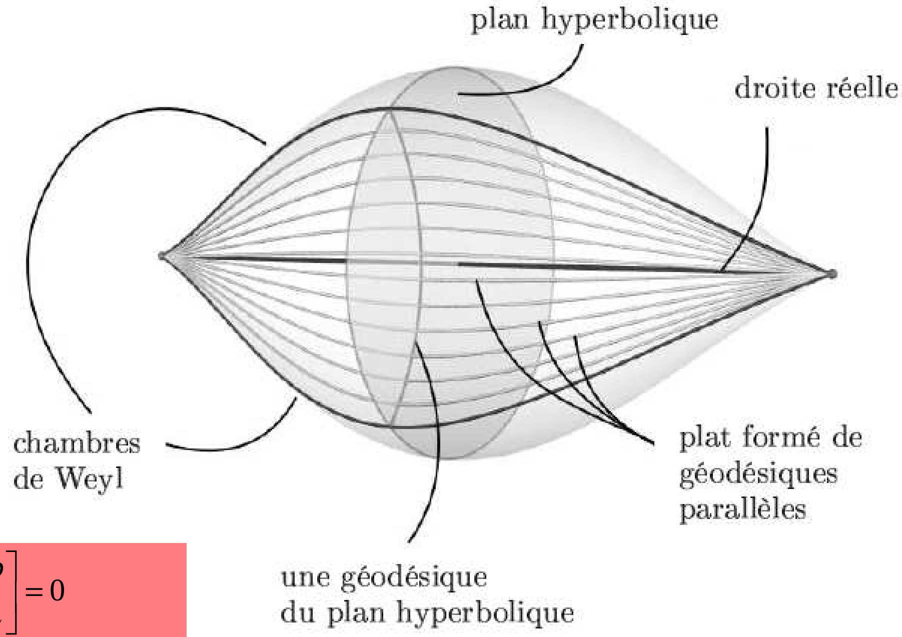
$$\mu \in D = \{z / |z| < 1\}$$

$$\Omega = h. \begin{bmatrix} 1 & \mu^* \\ \mu & 1 \end{bmatrix} > 0$$

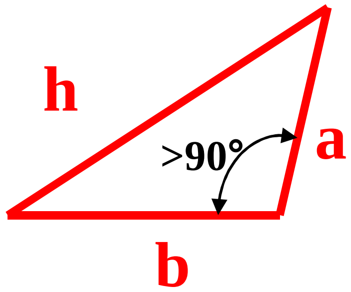
$$h \in R_+^*$$

$$\mu = \frac{a+ib}{h} \in D = \{z / |z| < 1\}$$

D : Poincaré Unit Disk



Hadamard Compactification



Ambligone Triangle
$h^2>a^2+b^2$

$$\det \begin{bmatrix} h-\lambda & a-ib \\ a+ib & h-\lambda \end{bmatrix} = 0$$

$$\Rightarrow \lambda^2 - 2h\lambda + \left(h^2 - \left(a^2 + b^2\right)\right) = 0$$

$$\Rightarrow \lambda = h \pm \sqrt{a^2 + b^2}$$

$$\lambda > 0 \Leftrightarrow h^2 > a^2 + b^2$$

Short Summary of the approach

$$\Omega = \begin{bmatrix} h & a-ib \\ a+ib & h \end{bmatrix}$$

Toeplitz HPD matrix

New parameterization $\{h,\mu\} \in R_+^* \times D$

$$\Omega = h.\begin{bmatrix} 1 & \mu^* \\ \mu & 1 \end{bmatrix} > 0$$

$$\Rightarrow \mu = \frac{a+ib}{h} \in D = \{z \in C / |z| < 1\}$$

Hadamard Compactification :

$\{\log(h), \mu\} \in R \times D$

$$ds^2 = \left(\frac{dh}{h}\right)^2 + \frac{|d\mu|^2}{\left[1-|\mu|^2\right]^2} = \begin{bmatrix} d\log(h) \\ d\mu \end{bmatrix}^+ \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\left[1-|\mu|^2\right]^2} \end{bmatrix} \begin{bmatrix} d\log(h) \\ d\mu \end{bmatrix}$$

$\{\log(h)_{Mean}, \mu_{mean}\}$ **FRECHET Barycenter in Metric Space**

$$= \arg \underset{\{\log(h),\mu\}}{Min} \sum_{i=1}^{N} d^{\,p}_{geodesique}\left(\{\log(h_i), \mu_i\}, \{\log(h), \mu\}\right)$$

$\{\log(h)_{Mean}, \mu_{mean}\}$

$$\Rightarrow \Omega_{Mean} = \begin{bmatrix} e^{\log(h_{mean})} & h_{mean}.\mu^*_{mean} \\ h_{mean}.\mu_{mean} & e^{\log(h_{mean})} \end{bmatrix}$$

Maurice Fréchet : Random Elements on Abstract Spaces

M. Fréchet, "les éléments aléatoires de nature quelconque dans un espace distancié", Annales de l'IHP, t.10, n°4, p.215-310, 1948(voir sur Numdam)

- **Il constitue la mise au point de certains des cours que nous avons fait à la Sorbonne en 1945 et en 1946**
- **Formidable *extension du Calcul [des probabilités qui] résulte de la simple introduction de la notion de distance de deux éléments aléatoires***
- **La nature, la science et la technique offrent de nombreux exemples d'éléments aléatoires qui ne sont, ni des nombres, ni des séries, ni des vecteurs, ni des fonctions. Telles sont par exemple, *la forme d'un fil jeté au hasard sur une table, la forme d'un oeuf pris au hasard dans un panier d'oeufs*. On a ainsi une *courbe aléatoire, une surface aléatoire*. On peut aussi considérer d'autres éléments mathématiques aléatoires : par exemple des *transformations aléatoires de courbe en courbe*.**
- **Il paraît certain que l'urbanisme conduira à étudier des éléments aléatoires tels que la *forme d'une ville prise au hasard*, vérifier par exemple, d'une manière scientifique l'hypothèse de la tendance au développement des villes vers l'Ouest.**
- **Parmi les premières notions à généraliser figurent celles de loi de probabilité, de valeurs typiques (moyenne, équiprobable, etc.), de dispersion, de convergence stochastique, que nous allons d'abord examiner.**
- **Nous indiquerons plus loin en Note additionnelle, un moyen très général d'étendre la notion de fonction de répartition à des éléments aléatoires de nature quelconque.**

Appartement ou vécu Fréchet 12 Square Desnouettes, 15ème, Paris

THALES

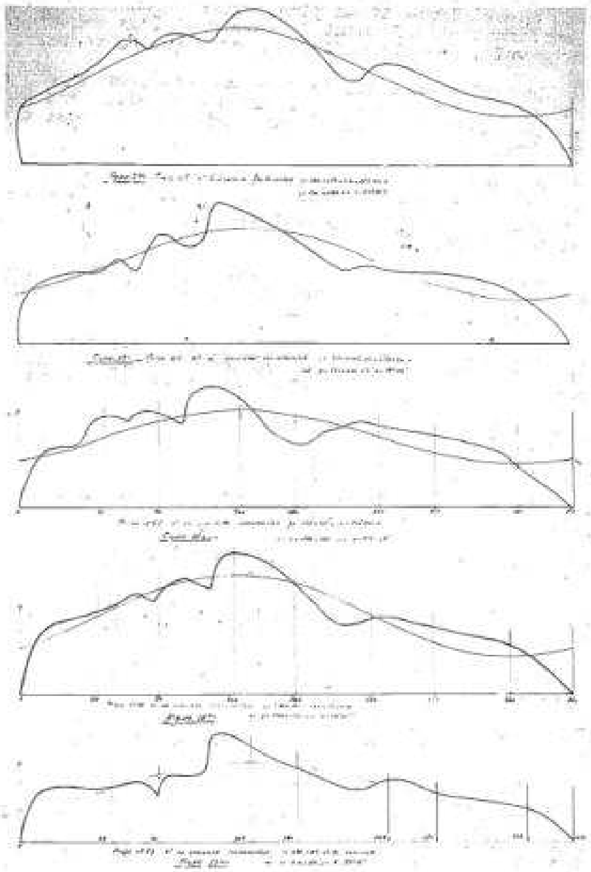
Maurice Fréchet : Random Elements on Abstract Space

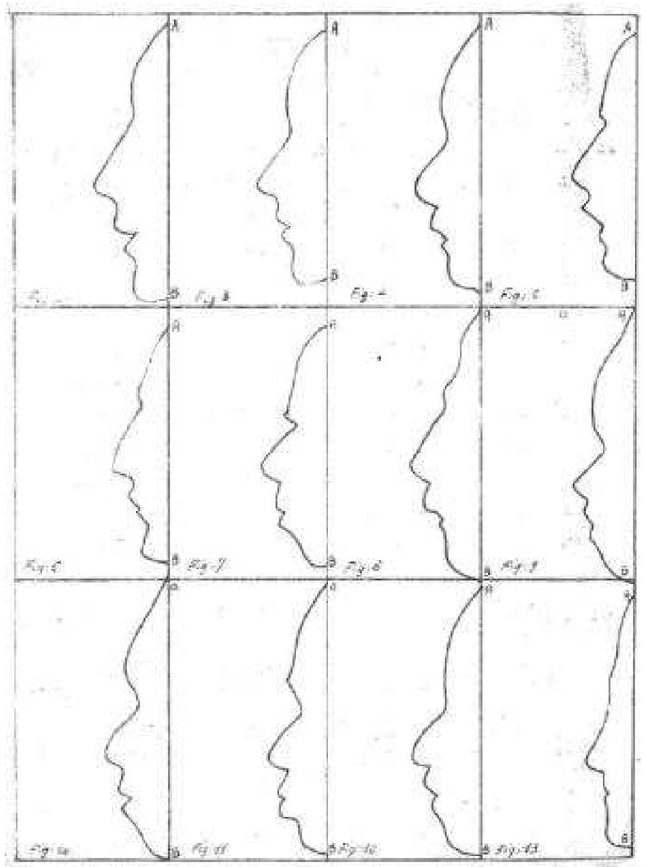
- Nous avons entrepris une étude expérimentale de quelques exemples particuliers. Ceux-ci sont actuellement : *la forme aléatoire d'un fil jeté sur une table, la forme aléatoire d'une section horizontale (prise à un niveau anatomique déterminé) d'un crâne humain* choisi au hasard.
- Pour ce second exemple, notre choix s'est porté sur deux collections de contours obtenus au "conformateur" d'une part sur des personnes vivantes (mais anonymes), *contours fournis gracieusement par la grande maison parisienne de chapellerie Sools*, d'autre part sur une *collection ethnographique de crânes* d'individus morts, collection mise aimablement à notre disposition par M. Lester, Directeur du laboratoire d'Ethnographie.
- L'étude statistique d'une *collection de contours crâniens préhistoriques*.
- Les fonctions représentatives sont *développées en série de fourier et l'on procédera à une analyse statistique de ces séries*. On déterminera les lois de fréquences des premiers coefficients pris isolément et l'on étudiera d'autre part leur corrélation au moyen de divers indices de corrélation et en particulier l'indice diagonal que nous avons défini récemment.
- Le laboratoire de Calcul de l'Institut Henri Poincaré opérant sur deux analyseurs harmoniques.
- Pour les fils, environ 120 coefficients (nous disposons de 100 de ces formes). Pour les contours crâniens dont les formes sont infiniment moins variées, nous nous sommes contentés de 12 harmoniques soient 25 coefficients pour chaque crâne.
- Dévelopement en série de Fourier de l'équation en coordonnées polaires d'un contour crânien

Fond Fréchet : French Science Academy Archive

M. Fréchet collaboration avec M. Ozil

- **M. Ozil a dès 1938 émis l'idée qu'on pourrait aider à la classification des races au moyen de l'analyse harmonique des profils humains. Les tableaux de M. Ozil ne correspondent pour des formes en fait assez compliqués qu'à 9 harmoniques et représentent seulement la moitié expressive du profil, c'est à dire une fonction non périodique.**





Documents numérisés à partir du Fond Fréchet des Archives de l'Académie des Sciences: Correspondance entre M. Fréchet et M. Ozil

THALES

www.thalesgroup.com

Metric Geometry of Hermitian Positive Definites Matrices (HPD):

What is the good metric ?

THALES

Probability Metrics: Distance between 2 random variables

Question in Probability and Statistic: Could we define distance between 2 random variables ?

$$Z^{(1)} = \begin{bmatrix} z_1^{(1)} \\ z_2^{(1)} \\ \vdots \\ z_{n-1}^{(1)} \\ z_n^{(1)} \end{bmatrix} \longleftrightarrow d\left(Z^{(1)}, Z^{(2)}\right) = ??? \longleftrightarrow Z^{(2)} = \begin{bmatrix} z_1^{(2)} \\ z_2^{(2)} \\ \vdots \\ z_{n-1}^{(2)} \\ z_n^{(2)} \end{bmatrix}$$

Or equivalently, could we define distance between 2 probability densities of these 2 random variables ?

$$p\left(Z^{(1)} / \theta^{(1)}\right) \longleftrightarrow d\left(\theta^{(1)}, \theta^{(2)}\right) = ??? \longleftrightarrow p\left(Z^{(2)} / \theta^{(2)}\right)$$

Point of views of Probabilists: Information Geometry of Rao

Information Geometry:

- **Cramer-Rao-Fréchet-Darmois Bound and Fisher Information Matrix**

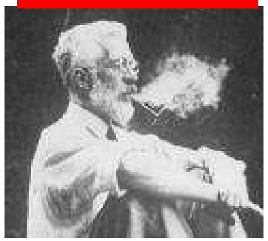
CRFD Bound

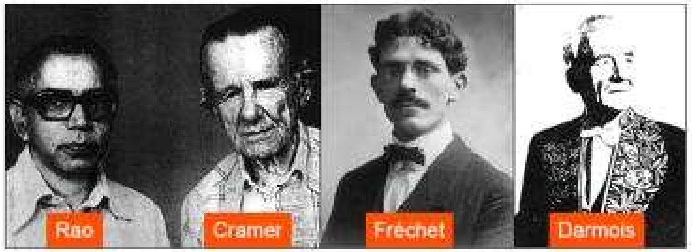
$$E\left[\left(\theta-\hat{\theta}\right)\left(\theta-\hat{\theta}\right)^{+}\right]\geq I(\theta)^{-1}$$

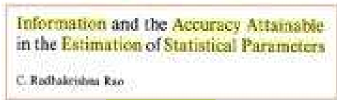
Fisher Information Matrix

$$[I(\theta)]_{i,j}=-E\left[\frac{\partial^2 \ln p(X/\theta)}{\partial\theta_i\partial\theta_j^*}\right]$$

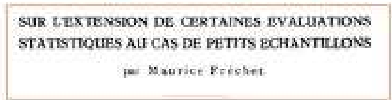
R.A. Fisher






Information and the Accuracy Attainable in the Estimation of Statistical Parameters
C. Radhakrishna Rao

1945


SUR L'EXTENSION DE CERTAINES EVALUATIONS STATISTIQUES AU CAS DE PETITS ECHANTILLONS
par Maurice Fréchet

1943
(IHP Lecture
1939)

- **Kulback-Leibler Divergence (variational definition by Donsker/Varadhan) :**

$$K(p,q)=\underset{\phi}{Sup}\left[E_p(\phi)-\ln E_q\left(e^{\phi}\right)\right]=\int p(x/\theta)\ln\left(\frac{p(x/\theta)}{q(x/\theta)}\right)dx$$

- **Rao-Chentsov Metric (invariance by non-singular change of parameterization)**

$$ds^2=K\left[p(X/\theta),p(X/\theta+d\theta)\right]=d\theta^{+}I(\theta)d\theta=\sum_{i,j}g_{i,j}d\theta_i d\theta_j^*$$

- **Invariance:** $w=W(\theta)\Rightarrow ds^2(w)=ds^2(\theta)$

N. N. Chentsov

Combinatorial Fundation of Kullback Divergence

- Kullback Divergence can be naturally introduced by combinatorial elements and stirling formula :

Let multinomial Law of N elements spread on M levels $\{n_i\}$

$$P_M(n_1,n_2,...,n_M / q_1,...,q_M) = N! \prod_{i=1}^{M} \frac{q_i^{n_i}}{n_i!}$$

with q_i **priors ,** $\sum_{i=1}^{M} n_i = N$ **and** $p_i = \frac{n_i}{N}$

Stirling formula gives : $n! \approx n^n.e^{-n}.\sqrt{2.\pi.n}$ when $n \to +\infty$

$$\underset{N\to+\infty}{Lim} \frac{1}{N} \log[P_M] = \sum_{i=1}^{M} p_i.\log\left[\frac{p_i}{q_i}\right] = K(p,q)$$

Variational Foundation of Kullback Divergence

- Donsker and Varadhan have proposed a variational definition of Kullback divergence :

$$K(p,q) = \underset{\phi}{Sup}\left[E_p(\phi) - \log E_q(e^{\phi})\right]$$

THALES

Kullback Divergence & VARADHAN's Variational Approach

- Donsker and Varadhan have proposed a variational definition of Kullback divergence :

$$K(p,q) = \sup_{\phi} \left[E_p(\phi) - \log E_q(e^{\phi}) \right]$$

$$\text{Consider} : \phi(\omega) = \ln\left(\frac{p(\omega)}{q(\omega)}\right)$$

$$\Rightarrow E_p(\phi) - \ln\left(E_q(e^{\phi})\right) = \sum_{\omega} p(\omega)\ln\left(\frac{p(\omega)}{q(\omega)}\right) - \ln\left[\sum_{\omega} q(\omega)\frac{p(\omega)}{q(\omega)}\right] = K(p,q) - \ln(1) = K(p,q)$$

This proves that the supremum over all ϕ is no smaller than the divergence

$$E_p(\phi) - \ln\left(E_q(e^{\phi})\right) = E_p\left[\ln\left(\frac{e^{\phi}}{E_q(e^{\phi})}\right)\right] = \sum_{\omega} p(\omega)\left(\ln\left[\frac{q^{\phi}(\omega)}{q(\omega)}\right]\right)$$

$$\text{with} \quad q^{\phi}(\omega) = \frac{q(\omega)e^{\phi(\omega)}}{\sum_{\theta} q(\theta)e^{\phi(\theta)}} \Rightarrow K(p,q) - \left[E_p(\phi) - \ln\left(E_q(e^{\phi})\right)\right] = \sum_{\omega} p(\omega)\left[\ln\left(\frac{p(\omega)}{q^{\phi}(\omega)}\right)\right] \geq 0$$

Using the divergence inequality,

- Link with « Large Deviation Theory » & Fenchel-Legendre Transform which gives that logarithm of generating function are dual to Kullback Divergence :

$$\log\left[\int e^{V(x)} q(x)dx\right] = \sup_{p}\left[\int V(x)\,p(x)dx - K(p,q)\right]$$

$$\Leftrightarrow K(p,q) = \sup_{V(.)}\left[\int V(x)\,p(x)dx - \log\left[\int e^{V(x)} q(x)dx\right]\right]$$

$$\Leftrightarrow K(p,q) = \sup_{V(.)}\left[E_p(V) - \log E_q\left[e^{V(x)}\right]\right]$$

Varadhan
Abel Prize 2007

THALES

Dual Information Geometry: Multivariate Gaussian Law

Dual Coordinates systems & Potential functions

- Potential Functions are Dual and related by **Legendre transformation** :

$$\text{Dual coordinates} \begin{cases} \tilde{\Theta} = (\theta, \Theta) = \left(\Sigma^{-1}m, (2\Sigma)^{-1}\right) \\ \tilde{H} = (\eta, H) = \left(m, -\Sigma + mm^T\right) \end{cases}$$

$$\Rightarrow \begin{cases} \tilde{\Psi}(\tilde{\Theta}) = 2^{-2}Tr\left(\Theta^{-1}\theta\theta^T\right) - 2^{-1}\log(\det \Theta) + 2^{-1} n \log(\pi) \\ \tilde{\Phi}(\tilde{H}) = -2^{-1}\log\left(1 + \eta^T H^{-1}\eta\right) - 2^{-1}\log(\det(-H)) - 2^{-1} n \log(2\pi e) \end{cases}$$

$$\begin{cases} \dfrac{\partial \tilde{\Psi}}{\partial \theta} = \eta \\ \dfrac{\partial \tilde{\Psi}}{\partial \Theta} = H \end{cases} \text{ and } \begin{cases} \dfrac{\partial \tilde{\Phi}}{\partial \eta} = \theta \\ \dfrac{\partial \tilde{\Phi}}{\partial H} = \Theta \end{cases}$$

$$\tilde{\Phi} \equiv \left\langle \tilde{\Theta}, \tilde{H} \right\rangle - \tilde{\Psi}$$

$$\text{with } \left\langle \tilde{\Theta}, \tilde{H} \right\rangle = Tr\left(\theta\eta^T + \Theta H^T\right)$$

$$\tilde{\Phi}(\tilde{H}) = E[\log p]$$ ← **Entropy**

- Hessians are convexe and define Riemannian metrics :

$$g_{ij} = \frac{\partial^2 \tilde{\Psi}}{\partial \Theta_i \partial \Theta_j} \text{ and } g^*_{ij} \equiv \frac{\partial^2 \tilde{\Phi}}{\partial H_i \partial H_j}$$

THALES

Link with Kullback Divergence

- For each Dual Geometry, we can built a divergence that is directly related to Kulback-Leibler Divergence :

$$Div[N(m_2,\Sigma_2),N(m_1,\Sigma_1)] \equiv \tilde{\Psi}(\tilde{\Theta}_2)+\tilde{\Phi}(\tilde{H}_1)-\langle \tilde{\Theta}_2,\tilde{H}_1\rangle \geq 0$$

$$Div^*[N(m_2,\Sigma_2),N(m_1,\Sigma_1)] \equiv \tilde{\Psi}(\tilde{\Theta}_1)+\tilde{\Phi}(\tilde{H}_2)-\langle \tilde{\Theta}_1,\tilde{H}_2\rangle \geq 0$$

$$Div[N(m_2,\Sigma_2),N(m_1,\Sigma_1)] = \int p(x/m_1,\Sigma_1)\log\frac{p(x/m_1,\Sigma_1)}{p(x/m_2,\Sigma_2)}dx$$

$$Div(N(m_2,\Sigma_2),N(m_1,\Sigma_1)) = \frac{1}{2}\begin{bmatrix} -\log\left(\det\left(\Sigma_1\Sigma_2^{-1}\right)\right)+Tr\left(\Sigma_1\left(\Sigma_2^{-1}-\Sigma_1^{-1}\right)\right) \\ +Tr\left(\Sigma_2^{-1}(m_1-m_2)(m_1-m_2)^T\right) \end{bmatrix}$$

Riemannian Metric

- As Potential are convexe, their Hessians define Riemannian Metrics :

$$ds^2 = \frac{1}{2}g_{ij}d\Theta_i d\Theta_j + O\left(|d\Theta_i|^3\right) = \frac{1}{2}g^{ij}dH_i dH_j + O\left(|dH_i|^3\right)$$

THALES

Information Geometry of Multivariate Gaussian Density

Multivariate Gaussian Distribution of zero mean

- Riemannian Information metric is given by Rao Metric

$$ds^2 = \frac{1}{2} Tr\left[\left(\Sigma^{-1} d\Sigma \right)^2 \right] = \frac{1}{2} \left\| \Sigma^{-1/2} d\Sigma \Sigma^{-1/2} \right\|^2$$

$$\text{with } \quad \|A\|^2 = \langle A, A \rangle \ \text{ et } \ \langle A, B \rangle = Tr\left(AB^T \right)$$

Rao Metric

$$\text{As } \quad \Sigma^{-1}\Sigma = I \Rightarrow d\Sigma^{-1}\Sigma = -\Sigma^{-1} d\Sigma$$

$$\text{then } \quad ds_{\Sigma}^2 = ds_{\Sigma^{-1}}^2$$

Invariance by inversion

$$D\left(\Sigma_1, \Sigma_2 \right) = D\left(\Sigma_1^{-1}, \Sigma_2^{-1} \right)$$

Multivariate Gaussian of non zero mean

- Isometries are given by following homeomorphisms :

$$(m, \Sigma) \rightarrow (m', \Sigma') = \left(A^T m + a, A^T \Sigma A \right) \text{ isometry}$$

$$ds^2 \mapsto ds'^2 = ds^2 \quad \text{with } \ (a, A) \in R^n \text{x} GL(n, R)$$

$$ds^2 = dm^T \Sigma^{-1} dm + \frac{1}{2} Tr\left[\left(\Sigma^{-1} d\Sigma \right)^2 \right]$$

THALES

Canonical Exemple: Monovariate Gaussian Law

Gauss-Laplace Law

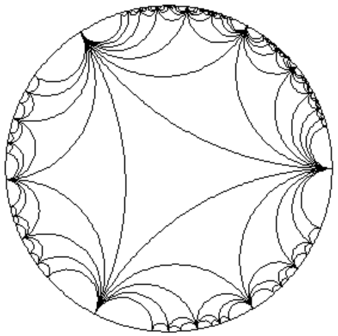
- Fisher Information Matrix for Gauss-Laplace Gaussian Law:

$$I(\theta)=\sigma^{-2}.\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix} \quad \text{with} \quad E\left[\left(\theta-\hat{\theta}\right)\left(\theta-\hat{\theta}\right)^T\right]\geq I(\theta)^{-1} \quad \text{and} \quad \theta=\begin{pmatrix}m\\ \sigma\end{pmatrix}$$

- Rao-Chentsov Metric of Information Geometry

$$ds^2=d\theta^T.I(\theta).d\theta=\frac{dm^2}{\sigma^2}+2.\frac{d\sigma^2}{\sigma^2}=2.\sigma^{-2}\left[\left(\frac{dm}{\sqrt{2}}\right)^2+(d\sigma)^2\right]$$

H. Poincaré



- Fisher Metric is equal to Poincaré Metric for Gaussian Laws:

$$z=\frac{m}{\sqrt{2}}+i.\sigma \qquad \omega=\frac{z-i}{z+i} \quad \left(|\omega|<1\right) \qquad \Rightarrow ds^2=8.\frac{|d\omega|^2}{\left(1-|\omega|^2\right)^2}$$

$$d^2\left(\{m_1,\sigma_1\},\{m_2,\sigma_2\}\right)=2.\left(\log\frac{1+\delta(\omega^{(1)},\omega^{(2)})}{1-\delta(\omega^{(1)},\omega^{(2)})}\right)^2$$

$$\text{with} \quad \delta(\omega^{(1)},\omega^{(2)})=\left|\frac{\omega^{(1)}-\omega^{(2)}}{1-\omega^{(1)}\omega^{(2)*}}\right|$$

Point of View of Geometers: Cartan-Siegel Domains Metric

« Il est clair que si l'on parvenait à démontrer que tous les domaines homogènes dont la forme

$$\Phi = \sum_{i,j} \frac{\partial^2 \log K(z,\bar{z})}{\partial z_i \partial \bar{z}_j} dz_i d\bar{z}_j$$

est définie positive sont symétriques, toute la théorie des domaines bornés homogènes serait élucidée. C'est là un problème de géométrie hermitienne certainement très intéressant »

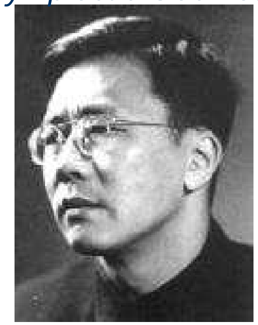
Dernière phrase de Elie Cartan, dans « Sur les domaines bornés de l'espace de n variables complexes », Abh. Math. Seminar Hamburg, 1935

Fine structure of
Information Geometry
(Hessian Geometry,
Kählerian Geometry)

Carl Ludwig Siegel
(Siegel domains in framework of Symplectic Geometry)

Henri Poincaré
(upper-half plane model of hyperbolic geometry)
n=1

Elie Cartan
(classification in 6 types of symmetric homogeneous bounded domains)
n<=3



Lookeng Hua
(Bergman, Cauchy and Poisson Kernels in Siegel domains)

Ernest Vinberg
(link with homogeneous convexe cones, Siegel domains of 2nd kind)

Jean-Louis Koszul
(canonical hermitian form of complex homogeneous spaces, a complex homogeneous space with positive definite canonical hermitian form is isomorphic to a bounded domain,,Study of Affine Transform Groups of locally flat manifolds)

Cartan-Siegel Symmetric Homogeneous Bounded Domains

$$\Omega^{I}_{1,1} = \Omega^{II}_{1} = \Omega^{III}_{1} = \Omega^{IV}_{1} = \{ z \in C / zz^{*} < 1 \}, K(z,w^{*}) = \frac{1}{(1 - zw^{*})^{2}}$$

Z : Complex Rectangular Matrix

$ZZ^{+} < I \quad (^{+} : \textit{transposed} - \textit{conjugate})$

Type I : $\Omega^{I}_{p,q}$ complex matrices with p lines and q rows

Type II : Ω^{II}_{p} complex symmetric matrices of order p

Type III : Ω^{III}_{p} complex skew symmetric matrices of order p

Type IV : Ω^{IV}_{n} complex matrices with n rows and 1 line :

$\left| ZZ^{t} \right| < 1, 1 + \left| ZZ^{t} \right|^{2} - 2ZZ^{+} > 0$

$$K(Z,W^{*}) = \frac{1}{\mu(\Omega)} \det(I - ZW^{+})^{-\nu} \quad \text{for} \begin{cases} \text{Type I} : \Omega^{I}_{p,q} , \nu = p + q \\ \text{Type II} : \Omega^{II}_{p} , \nu = p + 1 \\ \text{Type III} : \Omega^{III}_{p} , \nu = p - 1 \end{cases}$$

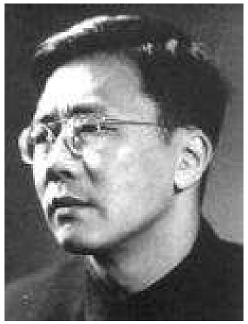
$$K(Z,W^{*}) = \frac{1}{\mu(\Omega)} (1 + ZZ^{t}W^{*}W^{+} - 2ZW^{*})^{-\nu} \text{ for Type IV} : \Omega^{IV}_{n} , \nu = n$$

where $\mu(\Omega)$ is euclidean volume of the domain.

Henri Poincaré
(n=1)

Elie Cartan
(n<=3)

Carl Ludwig Siegel



Lookeng Hua

THALES

Action of SL_2R Group on hyperbolic Poincaré Plan

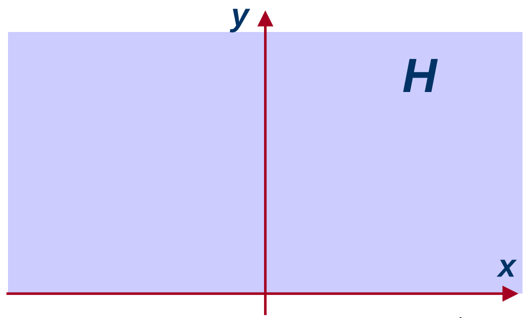
- Möbius Transform is a transitive action that transforms upper half-plane to itself (homogeneous space) :

$$M=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in SL_2R \quad , \quad M(z)=\frac{az+b}{cz+d} \text{ and } z\in H=\{z\in C:\operatorname{Im}(z)>0\}$$

- ***M*** and ***–M*** have same action then we consider the quotient Group :

$$PSL_2R=SL_2R/\langle \pm I_2\rangle$$

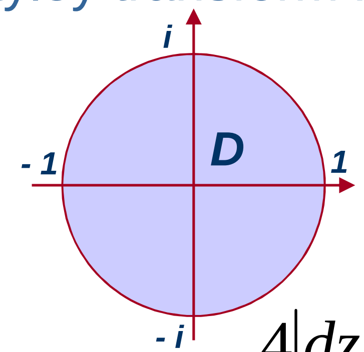
- Complex unit disk is link to upper half plane by Cayley transform :



$$ds^2=\frac{dx^2+dy^2}{y^2}=\frac{|dz|^2}{y^2}$$

$$D=\{z\in C:|z|<1\}$$

$$\left| \begin{array}{l} H\rightarrow D \\ z\mapsto \dfrac{z-i}{z+i} \end{array} \right. \quad \text{et} \quad \left| \begin{array}{l} D\rightarrow H \\ z\mapsto i\dfrac{1+z}{1-z} \end{array} \right.$$



$$ds^2=\frac{4|dz|^2}{\left(1-|z|^2\right)^2}$$

Carl Ludwig Siegel Contribution in his seminal book « Symplectic Geometry » : a generalization of Poincaré Space

Invariance by Automorphisms: Siegel Metric of SH_n

- **Siegel Metric for the Siegel Upper-Half Plane:**
  - Upper-Half Plane : $SH_n = \{Z = X + iY \in Sym(n,C) / \mathrm{Im}(Z) = Y > 0\}$
    - Isometries of SH_n are given by the quotient group:

$PSp(n,R) \equiv Sp(n,R) / \{\pm I_{2n}\}$ with $Sp(n,F)$ the Symplectic Group:

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \Rightarrow M(Z) = (AZ+B)(CZ+D)^{-1}$$

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(n,F) \Leftrightarrow \begin{cases} A^T C \text{ et } B^T D \quad \text{symmetric} \\ A^T D - C^T B = I_n \end{cases}$$

$$Sp(n,F) \equiv \{M \in GL(2n,F) / M^T JM = J\},\ J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \in SL(2n,R)$$

- Unique Metric invariante by $M(Z)$:

$$ds^2_{Siegel} = Tr\left(Y^{-1}(dZ)Y^{-1}(d\bar{Z})\right) \qquad Z = X + iY$$

$$\begin{cases} X = 0 \\ Y = R_n \end{cases} \Longrightarrow ds^2 = Tr\left(\left(R_n^{-1}(dR_n)\right)^2\right)$$

Siegel Uppel-Half Plane Geometry

Siegel Upper Half Space

$$SH_n = \{ Z = X + iY \in Sym(n,C) \,/\, \mathrm{Im}(Z) = Y > 0 \}$$

$$\det\left(R_1^{-1/2}.R_2.R_1^{-1/2} - \lambda.I\right) = 0$$

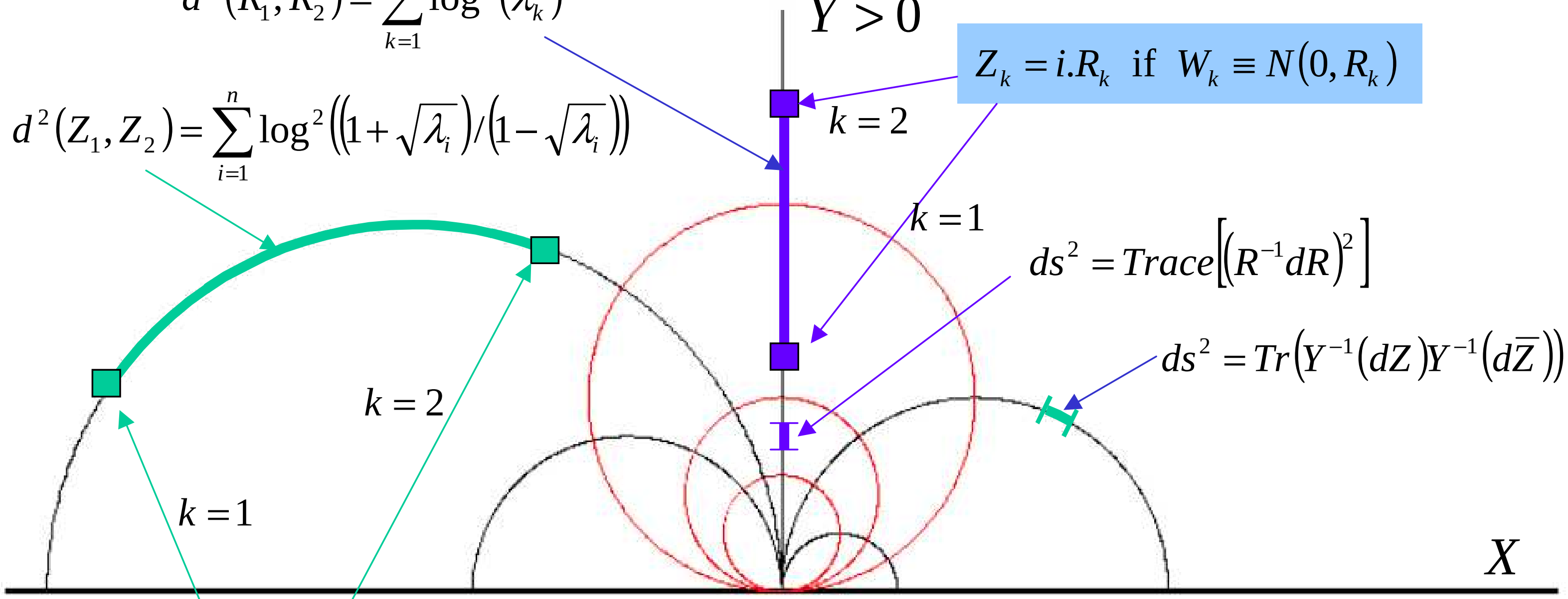
$$d^2(R_1,R_2) = \sum_{k=1}^{n} \log^2(\lambda_k)$$

$$d^2(Z_1,Z_2) = \sum_{i=1}^{n} \log^2\left(\left(1+\sqrt{\lambda_i}\right)/\left(1-\sqrt{\lambda_i}\right)\right)$$

$$Z_k = i.R_k \ \text{ if } \ W_k \equiv N(0,R_k)$$

$$ds^2 = Trace\left[\left(R^{-1}dR\right)^2\right]$$

$$ds^2 = Tr\left(Y^{-1}(dZ)Y^{-1}(d\overline{Z})\right)$$



$$Z_k = X_k + i.Y_k$$

$$R(Z_1,Z_2) = (Z_1 - Z_2)(Z_1 - \overline{Z}_2)^{-1}(\overline{Z}_1 - \overline{Z}_2)(\overline{Z}_1 - Z_2)^{-1}$$

$$\det\left(R(Z_1,Z_2) - \lambda.I\right) = 0$$

THALES

QUANTIZATION IN COMPLEX SYMMETRIC SPACES

$$SD_n = \{Z / ZZ^+ < I\}$$

$$\text{with } g = \begin{bmatrix} A & B \\ B^* & B^* \end{bmatrix} \text{ and } g^t J g = J \text{ with } J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

$$\text{where } \begin{cases} A^+ A - B^t B^* = I \\ B^+ A - A^t B^* = 0 \end{cases} \text{ with } g(Z) = (AZ + B)(B^* Z + A^*)^{-1}$$

$$\text{Kähler potential}: F(z) = -\log\det(I - Z^+ Z) = -trace \log(I - Z^+ Z)$$

$$F(g(Z)) = F(Z) + 2\,\mathrm{Re}\, trace \log(A^* + B^* Z) \Rightarrow \partial\partial^* F(g(Z)) = \partial\partial^* F(Z)$$

$$g(0) = B(A^*)^{-1} \Rightarrow F(g(0)) = \log\det(I + B^+ B) = trace \log(I + B^+ B)$$

F. Berezin

$$\langle f, g \rangle = c(h) \int f(Z) g(Z) \left[\frac{K(Z, Z^*)}{K(0,0)} \right]^{-1/h} d\mu(Z, Z^*)$$

$$c(h)^{-1} = \int \left[\frac{K(Z, Z^*)}{K(0,0)} \right]^{-1/h} d\mu(Z, Z^*) \quad \text{with} \quad K(gZ, gZ^*) j(g, z) j(g, Z)^* = K(Z, Z^*) \ , \ j(g, Z) = \frac{\partial gZ}{\partial Z}$$

$$d\mu(W, W^*) = F(W, W^*) \frac{d\mu_L(W, W^*)}{\pi^n}$$

$$ds^2 = \sum_{\alpha,\beta} g_{\alpha,\beta} dW^\alpha dW^{\beta *} \text{ with } g_{\alpha\beta} = -\frac{\partial^2 \log F(W, W^*)}{\partial W^\alpha \partial W^{\beta *}} \text{ where } F(W, W^*) = \det(I + WW^+)^{-\nu}$$

Siegel HPD Matrix Metric = IG Metric of Multiv. Gaussian Law

- **Information Geometry for Multivariate Gaussian Law of zero Mean and intrinsec Geometry of Hermitian Positive Definite Matrices (particular case of Siegel Upper-Half Plane) provide the same metric**
  - Information Geometry:

$$p(Z_n / R_n) = (\pi)^{-n} . \left| R_n \right|^{-1} . e^{-Tr\left[\hat{R}_n . R_n^{-1}\right]} \quad \text{with} \quad g_{ij}(\theta) = -E\left[\frac{\partial^2 \ln p(Z_n / \theta_n)}{\partial \theta_i . \partial \theta_j^*} \right]$$

$$\text{with} \quad \hat{R}_n = (Z_n - m_n).(Z_n - m_n)^+$$

$$\text{and} \quad E\left[\hat{R}_n\right] = R_n \qquad m_n = 0 \Longrightarrow ds^2 = Tr\left(\left(R_n^{-1}(dR_n)\right)^2\right)$$

  - Geometry of Siegel Upper-Half Plane:

$$SH_n = \left\{ Z = X + iY \in Sym(n, C) / \operatorname{Im}(Z) = Y > 0 \right\}$$

$$ds^2_{Siegel} = Tr\left(Y^{-1}(dZ)Y^{-1}\left(d\overline{Z}\right)\right) \quad \text{with} \quad Z = X + iY$$

$$\begin{cases} X = 0 \\ Y = R_n \end{cases} \Longrightarrow ds^2 = Tr\left(\left(R_n^{-1}(dR_n)\right)^2\right)$$

THALES

Distance between HPD Matrices: particular case of Siegel

- **Siegel Distance:**
  - Particular Case (X=0) and General Case:
    - Particular Case (pure imaginary axis) : $Z = iR \quad \text{avec} \quad R > 0$

$$d^2(R_1, R_2) = \left\| \log\left(R_1^{-1/2}.R_2.R_1^{-1/2}\right) \right\|^2 = \sum_{k=1}^{n} \log^2(\lambda_k)$$

with $\det\left(R_2 - \lambda R_1\right) = 0$

    - General Case of Siegel Upper-Half Plane Distance:

$$Z = X + iY \in SH_n \quad \text{with} \quad X \neq 0$$

$$d_{Siegel}^2(Z_1, Z_2) = \left(\sum_{k=1}^{n} \log^2\left(\frac{1+\sqrt{\lambda_k}}{1-\sqrt{\lambda_k}} \right) \right) \quad \text{with} \quad Z_1, Z_2 \in SH_n$$

with $\det\left(R(Z_1, Z_2) - \lambda.I\right) = 0$

$$R(Z_1, Z_2) = (Z_1 - Z_2)\left(Z_1 - \overline{Z}_2\right)^{-1}\left(\overline{Z}_1 - \overline{Z}_2\right)\left(\overline{Z}_1 - Z_2\right)^{-1}$$

THALES

Another distance deduced from Siegel Work

$$ds^2 = Tr\left[\left(\Sigma . d\Sigma^{-1}\right)^2\right]$$

Siegel has deduced an other distance from :

$$T_{12} = \left(\Sigma_1^{-1} - \Sigma_2^{-1}\right).\left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1} \quad \text{and} \quad R = T_{12}^2 = T_{12}^+ . T_{12}$$

R is hermitian positive definite matrix with eigen-values :

$$r_k = \left(\frac{1-\lambda_k}{1+\lambda_k}\right)^2 \quad \text{with } r_k = \sigma(R) \quad \text{and} \quad \lambda_k = \sigma(R_n^{(1)1/2} . R_n^{(2)-1} . R_n^{(1)1/2})$$

We deduce that :

$$d^2(\Sigma_1, \Sigma_2) = Tr\left[\ln^2\left(\frac{I_n + R^{1/2}}{I_n - R^{1/2}}\right)\right] = \sum_{k=1}^{n} \ln^2\left(\frac{1 + r_k^{1/2}}{1 - r_k^{1/2}}\right)$$

because :

$$\ln^2\left(\frac{I_n + R^{1/2}}{I_n - R^{1/2}}\right) = 4.\left[\tanh^{-1} R^{1/2}\right]^2 = 4.R.\left(\sum_{k=0}^{\infty} \frac{R^k}{2.k+1}\right)^2 \quad \text{and} \quad Tr\left[R^j\right] = \sum_{k=1}^{n} r_k^j$$

Geometry of Hermitian Positive Definite Matrices given by:

- **Geodesic :** $d(R_X,\gamma(t)) = t.d(R_X,R_Y)$ with $t \in [0,1]$

$$\gamma(t) = R_X^{1/2} e^{t\log\left(R_X^{-1/2} R_Y R_X^{-1/2}\right)} R_X^{1/2} = R_X^{1/2}\left(R_X^{-1/2} R_Y R_X^{-1/2}\right)^t R_X^{1/2}$$

$$\gamma(0) = R_X \quad , \quad \gamma(1) = R_Y \text{ and } \gamma(1/2) = R_X \circ R_Y$$

Properties of this space

- **Symetric Space as studied by Elie Cartan : Existence of bijective geodesic isometry**

$$G_{(A,B)}X = (A\circ B)X^{-1}(A\circ B) \quad \text{avec} \quad A\circ B = A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{1/2}A^{1/2}$$

- **Bruhat-Tits Space : semi-parallelogram inequality**

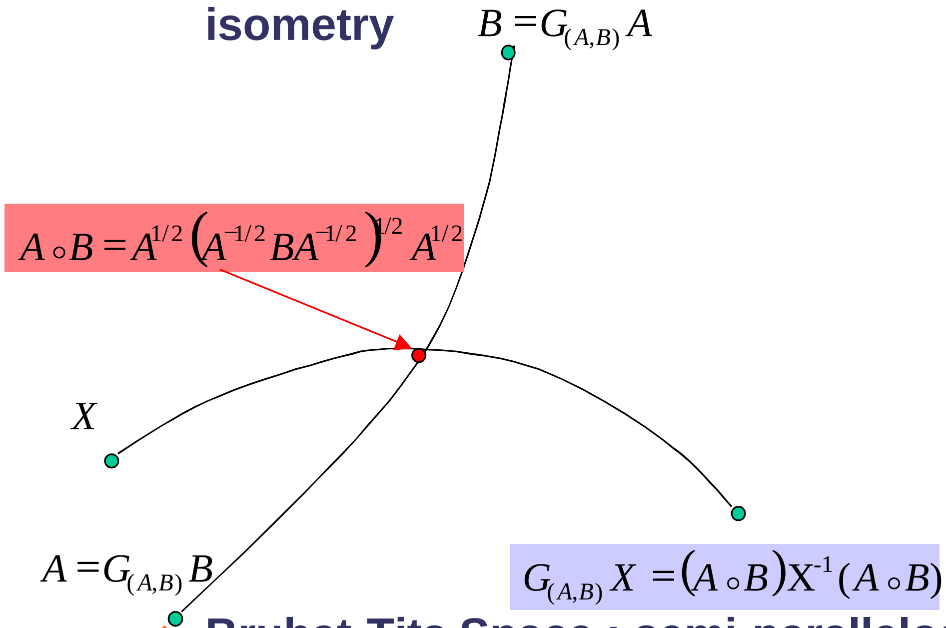
$$\forall x_1, x_2 \quad \exists z \ \text{ tel que } \forall x$$

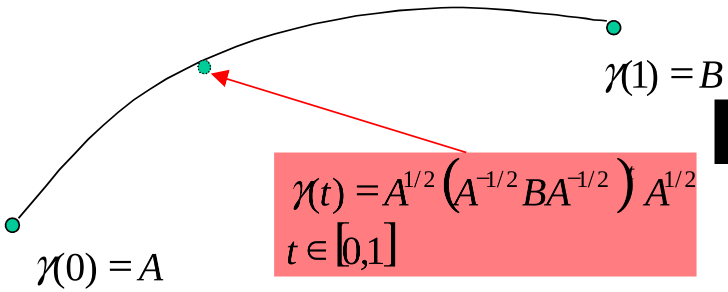
$$d(x_1,x_2)^2 + 4d(x,z)^2 \leq 2d(x,x_1)^2 + 2d(x,x_2)^2 \quad \forall x \in X$$

- **Cartan-Hadamard Space (Complete, simply connected with negative sectional curvature Manifold)**

Cartan's Symmetric Space

- Symetric Space as studied by Elie Cartan : Existence bijective geodesic isometry



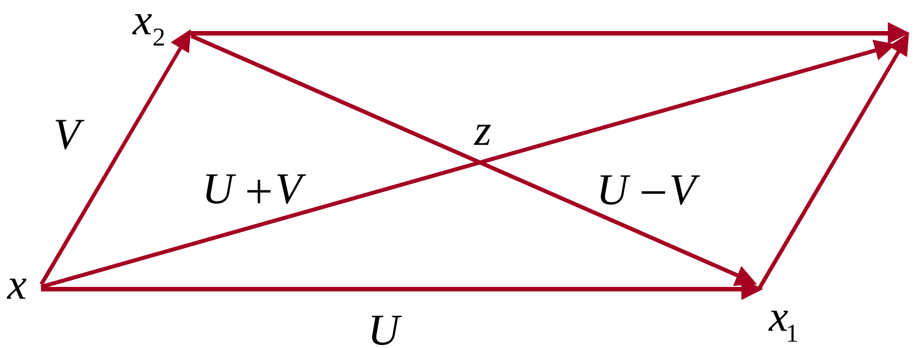


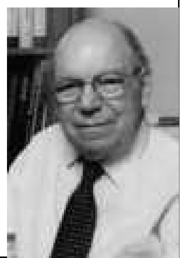

E. J. Cartan


M. Berger

- Bruhat-Tits Space : semi-parallelogram inequality

$$d(x_1,x_2)^2 + 4d(x,z)^2 \le 2d(x,x_1)^2 + 2d(x,x_2)^2$$




J. Tits

$$\|U-V\|^2 + \|U+V\|^2 = 2\|U\|^2 + 2\|V\|^2$$

Symmetric Space

- This isometry for metric space :

$$G_{(A,B)}X = (A \bullet B)\mathrm{X}^{-1}(A \bullet B) \quad \text{with} \quad \begin{cases} A \bullet B = A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{1/2}A^{1/2} \\ \delta^2(A,B) = \left\|\log\left(A^{-1/2}BA^{-1/2}\right)\right\|^2 \end{cases}$$

- Is an extension of this one :

$$G_{(a,b)}x = (a \circ b)x^{-1}(a \circ b) \quad \text{with} \quad \begin{cases} a \circ b = \sqrt{ab} \\ \delta^2(a,b) = \left|\log\left(ab^{-1}\right)\right|^2 \end{cases}$$

- To be compared with Euclidean « symmetric » space

$$G_{(a,b)}x = \left(\frac{a+b}{2}\right) - x + \left(\frac{a+b}{2}\right) \quad \text{with} \quad \begin{cases} a \circ b = \dfrac{a+b}{2} \\ \delta^2(a,b) = |a-b|^2 \end{cases}$$

Unicity of fixed point

- $A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{1/2}A^{1/2}$ **is the only fixed point because :**

$$CX^{-1}C = X \Rightarrow \left(X^{-1/2}CX^{-1/2}\right)\left(X^{-1/2}CX^{-1/2}\right) = I$$

$$\underset{\text{unicity of square root}}{\Rightarrow} X^{-1/2}CX^{-1/2} = I \Rightarrow X = C$$

$$d\left(\left(A \bullet B\right)\mathrm{X}^{-1}(A \bullet B), X\right) = d\left(\left(A \bullet B\right)\mathrm{X}^{-1}(A \bullet B)X^{-1}, I\right) = 2d\left(X, (A \bullet B)\right)$$

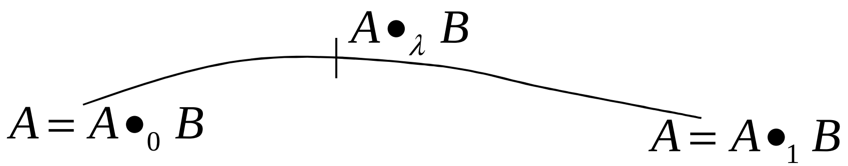
- **due to trace property of :**

$$d(X,(A \bullet B)) = \left(\sum_{k=1}^{n} \log^2 \lambda_i\right)$$

with $\{\lambda_i\}_{i=1}^{n}$ eigenvalues of $(A \bullet B)X^{-1}$

Convexity

- For space of Symmetric Positive Definite matrices, the analogue of $\lambda a + (1-\lambda) b$ is : $$A \bullet_\lambda B = A^{1/2}\left(A^{-1/2} B A^{-1/2}\right)^{\lambda} A^{1/2}$$



- It constitutes a natural framework for a generalized theory of convexity, where the role of arithmetic mean is played by a midpoint pairing :
- f is called $\left(\bullet_1, \bullet_2\right)$-convex if
- For space of symmetric positive definite matrices, an affine function is given by : $$f\left(X \bullet_1 Y\right) \leq f\left(X\right) \bullet_2 f\left(Y\right)$$

- and is closely related to entropy : $f = \log \det$

$$Entropy = -\log \det\left(R\right) + cste$$

THALES

Bruhat-Tits Space & Semi-parallelogram Law

- **Bruhat-Tits Space :**
  - Space (Sym^{++},δ) is called Bruhat-Tits Space, where, according to δ metric, point at equal distance of 2 points is unique and given by geometric mean:

« Bruhat-Tits Space» (Metric Space)

$$\forall A, B \in Sym(n,R) \text{ positive definite}$$

$$\delta(A,B) = \left(\sum_{i=1}^{n} \log^2 \lambda_i \right) \quad \text{with} \quad \det\left(AB^{-1} - \lambda I\right) = 0$$

$$\delta(A,B) = 2\delta(A, A \circ B) = 2\delta(B, A \circ B)$$

$$\delta\left(C^T AC, C^T BC\right) = \delta(A,B) = \delta(A^{-1}, B^{-1})$$

- Bruhat-Tits Space is a complete metric space that verifies semi-parallelogram law :

$$\forall x_1, x_2 \in \mathrm{X} \quad \exists z \ \text{ such that :}$$

$$\delta(x_1,x_2)^2 + 4\delta(x,z)^2 \leq 2\delta(x,x_1)^2 + 2\delta(x,x_2)^2 \ \text{ for all } x \in X$$

- Bruhat-Tits Space are particular cases of **Cartan-Hadamard Manifold**

[1] F. Bruhat & J. Tits, « Groupes réductifs sur un corps local », IHES, n°41, pp.5-251, 1972

THALES

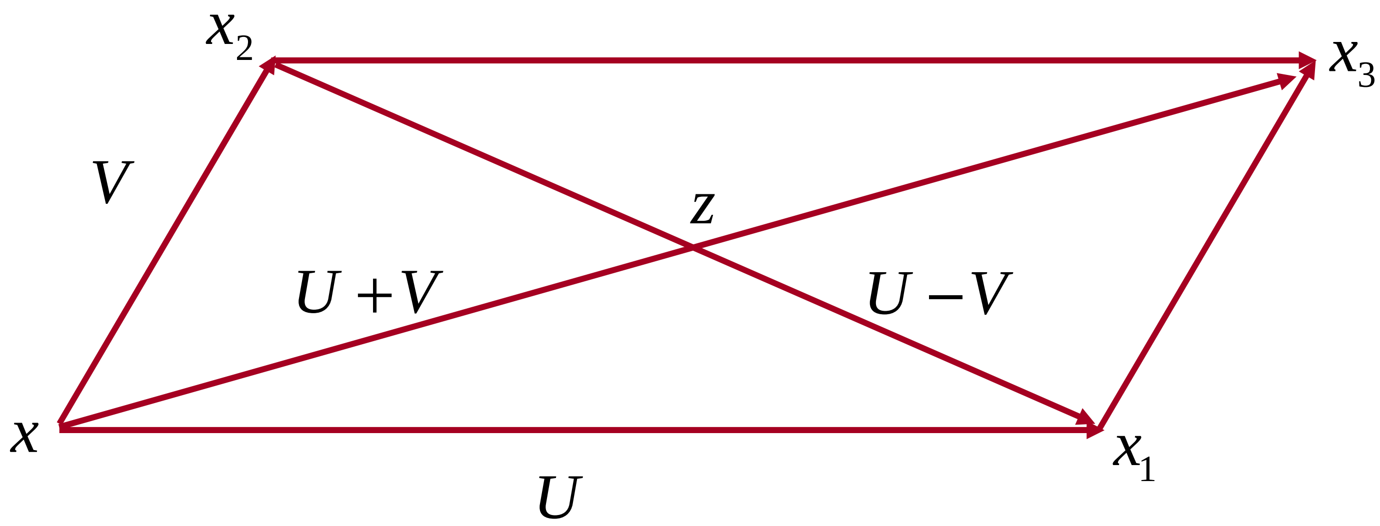
Bruhat-Tits Space & Semi-Parallelogram Law

Bruhat-Tits Space :

- Semi-parallelogram Law :

$$\forall x_1, x_2 \in X \quad \exists z \text{ such that :}$$

$$\delta(x_1,x_2)^2 + 4\delta(x,z)^2 \le 2\delta(x,x_1)^2 + 2\delta(x,x_2)^2 \quad \text{for all } x \in X$$



- Deduced from parallelogram law :

$$\|U-V\|^2 + \|U+V\|^2 = 2\|U\|^2 + 2\|V\|^2$$

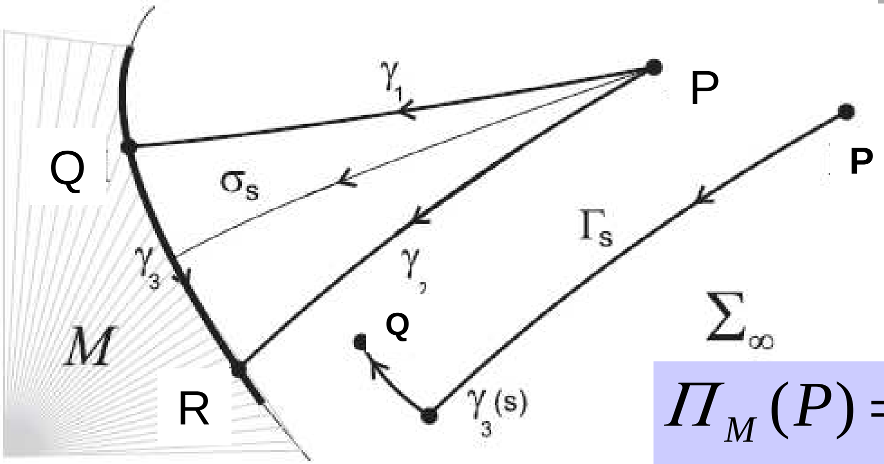
$$d_2(x_1,x_2)^2 + d_2(x,x_3)^2 = 2d_2(x,x_1)^2 + 2d_2(x,x_2)^2$$

$$2.d_2(x,z) = d_2(x,x_3) \Rightarrow d_2(x_1,x_2)^2 + 4d_2(x,z)^2 = 2d_2(x,x_1)^2 + 2d_2(x,x_2)^2$$

THALES

Geodesic Projection

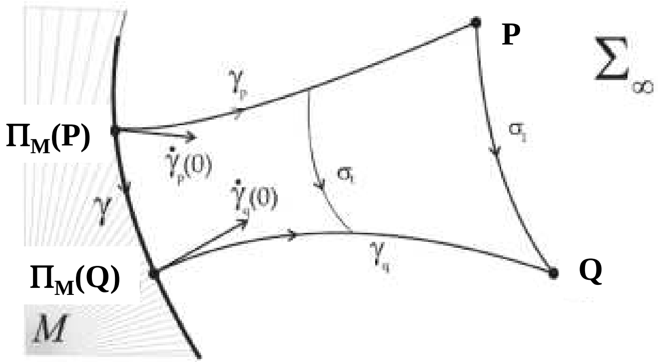
M : geodesically convex closed submanifold of Hadamard Manifold



$$\Pi_M(P) = \arg\min_{S \in M} dist(P,S)$$

- **Geodesic between matrix P and Geodesic between Q & R :**

$$\sigma_s(t) = \sigma(s,t) = P^{1/2}\left[P^{-1/2}Q^{1/2}\left(Q^{-1/2}RQ^{-1/2}\right)^s Q^{1/2}P^{-1/2}\right]^t P^{1/2}$$



- **The geodesic projection is a contraction :**

$$dist(\Pi_M(P), \Pi_M(Q)) \le dist(P,Q)$$

Kähler Potential & Bergman Kernel

- **We can find Bergman Kernel of Kähler geometry for Siegel Upper-half space by mean of Siegel Unit Disk :**

$$\Psi : SH_n \to SD_n$$
$$Z \mapsto (Z - iI_n)(Z + iI_n)^{-1}$$

$$SD_n = \{ Z \in Sym(n,C) / \|Z\|_2 < I \} \text{ with } \|Z\|_2 = Z\bar{Z}$$

$$\Phi = -\ln K_B(Z,Z)$$

With Bergman Kernel $K_B(Z,W) = \left(I - Z\bar{W}\right)$ on $\{ Z / I\text{-}Z\bar{Z} > 0 \}$

$$ds^2 = \frac{\partial^2 \Phi}{\partial Z \partial \bar{Z}} dZ d\bar{Z} = Tr\left(\left(I - Z\bar{Z}\right)^{-1} dZ \left(I - Z\bar{Z}\right)^{-1} d\bar{Z} \right)$$

- Analogy with Bergman Kernel for Poincaré unit disk :

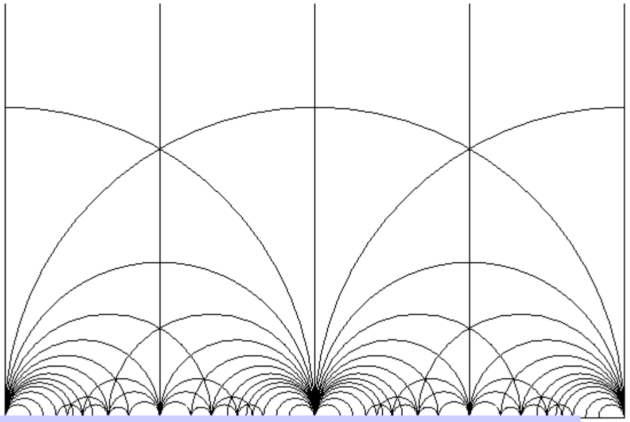
$$\Phi = -\ln K_B(z,z) \quad \text{with} \quad K_B(z,w) = 1 - z\bar{w} \quad \text{on} \{ z / |z|^2 < 1 \}$$

$$\Phi = -\ln\left(1 - |z|^2\right) \Rightarrow \frac{\partial \Phi}{\partial \bar{z}} = \frac{z}{1 - |z|^2} \Rightarrow g = \frac{\partial^2 \Phi}{\partial z \partial \bar{z}} = \frac{\left(1 - |z|^2\right) - z(-\bar{z})}{\left(1 - |z|^2\right)^2} = \frac{1}{\left(1 - |z|^2\right)^2}$$

$$ds^2 = g dz d\bar{z} = \frac{\partial^2 \Phi}{\partial z \partial \bar{z}} dz d\bar{z} = \frac{|dz|^2}{\left(1 - |z|^2\right)^2}$$

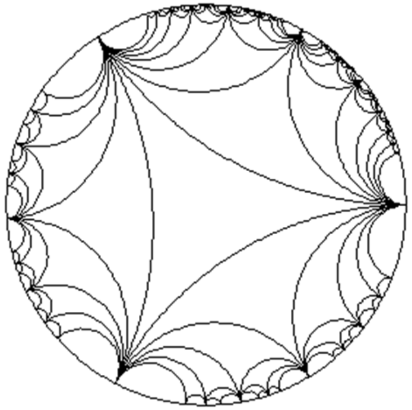
THALES

Upper-Half Planes & Disks of H. Poincaré and C.L. Siegel



Upper-Half Plane of H. Poincaré

$$w=\frac{z-i}{z+i}$$



Unit Disk of Poincaré

$$ds^2=\frac{dx^2+dy^2}{y^2}=\frac{|dz|^2}{y^2}$$

$$ds^2=y^{-1}dzy^{-1}dz^*$$

with $z=x+iy$ and $y>0$

$$ds^2=\frac{|dw|^2}{\left(1-|w|^2\right)^2}$$

$$ds^2=\left(1-ww^*\right)^{-1}dw\left(1-ww^*\right)^{-1}dw^*$$

Siegel Upper-Half Plane

$$W=(Z-iI)(Z+iI)^{-1}$$

Unit Siegel Disk

$$ds^2=Tr\left(Y^{-1}dZY^{-1}d\overline{Z}\right)$$

with $Z=X+iY$

$X\in Herm(n,C)$ and $Y\in HPD(n,C)$

$$ds^2=Tr\left[\left(I-WW^+\right)^{-1}dW\left(I-W^+W\right)^{-1}dW^+\right]$$

THALES

- **Koszul Forms**
  - 1st Koszul Form : $\alpha = -\frac{1}{4} d\Psi(X)$

$$\Psi(X) = Tr_{g/b}\left[ad(JX) - Jad(X)\right] \quad \forall X \in g$$

  - 2nd Koszul Form: $\beta = D\alpha$

- **Application for Poincaré Upper-Hal Plane:**

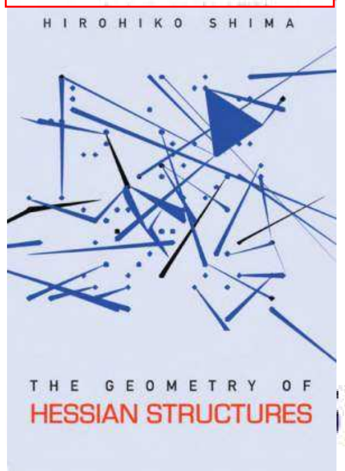
$$V = \{z = x + iy \,/\, y > 0\}$$

  - With $X = y\frac{d}{dx}$ and $Y = y\frac{d}{dy} \Rightarrow \begin{cases} ad(Y).Z = [Y, Z] \\ [X, Y] = -Y \\ JX = Y \end{cases}$

  - We can deduce that $\begin{cases} Tr\left[ad(JX) - Jad(X)\right] = 2 \\ Tr\left[ad(JY) - Jad(Y)\right] = 0 \end{cases}$

$$\Psi(X) = 2\frac{dx}{y} \Rightarrow \alpha = -\frac{1}{4} d\Psi = -\frac{1}{2}\frac{dx \wedge dy}{y^2}$$

$$\Rightarrow ds^2 = \frac{dx^2 + dy^2}{2y^2} \qquad \Omega = \frac{1}{y^2} dx \wedge dy$$



- **Koszul form for Siegel Upper-Half Plane:**

$$V=\{Z=X+iY\ /\ Y>0\}$$

  - Symplectic Group :

$$\begin{cases} SZ=(AZ+B)D^{-1} \\ A^T D=I, B^T D=D^T B \end{cases} \text{ with } S=\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \text{ and } J=\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \text{ with } \begin{cases} b=b^T \\ d=-a^T \end{cases} \text{ and base } \alpha_{ij}=\begin{pmatrix} e_{ij} & 0 \\ 0 & -e_{ji} \end{pmatrix}, \beta_{ij}=\begin{pmatrix} 0 & e_{ij}+e_{ji} \\ 0 & 0 \end{pmatrix}$$

  - Associated Lie Algebra:

$$\begin{cases} \Psi(\alpha_{ij})=0 \\ \Psi(\beta_{ij})=\delta_{ij}(3p+1) \end{cases} \Rightarrow \begin{cases} \Psi(dX+idY)=\dfrac{3p+1}{2}Tr\left(Y^{-1}dX\right) \\ \Omega=-\dfrac{1}{4}d\Psi=\dfrac{3P+1}{8}Tr\left(Y^{-1}dZ\wedge Y^{-1}d\overline{Z}\right) \\ ds^2=\dfrac{(3p+1)}{8}Tr\left(Y^{-1}dZY^{-1}d\overline{Z}\right) \end{cases}$$

www.thalesgroup.com

Geometry of Optimal Transport and 4th « Forgotten » distance of Maurice Fréchet

THALES

Optimal Transport Theory: Wasserstein distance

Wasserstein Distance

- Classical form of Wasserstein distance in Optimal transport

$$W_n(\mu,\nu)=\left(\underset{\pi\in\Pi(\mu,\nu)}{Inf}\int_{\wp} d(x,y)^n d\pi(x,y)\right)^{1/n}$$

$$W_n(\mu,\nu)= Inf\left\{\left(E\left[d(X,Y)^n\right]\right)^{1/n}, law(X)=\mu, law(Y)=\nu\right\}$$

Particular case of Wasserstein distance

- Case *n = 1* : Monge-Kantorovich-Rubinstein Distance

$$W_1(\mu,\nu)=d(P,Q)=\underset{X,Y}{Inf}\ E\left[\left|X-Y\right|\right]$$

- Case *n=2* : distance de Levy-Fréchet Distance

$$W_2(\mu,\nu)=d(P,Q)=\underset{X,Y}{Inf}\ E\left[\left|X-Y\right|^2\right]$$

THALES

Fréchet-Wasserstein Distance for Multivariate Gaussian LAw

- **Fréchet-Wasserstein Distance**

$$E\left[\|X-Y\|^2\right]=E\left[Tr\left((X-Y)^+(X-Y)\right)\right]$$
$$E\left[\|X-Y\|^2\right]=|E(X)-E(Y)|^2+Tr\left[R_X\right]+Tr\left[R_Y\right]-2Tr\left[R_{X,Y}\right]$$

- **Proof:**
  - Let $W=\begin{bmatrix} X \\ Y \end{bmatrix} \Rightarrow R_W=\begin{bmatrix} R_X & \Psi \\ \Psi^+ & R_Y \end{bmatrix}\geq 0$

$$\underset{R_X-\Psi R_Y^{-1}\Psi^+\geq 0}{Sup}\ Tr\left(\Psi+\Psi^+\right)\Rightarrow \Psi=R_X^{1/2}\left(R_X^{1/2}.R_Y.R_X^{1/2}\right)^{1/2}R_X^{-1/2}$$

  - Solution is given by:

$$Tr\left(\Psi+\Psi^+\right)=2.Tr\left[\left(R_X^{1/2}.R_Y.R_X^{1/2}\right)^{1/2}\right]$$

$$Y=R_X^{1/2}\left(R_X^{1/2}.R_Y.R_X^{1/2}\right)^{-1/2}R_X^{1/2}.X$$

THALES

Optimal Transport: Fréchet-Wasserstein Distance

Wasserstein Distance for Multivariate Gaussian Law with non-zero mean

$$d^2((m_X,R_X),(m_Y,R_Y)) = |m_X - m_Y|^2 + Tr[R_X] + Tr[R_Y] - 2.Tr\left[\left(R_X^{1/2} R_Y R_X^{1/2}\right)^{1/2}\right]$$

$$Tr[R_X] + Tr[R_Y] - 2.Tr\left[\left(R_X^{1/2} R_Y R_X^{1/2}\right)^{1/2}\right] = Tr[MM^+] \text{ with } M = \left[R_X^{1/2}\left(R_X^{1/2} R_Y R_X^{1/2}\right)^{-1/2} R_X^{1/2} - I\right] R_Y^{1/2}$$

- **Geodesic:** $D_{X,Y} = R_X^{1/2}\left(R_X^{1/2} R_Y R_X^{1/2}\right)^{-1/2} R_X^{1/2} = R_X \circ R_Y^{-1}$
  - If we let: $\begin{cases} m_{(t)} = (1-t)m_Y + t.m_X \\ R_{(t)} = ((1-t)I_k + t.D_{X,Y})R_Y((1-t)I_k + t.D_{X,Y}) \end{cases}$

$$W_2\left[(m_{(s)},R_{(s)}),(m_{(t)},R_{(t)})\right] \leq (t-s).W_2\left[(m_{(0)},R_{(0)}),(m_{(1)},R_{(1)})\right]$$

- **Optimal Transport :**
  - $(I_k,\nabla\psi)_\# N(m_Y,R_Y)$ transport from $N(m_Y,R_Y)$ to $N(m_X,R_X)$

$$\psi(v) = \frac{1}{2}(v-m_Y)^+ D_{X,Y}(v-m_Y) + (m_X - v)$$

$$x = \nabla\psi(y) = D_{X,Y}(y-m_Y) + m_X$$

$$(x-m_x)^+ R_X^{-1}(x-m_X) = (y-m_Y)^+ R_Y^{-1}(y-m_Y)$$

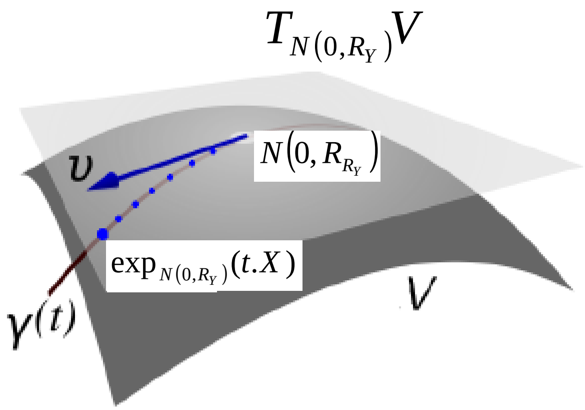
THALES

Optimal Transport: Fréchet-Wasserstein Distance

Characterization of this space (cas m=0)

- **Wasserstein Metric :**

$$g_{N(0,R_Y)}(X,Y) = Tr\left(X.R_Y.Y\right)$$



- **Tangent Space & Exponential Map:**

$$\exp_{N(0,R_Y)}(t.X) = N\left(0, R_{R_Y,(t)}\right)$$

$$R_{R_Y,(t)} = ((1-t)I_k + t.X)R_Y((1-t)I_k + t.X)$$

Properties of this space

- **Alexandrov Space**

$$d(\alpha,\gamma(t))^2 \geq (1-t).d(\alpha,\gamma(0))^2 + t.d(\alpha,\gamma(t))^2 - t(1-t).d(\gamma(0),\gamma(1))^2$$

- **Space of Positive sectionnal curvature (geodesically convexe and simply connected). Sectional Curvature is given by:**

$$K_{N(V)}(X,Y) = (3/4).Tr\left(([Y,X-S])V([Y,X]-S)\right)^T$$

Where matrix S is a symetric matrix such that:

$([Y,X]-S)V$ **is anti-symmetric.**

THALES

Wasserstein Barycenter of multivariate gaussian laws (Cas m=0)

- **Wasserstein Barycenter**

$$\underset{\mu}{Inf}\sum_{k=1}^{N}W_2^2(\mu,\nu_k) \quad \text{with} \quad W_2(\mu,\nu)=\underset{X,Y}{Inf}\ E\left[\|X-Y|^2\right]$$

- **Solution of Wasserstein Barycenter for N Mulitvariate gaussian laws of zero mean :** $\{N(0,R_k)\}_{k=1}^{N}$
  - Constrained for the barycenter:

$$\sum_{k=1}^{N}\left(R^{1/2}R_kR^{1/2}\right)^{1/2}=R$$

- **Iterative solution (paris-dauphine university)**
  - Fixed point (convergence for d=2, d>3 conditional convergence to initiation)

$$K^{(n+1)}=\left(\sum_{k=1}^{N}\left(K^{(n)}K_k^2K^{(n)}\right)^{1/2}\right)^{1/2} \quad \text{with} \quad K_i=R_i^{1/2}$$

THALES

Optimal Transport: Fréchet-Wasserstein Distance

Levy-Fréchet Distance

- In 1957, Maurice René Fréchet introduced this distance, based on a previous Paul Levy's paper from 1950

ACADÉMIE DES SCIENCES

SÉANCE DU LUNDI 4 FÉVRIER 1957.

CALCUL DES PROBABILITÉS. — *Sur la distance de deux lois de probabilité.* Note de M. MAURICE FRÉCHET.

Une formule explicite et simple est donnée pour représenter la distance de deux lois de probabilités quand on utilise la première des trois définitions de cette distance proposées par Paul Lévy. Une quatrième définition est proposée.

Nous prendrons ici, pour cette distance, l'écart quadratique moyen de X et de Y. En appelant $F(x)$, $G(y)$, $H(x, y)$ les fonctions de répartition respectives de X, de Y et du couple (X, Y), cet écart quadratique moyen D_H a pour carré

$$D_H^2 = \mathfrak{M}_H (X - Y)^2 = \iint_P (x-y)^2 \, d_x d_y H(x, y),$$

$$d^2(F,G) = \underset{X,Y}{Inf}\ E\left[|X-Y|^2\right] = \iint (x-y)^2 d_x d_y H(x,y)$$

THALES

Levy-Fréchet Distance

- **1957 Fréchet paper in CRAS :**

Paul Lévy a proposé (1) trois définitions de la distance de deux lois de probabilité L, L′.

Nous examinerons ici la première, qui est la plus intuitive et qui, contrairement à ce que l'on aurait pu attendre, conduit à des formules très simples.

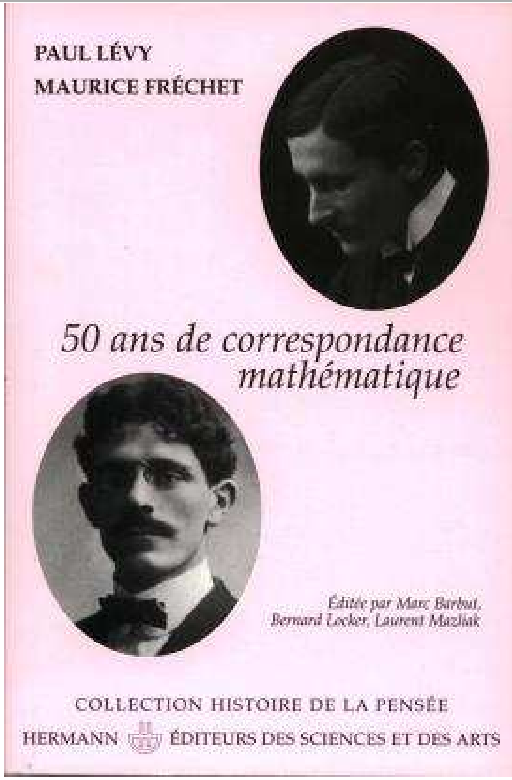
Selon cette première définition, la distance (L, L′) de ces deux lois est la borne inférieure de la « distance globale »

$$([X], [Y])$$

de deux nombres aléatoires X, Y qui ont respectivement L et L′ comme lois de probabilités individuelles, quand la corrélation entre X et Y varie.

Il est clair que la distance (L, L′) va dépendre de la définition adoptée pour la distance globale de X et de Y.

C.R., 1957, 1er *Semestre*. (T. 244, N° 6.) 44



- **Letter from Paul Levy to Maurice Fréchet (2 April 1958)**
  - … J'ai ainsi pu apprécier ce que vous aviez fait, en prenant comme point de départ de votre mémoire ce que vous appelez ma première définition de la distance de deux lois de probabilité (en fait ce n'était pas la première). Vous l'avez d'ailleurs généralisée, en ce sens que je ne l'avais associée qu'à une de vos définitions de deux variables aléatoires. Et j'ai beaucoup admiré comment avec votre quatrième définition, vous arrivez à faire quelque chose de maniable d'une idée qui pour moi était surtout théorique, vu la difficulté de déterminer le minimum de la distance de deux variables aléatoires ayant les répartitions marginales données.

4th « Forgotten » distance of Fréchet : Frechet Extreme Copulas

4th distance of Fréchet

$$d^2(F,G) = \iint (x-y)^2 \, d_x d_y H_1(x,y)$$

with $H_1(x,y) = Min[F(x), G(y)]$

- Extreme Copulas of Fréchet

Nous pouvons, en effet, considérer $H(x, y)$ comme définissant un « tableau de corrélation » dont les « marges » sont définies par les fonctions $F(x)$, $G(y)$. Or nous avons montré ([2]) que l'ensemble des fonctions $H(x, y)$ est identique à l'ensemble des fonctions de répartition dont les valeurs sont comprises entre deux d'entres elles, à savoir

$$(4) \qquad \begin{cases} H_0(x, y) = Max[F(x) + G(y) - 1, 0], \\ H_1(x, y) = Min[F(x), G(y)]. \end{cases}$$

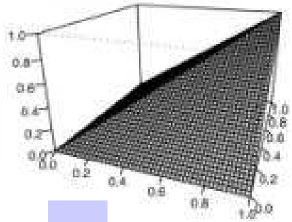
Poursuivant cette étude (d'ailleurs dans un autre but), Salvemini avait conjecturé que $\mathfrak{M}_H(X - Y)^2$ atteignait sa borne inférieure pour $H \equiv H_1$. Bass a énoncé ([3]) le résultat correspondant pour r_H dans le cas où X et Y sont bornés (et m'en a communiqué la démonstration). Un peu plus tard, Dall'Aglio ([4]) a validé la conjecture de Salvemini dans un cas plus général encore.

$$H_1(x,y) \le H(x,y) \le H_0(x,y)$$

with

$$H_0(x,y) = Max[F(x) + G(y) - 1, 0]$$

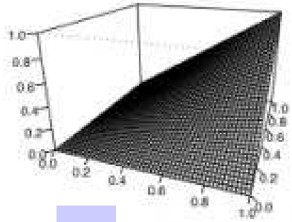
$$H_1(x,y) = Min[F(x), G(y)]$$



$H_0(x,y)$

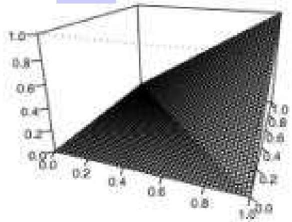
$\geq$

Extreme Fréchet Copulas



$H(x,y)$

$\geq$



$H_1(x,y)$

- 4th distance of Fréchet

Pour esquiver ces deux difficultés, nous allons proposer une quatrième définition. Si celle-ci les supprime, en effet, il faut reconnaître qu'elle est moins intuitive que celle de Lévy.

Nous poserons, *a priori*, sans explication

$$(L, L') = ([X], [Y])_{H_1}.$$

On peut en donner deux justifications. D'une part, elle coïncide avec celle de Lévy, au moins dans le cas, examiné plus haut, où la distance globale de X et Y est égale à leur écart quadratique moyen. D'autre part, on peut prouver que cette valeur de (L, L') vérifie bien même dans le cas général les trois conditions imposées à la notion de distance.

4th distance of Fréchet & Romane School of statistic (Pompilj, Leti, Rizzi, Dall'Aglio,Salvemini, Gini…)

- **G. Dall'Aglio (Université La Sapienza à Rome), « *Fréchet Classes: the Beginnings* », Advanced in Probability Distributions with Given Marginals Beyond the Copulas, Rome 1991, Kluwer**
  - With help of Giuseppe Pompilj, G. Dall'Aglio rmet Maurice Fréchet, 80 years old, in 1956 in Roma, during visit to l'Istituto di Calcolo delle Probabilita
  - Dall'Aglio met a 2nd time Fréchet in Paris and is in contact with Paul Levy
  - Dall'Aglio write : « ***Levy also showed some interest in distributions with given marginals. In a note in 1960 he refers to a formula by Poincaré for utilization in n-dimensional distributions with givens marginals, without developping the idea*** »
    - Paul Levy, « Sur les conditions de compatibilité des données marginales relatives aux lois de probabilité » CRAS, t. 250 pp.2507-2509, 1960, note présentée par M. Jacques Hadamard

SÉANCE DU 4 AVRIL 1960. 2507

CALCUL DES PROBABILITÉS. — *Sur les conditions de compatibilité des données marginales relatives aux lois de probabilité*. Note de M. **Paul Lévy**, présentée par M. Jacques Hadamard.

L'auteur se propose d'attirer l'attention sur le fait qu'une formule de Poincaré résout le problème dans le cas de *n* variables aléatoires n'ayant chacune que deux valeurs possibles, lorsque toutes les lois marginales sont données. Il indique ensuite le principe d'une méthode permettant de ramener à ce cas le cas le plus général. Mais les calculs sont vite impraticables.

THALES

4th distance of Fréchet & Romane School of statistic (Pompilj, Leti, Rizzi, Dall'Aglio,Salvemini, Gini…)

- **G. Dall'Aglio (Université La Sapienza à Rome), « *Fréchet Classes: the Beginnings* », Advanced in Probability Distributions with Given Marginals Beyond the Copulas, Rome 1991, Kluwer**
  - By integration by parts:

$$E\left[\left|X-Y\right|\right]=\int_R \left(F(z)+G(z)-H(z,z)\right)dz$$

Minimlizing Repartition function :

$$H^*(x,y)=\begin{cases} F(x)-\max\left\{\underset{x\le z\le y}{Inf}\left(F(z)-G(z)\right),0\right\} & \text{if } x\le y \\ G(y)-\max\left\{\underset{y\le z\le x}{Inf}\left(G(z)-F(z)\right),0\right\} & \text{if } x\ge y \end{cases}$$

For $\alpha>1$:

$$E\left[\left|X-Y\right|^{\alpha}\right]=\alpha(\alpha-1)\int_{u>v}\left(G(v)-H(u,v)\right)(u-v)^{\alpha-2}dudv+$$
$$\alpha(\alpha-1)\int_{u<v}\left(F(u)-F(u,v)\right)(v-u)^{\alpha-2}dudv$$

Minimizing Repartition Function :

$$H_1(x,y)=Min\left\{F(x),G(y)\right\}$$

  - $H^*(x,y)$ assigns to all sub-sets of diagonal y=x, maximum of probability compliant with marginals: S. Bertino, « Su di una sottoclasse della classe di Fréchet », Statistica 25, pp. 511-542, 1968
  - Gini Work : in 1914, Gini introduced « linear dissimilarity parameter » (solution in discret case for α = 1, 2)
  - Tommaso Salvemini (+ Leti) Work : construction of « tabella di cograduazione » and « di contrograduazione » (equivalent to Fréchet Extreme Repartition Functions)
  - Generalization to multivariate by Rizzi in 1957, Dall'Aglio (1960)

THALES

Comparison Information Geometry versus Optimal Transport

Complex Gaussian Circular Law of zero mean: $p(Z/R)=\frac{1}{\pi^n \det(R)}e^{-Tr(\hat{R}R^{-1})}$ with $\begin{cases}\hat{R}=ZZ^+\\ E[\hat{R}]=R\end{cases}$	
Information Geometry	**Optimal Transport Theory**
Distance between random variables: $ds^2=d\theta^+ I(\theta)d\theta=Tr\left((R^{-1}dR)^2\right)=\lVert R^{-1/2}dRR^{-1/2}\rVert_F^2$ $I(\theta)=\left[g_{i\bar{j}}\right]_{i,j},\ g_{i\bar{j}}=-E\left[\frac{\partial^2 \log p(X/\theta)}{\partial\theta_i\partial\theta_j^*}\right]$	**Distance between random variables:** $d^2(P,Q)=\underset{X,Y}{Inf}\ E\left[\lvert X-Y\rvert^2\right]$
Metric: $g_P(R_X,R_Y)=Tr\left(P^{-1}.R_X.P^{-1}.R_Y\right)$	**Metric:** $g_P(R_X,R_Y)=Tr\left(R_X.P.R_Y\right)$
Tangent space and Exponential map: $\exp_X\left(v_X^Y,t\right)=X^{1/2}e^{t\log\left(X^{-1/2}YX^{-1/2}\right)}X^{1/2}$ $\begin{cases}v_X^Y=grad_X^Y(V)=-\exp_X^{-1}(V)\\ v_X^Y=-X^{1/2}\log\left(X^{-1/2}YX^{-1/2}\right)X^{1/2}\\ \exp_X\left(v_X^Y,t\right)=X^{1/2}e^{-t\left(X^{-1/2}.v_X^Y.X^{-1/2}\right)}X^{1/2}\end{cases}$	**Tangent space and Exponential map:** $\exp_{N(0,R_Y)}(t.X)=N\left(0,R_{R_Y,(t)}\right)$ $R_{R_Y,(t)}=\left((1-t)I_k+t.X\right)R_Y\left((1-t)I_k+t.X\right)$
Distance between covariance matrices: $d^2(R_X,R_Y)=\lVert\log\left(R_X^{-1/2}.R_Y.R_X^{-1/2}\right)\rVert_F^2$ $d^2(R_X,R_Y)=\sum_{k=1}^{n}\log^2(\lambda_k)$ $\det\left(R_X^{-1/2}.R_Y.R_X^{-1/2}-\lambda.I\right)=\det(R_Y-\lambda R_X)=0$	**Distance between covariance matrices:** $d^2(R_X,R_Y)=Tr[R_X]+Tr[R_Y]-2.Tr\left[\left(R_X^{1/2}R_YR_X^{1/2}\right)^{1/2}\right]$
Geodesic between covariance matrices: $\gamma(t)=R_X^{1/2}e^{t\log\left(R_X^{-1/2}R_YR_X^{-1/2}\right)}R_X^{1/2}$ $\gamma(t)=R_X^{1/2}\left(R_X^{-1/2}R_YR_X^{-1/2}\right)^t R_X^{1/2}$ $\gamma(0)=R_X\ ,\ \gamma(1)=R_Y$ and $\gamma(1/2)=R_X\circ R_Y$	**Geodesic between covariance matrices:** $\gamma(t)=\left((1-t)I_k+t.D_{X,Y}\right)R_Y\left((1-t)I_k+t.D_{X,Y}\right)$ with $D_{X,Y}=R_X^{1/2}\left(R_X^{1/2}R_YR_X^{1/2}\right)^{-1/2}R_X^{1/2}=R_X\circ R_Y^{-1}$ $d(\gamma(s),\gamma(t))\le(t-s).d(\gamma(0),\gamma(1))$

Cartan Symmetric Space: $G_{(R_X,R_Y)}R_Z=(R_X\circ R_Y)R_Z^{-1}(R_X\circ R_Y)$ with $R_X\circ R_Y=R_X^{1/2}\left(R_X^{-1/2}R_YR_X^{-1/2}\right)^{1/2}R_X^{1/2}$	**Space associated to Optimal transport:** $(x-m_x)^+R_X^{-1}(x-m_X)=\left[R_X^{-1/2}(x-m_X)\right]^+\left[R_X^{-1/2}(x-m_X)\right]$ $=(y-m_Y)^+R_Y^{-1}(y-m_Y)$ $x=\nabla\psi(y)=D_{X,Y}(y-m_Y)+m_X$ $\psi(v)=\frac{1}{2}(v-m_Y)^+D_{X,Y}(v-m_Y)+(m_X-v)$
Bruhat-Tits Space (semi-parallelogram inequality): $\forall x_1,x_2\ \ \exists z$ such that $\forall x\in X$ $d(x_1,x_2)^2+4d(x,z)^2\le 2d(x,x_1)^2+2d(x,x_2)^2$	**Alexandrov Space:** $d(\alpha,\gamma(t))^2\ge$ $(1-t).d(\alpha,\gamma(0))^2+t.d(\alpha,\gamma(t))^2-t(1-t).d(\gamma(0),\gamma(1))^2$
Cartan-Hadamard Space: Complete, simply connected with negative sectional curvature Manifold	**Wasserstein Space:** Non-negative sectional curvature Manifold
Sectional curvature: $K=-Tr\left((I-ZZ^+)^{-1}M(I-ZZ^+)^{-1}M^+\right)$ with $(I-ZZ^+)^{-1}=PP^+$ when $I-ZZ^+>0$ $K=-Tr(TT^+)<0$ where $T=P^+MP$	**Sectional curvature:** $K_{N(V)}(X,Y)=(3/4).Tr\left(([Y,X-S])V([Y,X]-S)\right)^T$
Barycenter of N covariances matrices: $\sum_{k=1}^{N}\log\left(R^{-1/2}R_kR^{-1/2}\right)=0$ $R_{(n+1)}=R_{(n)}^{1/2}e^{\varepsilon\left(\sum_{k=1}^{N}\log\left(R_{(n)}^{-1/2}R_kR_{(n)}^{-1/2}\right)\right)}R_{(n)}^{1/2}$	**Barycenter of N covariances matrices:** $\sum_{k=1}^{N}\left(R^{1/2}R_kR^{1/2}\right)^{1/2}=R$ $K_{(n+1)}=\left(\sum_{k=1}^{N}\left(K_{(n)}K_k^2K_{(n)}\right)^{1/2}\right)^{1/2}$ with $K_i=R_i^{1/2}$

For transport optimal:

A. Takatsu. On Wasserstein Geometry of the Space of Gaussian Measures, to appear in Osaka J. Math., 2011, http://arxiv.org/abs/0801.2250

www.thalesgroup.com

Karcher Flows for HPD matrices: Mean & Median Computation

THALES

Geometric Mean Extension for Symmetric Positive Definite Matrices :

- Cauchy series given by harmonic and arithmetic mean :

$$A \circ B = \lim_{n\to\infty} A_n = \lim_{n\to\infty} B_n \quad \text{with} \quad A_0 = A \;,\; B_0 = B$$

$$A_{n+1} = 2\left(A_n^{-1} + B_n^{-1}\right)^{-1} \quad \text{and} \quad B_{n+1} = \left(A_n + B_n\right)/2$$

- Solution of the following ODE :

$$\forall A, B \in Sym^{++}, A \circ B = \lim_{t\to\infty} X(t) \quad \text{with} \begin{cases} \dfrac{dX}{dt} = B - XA^{-1}X \\ X(0) \in Sym^{++} \end{cases}$$

- Solution to the inequality :

$$\forall A, B \in Sym^{++}, A \circ B \text{ higher matrix (Loewner) s.t. } \begin{pmatrix} A & X \\ X & B \end{pmatrix} > 0$$

- Solution of the equation :

$$F''(X)(A) = X^{-1}AX = B^{-1} \quad \text{with} \quad F(X) = -\log\det(X)$$

THALES

Geometric Mean by Ricatti Equation

- In scalar case, geometric mean of a,b>0 satisfies for x>=0 :

$$x=\sqrt{ab} \Leftrightarrow x^2=ab \Leftrightarrow xa^{-1}x=b$$

- The last « symmetrized » version is suitable for generalization to noncommutative settings : the geometric mean is the unique solution of $xa^{-1}x=b$, if the unique solution exists.
- We define a group equipped with the operation $x \bullet y = xy^{-1}x$:

$$xa^{-1}x=b \Leftrightarrow x \bullet a = b \Leftrightarrow x = a \circ b$$

- To solve $x \bullet e = a$, we need $xex = x^2 = e$ to have a unique solution. Every element must have a unique square root : $e \circ a = a^{1/2}$
- Extension for Matrix Geometric Mean :

$$XA^{-1}X = B \Rightarrow A^{-1/2}(XA^{-1/2}A^{-1/2}X)A^{-1/2} = A^{-1/2}BA^{-1/2}$$

$$\Rightarrow (A^{-1/2}XA^{-1/2})(A^{-1/2}XA^{-1/2}) = A^{-1/2}BA^{-1/2}$$

$$\Rightarrow A^{-1/2}XA^{-1/2} = \left(A^{-1/2}BA^{-1/2}\right)^{1/2}$$

$$\Rightarrow \boxed{X = A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^{1/2}A^{1/2} = A \circ B}$$

Key Enablers : Matrix Square Root

Iterative Method for Square Root Computation

- **Denman-Beabers Iteration :**

$$X_{k+1} = \begin{bmatrix} 0 & Y_{k+1} \\ Z_{k+1} & 0 \end{bmatrix} = \frac{1}{2}\left(\begin{bmatrix} 0 & Y_k \\ Z_k & 0 \end{bmatrix} + \begin{bmatrix} 0 & Z_k^{-1} \\ Y_k^{-1} & 0 \end{bmatrix} \right)$$

with $X_0 = \begin{bmatrix} 0 & A \\ I & 0 \end{bmatrix}$ **then** $\underset{k\to\infty}{Lim}\, X_k = \begin{bmatrix} 0 & A^{1/2} \\ A^{-1/2} & 0 \end{bmatrix}$

- **Schultz Iteration (without matrix inversion) :**

$$X_{k+1} = \begin{bmatrix} 0 & Y_{k+1} \\ Z_{k+1} & 0 \end{bmatrix} = \frac{1}{2} X_k (3I - X_k^2)$$

- **Björk Method by Schur Decomposition :**

$Q^+ AQ = T$ **with *T* Upper triangular matrix**

By recurrence, we compute : $U^2 = T \Rightarrow A^{1/2} = QUQ^+$

The arithmetic-geometric mean : C.F. Gauss result

- **Gauss (20 years old) & Lagrange independently proved that**
  - Arithmetic-geometric mean of a and b :

$$\forall a \geq b > 0 \quad , \; (a_0, b_0) = (a, b) \text{ and } \begin{cases} a_{n+1} = \dfrac{a_n + b_n}{2} \\ b_{n+1} = \sqrt{a_n b_n} \end{cases}$$

$$a_{n+1}^2 - b_{n+1}^2 = \left(\frac{a_n + b_n}{2}\right)^2 - a_n b_n = \left(\frac{a_n - b_n}{2}\right)^2 \geq 0$$

$$\Rightarrow \forall n \geq 0 \quad b = b_0 \leq \ldots \leq b_n \leq b_{n+1} \leq a_{n+1} \leq a_n \leq \ldots \leq a_0 = a$$

  - Is related to elliptic integral :

$$I(a,b) = \frac{\pi/2}{M(a,b)} \quad \text{with} \quad \begin{cases} M(a,b) = \lim\limits_{n \to \infty} a_n = \lim\limits_{n \to \infty} b_n \\ I(a,b) = \displaystyle\int_0^{\pi/2} \frac{dt}{\sqrt{a^2 \cos^2 t + b^2 \sin^2 t}} \end{cases}$$

  - Landen transformation gives also :

$$\begin{cases} I(a_1, b_1) = I\left(\dfrac{a+b}{2}, \sqrt{ab}\right) = I(a,b) \\ 2J(a_1, b_1) = J(a,b) + abI(a,b) \end{cases} \quad \text{with} \quad J(a,b) = \int_0^{\pi/2} \sqrt{a^2 \cos^2 t + b^2 \sin^2 t}\, dt$$

4.J(a,b) : Perimeter of Ellipse of half-axis a & b

$$\text{Landen transf.}: u = t + Arc\tan\left(\frac{b}{a}\tan t\right)$$

THALES

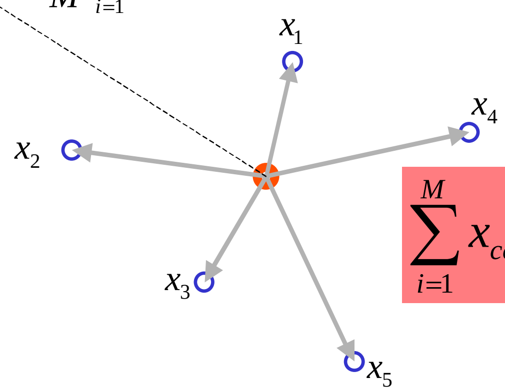
Center of Mass : Arithmetic Mean and Median in R^n

In R^n, the center of mass is defined for finite set of points $\{x_i\}_{i=1,\ldots,M}$

- Arithmetic mean: $x_{center} = \frac{1}{M}\sum_{i=1}^{M} x_i$
- This point minimizes the function of distances: $x_{center} = \arg\underset{x}{Min}\sum_{i=1}^{M} d^2(x, x_i)$
- The median (Fermat-Weber Point) minimizes : $x_{median} = \arg\underset{x}{Min}\sum_{i=1}^{M} d(x, x_i)$

$$m_{mean} = \underset{m}{Min}\, E\left[\left\|x - m\right\|^2\right]$$

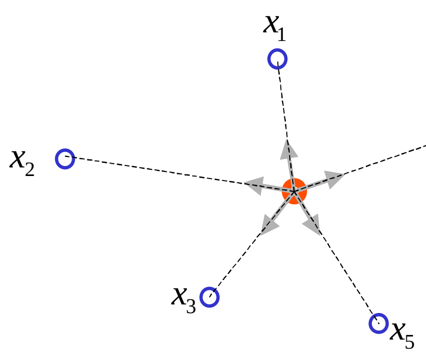
$$x_{center} = \frac{1}{M}\sum_{i=1}^{M} x_i$$



$$\sum_{i=1}^{M} \overrightarrow{x_{center} x_i} = 0$$

$$x_{center} = \arg\underset{x}{Min}\sum_{i=1}^{M} d^2(x, x_i)$$

$$m_{median} = \underset{m}{Min}\, E\left[\left\|x - m\right\|\right]$$



$$\sum_{i=1}^{M} \frac{\overrightarrow{x_{center} x_i}}{\left\|\overrightarrow{x_{center} x_i}\right\|} = 0$$

$$x_{median} = \arg\underset{x}{Min}\sum_{i=1}^{M} d(x, x_i)$$

Cartan Center of Mass and Karcher Flow

Cartan Center of Mass

- **Elie Cartan has proved that the following functional :**

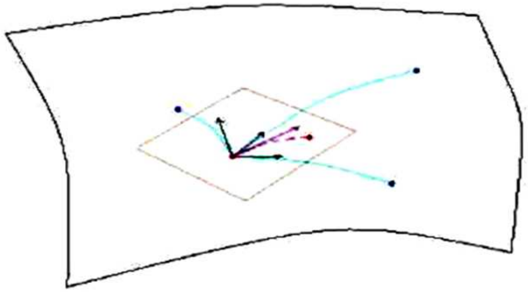
$$f : m \in M \mapsto \int_A d^2(m,a)da$$

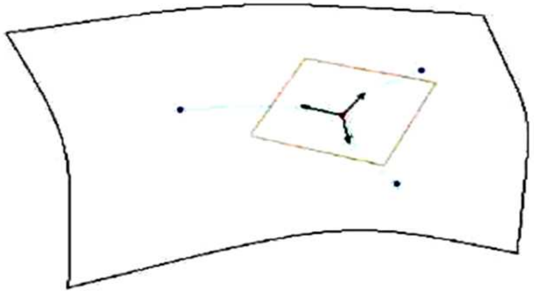
is strictly convexe and has only one minimum (center of mass of *A* for distribution *da*) for a manifold of negative curvature

Karcher Flow

- **Hermann Karcher has proved the convergence of the following flow to the Center of Mass :**

$$m_{n+1} = \gamma_n(t_n) = \exp_{m_n}\left(-t_n.\nabla f(m_n)\right) \quad \text{avec} \quad \dot{\gamma}_n(0) = -\nabla f(m_n)$$





$$\nabla f = -\int_A \exp_m^{-1}(a)da$$

THALES

probability on a manifold : Emery's exponential barycenter

- **Maurice René Fréchet, inventor of Cramer-Rao bound in 1939, has also introduced the entire concept of** Metric Spaces Geometry **and functional theory on this space (any normed vector space is a metric space by defining but not the contrary). On this base, Fréchet has then extended probability in abstract spaces.** $d(x,y)=\|y-x\|$

M. R. Fréchet, "Les éléments aléatoires de nature quelconque dans un espace distancié", Annales de l'Institut Henri Poincaré, n°10, pp.215-310, 1948

- **In this framework, expectation** $b=E[g(x)]$ **of an abstract probabilistic variable** $g(x)$ **where** x **lies on a manifold is introduced by Emery as an exponential barycenter :** $\int_M \exp_b^{-1}(g(x))P(dx)=0$

- **In Classical Euclidean space, we recover classical definition of Expectation *E[.]* :**

$$p,q\in R^n \Rightarrow \exp_p^{-1}(q)=q-p \Rightarrow E[g(x)]=\int_{R^n} g(x)P(dx)=\int_{R^n} g(x)p_X(x)dx$$

M. Emery & G. Mokobodzki, "Sur le barycentre d'une probabilité sur une variété", Séminaire de Proba. XXV, Lectures note in Math. 1485, pp.220-233, Springer, 1991

THALES

Mean & Median of N matrices HPD(n)

Mean of N Hermitian Positive Definite Matrices HPD(n)

- Solution given by Karcher Flow with Information Geometry metric

$$X_{moyenne} = \underset{X}{Arg\ Min} \sum_{k=1}^{N} d^2(X, B_k) = \sum_{k=1}^{N} \left\| \log\left(X^{-1/2}.B_k.X^{-1/2}\right) \right\|^2$$

$$X_{n+1} = X_n^{1/2} e^{\varepsilon \sum_{k=1}^{N} \log\left(X_n^{-1/2} B_k X_n^{-1/2}\right)} X_n^{1/2}$$

Median (Fermat-Weber Point) of N matrices HPD(n)

$$X_{mediane} = \underset{X}{Arg\ Min} \sum_{k=1}^{N} d(X, B_k) = \sum_{k=1}^{N} \left\| \log\left(X^{-1/2}.B_k.X^{-1/2}\right) \right\|$$

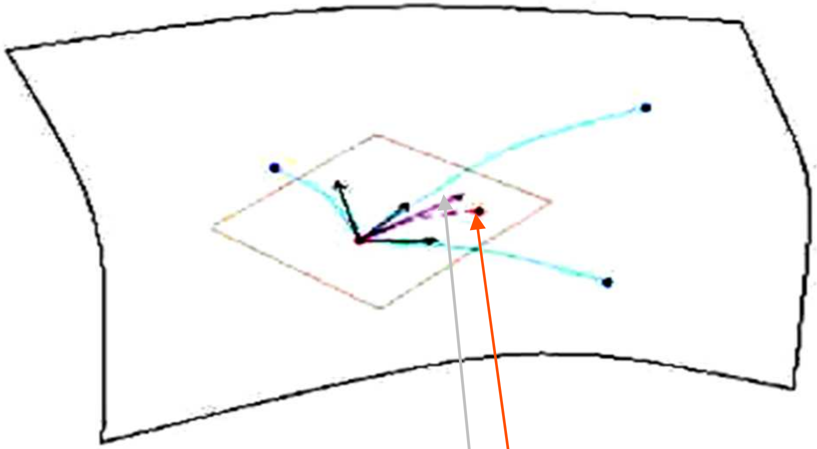
$$X_{n+1} = X_n^{1/2} e^{\varepsilon \sum_{k \in S_n} \frac{\log\left(X_n^{-1/2} B_k X_n^{-1/2}\right)}{\left\| \log\left(X_n^{-1/2} B_k X_n^{-1/2}\right) \right\|}} X_n^{1/2} \quad \text{with} \quad S_n = \left\{ k \,/\, X_n \neq B_k \right\}$$

- PhD Yang Le supervised by Marc Arnaudon (Univ. Poitiers/Thales)

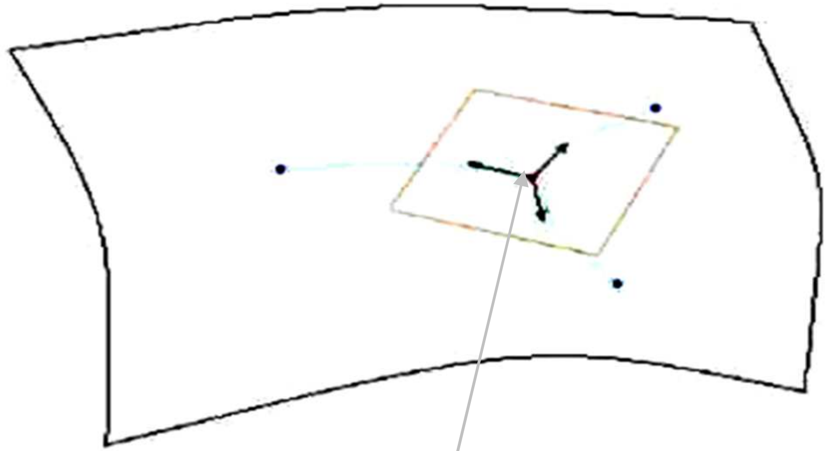
THALES

Median on a Manifold: Karcher-Cartan-Fréchet

- **Gradient flow on Surface/Manifold**



Gradient Flow :Pushed by Sum of Normalized Tangent vectors of Geodesics



Fréchet-Karcher Barycenter : Sum of Normalized Tangent vectors of Geodesics is equal to zero

$$h : m \in \mathrm{M} \mapsto \frac{1}{2}\int_A d(m,a)da \underset{Min}{\Rightarrow} \nabla h = -\int_A \frac{\exp_m^{-1}(a)}{\left\|\exp_m^{-1}(a)\right\|} da$$

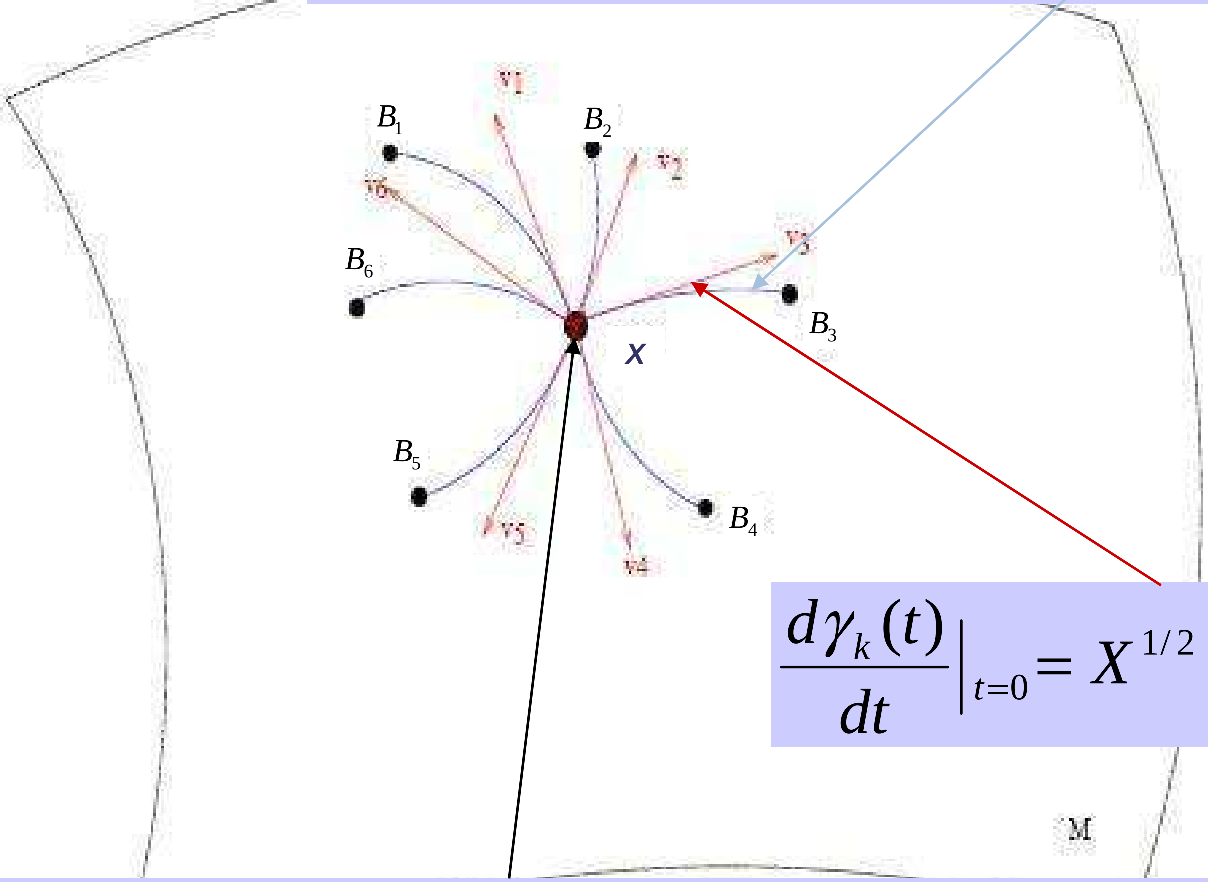
Sum of Normalized Tangent Vectors

Compute point on the surface In the direction of sum of Normalized Tangent Vector

$$m_{n+1} = \exp_{m_n}\left(t.\sum_{k=1}^{M} \frac{\exp_{m_n}^{-1}(x_k)}{\left\|\exp_{m_n}^{-1}(x_k)\right\|} \right)$$

Mean : Karcher Barycenter

$$\gamma_k(t) = X^{1/2}\left(X^{-1/2}B_k X^{-1/2}\right)^t X^{1/2} = X^{1/2}e^{t\log\left(X^{-1/2}B_k X^{-1/2}\right)}X^{1/2}$$



$$\left.\frac{d\gamma_k(t)}{dt}\right|_{t=0} = X^{1/2}\log\left(X^{-1/2}B_k X^{-1/2}\right)X^{1/2}$$

$$\sum_{k=1}^{N}\left.\frac{d\gamma_k(t)}{dt}\right|_{t=0} = X^{1/2}\left(\sum_{k=1}^{N}\log\left(X^{-1/2}B_k X^{-1/2}\right)\right)X^{1/2} = 0$$

Gradient Flow & Toeplitz structure

$$A_{t+1} = A_t^{1/2} e^{-\varepsilon \sum_{k=1}^{N} \log\left(A_t^{-1/2} B_k A_t^{-1/2}\right)} A_t^{1/2}$$

- **The following structure of a Matrix *M* is invariant by this flow :**

$$JMJ = M$$

- **where *J* is antidiagonal matrix. We can prove that :**

$$E = \left\{ M \in HPD_n(C) \,/\, JMJ = M \right\}$$

- **is a close submanifold of $HPD_n(C)$ (Hermitian Positive Definite Matrix) and sub-group of $GL_n(C)$**

- **For Siegel metric, this set is convex in a “Geodesic” meaning (all geodesics between two elements of *E* are in *E*). Then barycenter of N elements of *E* is in *E*.**

$$e^{JMJ} = e^{JMJ^{-1}} = Je^{M}J^{-1} = e^{M}$$

www.thalesgroup.com

Karcher Flows for Toeplitz HPD matrices (THPD): Trench/Verblunsky Theorem

THALES

Additional Constraint : Toeplitz Structure

- **Covariances matrices are** structured matrices **that verify the following constraints :**
  - Toeplitz Structure (for stationary signal) : $\forall n,\quad E\left[z_n z_{n-k}^*\right] = r_k$

$$R_n = \begin{bmatrix} r_0 & r_1^* & r_2^* & \cdots & r_{n-1}^* \\ r_1 & r_0 & r_1^* & \ddots & \vdots \\ r_2 & r_1 & \ddots & \ddots & r_2^* \\ \vdots & \ddots & \ddots & r_0 & r_1^* \\ r_{n-1} & \cdots & r_2 & r_1 & r_0 \end{bmatrix} = \left[\begin{array}{cccc|c} & & & & r_{n-1}^* \\ & R_{n-1} & & & \vdots \\ & & & & r_2^* \\ & & & & r_1^* \\ \hline r_{n-1} & \cdots & r_2 & r_1 & r_0 \end{array}\right]$$

  - Hermitian structure :

$$R_n^+ = R_n \quad \text{with} \quad + : \text{transposed} \quad \text{and} \quad \text{conjuguate}$$

  - Positive Definite Structure :

$$\forall Z \in C^n, Z^+ R_n Z > 0, \forall \lambda \text{ such that } \det\left(R_n - \lambda . I_n\right) = 0 \Rightarrow \lambda > 0$$

- **How to built a flow that preserves the Toeplitz structure ?**

If you remember first example on « Right Triangle », in this case Pythagore constraint is replaced by HPD & Toeplitz constraints

Complex Circular Multivariate Gaussian Model

Radar Signal Model :

- Complex Circular Multivariate Model :

$$p(Y_n / R_n) = (\pi)^{-n}.\left|R_n\right|^{-1}.e^{-Tr\left[\hat{R}_n.R_n^{-1}\right]}$$

$$\text{with} \quad \hat{R}_n = (Y_n - m_n).(Y_n - m_n)^+ \quad \text{et} \quad E\left[\hat{R}_n\right] = R_n$$

- Radar Model :

$$m_n = 0 \quad \text{zero mean process}$$

$$R_n = E\left[Y_n Y_n^+\right] \quad \text{Toeplitz Hermitian Definite Positive}$$

$$Y_n = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \text{with} \quad y_k = a_k + ib_k = \rho_k e^{i\phi_k}$$

$$y_n = -\sum_{k=1}^{N} a_k^{(N)} y_{n-k} + b_n \quad \text{with } E\left[b_n b_{n-k}^*\right] = \delta_{k,0}\sigma^2 \text{ and } A_N = \left[a_1^{(N)} \ldots a_N^{(N)}\right]^T$$

- Close Link with Issai Schur's algorithm (1875-1941)

  - *[Alpay] D. Alpay, « Algorithme de Schur, espaces à noyau reproduisant et théorie des systèmes », Panoramas et synthèse, n°6, Société Mathématique de France, 1998*

Block matrix structure of CAR model

- **We can exploit Block structure of covariance matrix to compute Rao metric for CAR model :**
  - Rao metric for Complex Multivariate Gaussian model of zero mean :

$$ds^2 = Tr\left(\left(R^{-1}dR\right)^2\right) = \left\|R^{-1/2}dRR^{-1/2}\right\|^2 \text{ with } \left\|A\right\|^2 = \langle A, A\rangle \text{ et } \langle A, B\rangle = Tr\left(AB^T\right)$$

$$D^2(R_1, R_2) = \left\|\log\left(R_1^{-1/2}R_2R_1^{-1/2}\right)\right\|^2 = \sum_{k=1}^{n}\log^2(\lambda_k) \text{ et } \det\left(R_1^{-1/2}R_2R_1^{-1/2} - \lambda I\right) = 0$$

  - Block structure of covariance matrix for CAR model :

$$R_n^{-1} = \begin{bmatrix} \alpha_{n-1} & \alpha_{n-1}.A_{n-1}^+ \\ \alpha_{n-1}.A_{n-1} & R_{n-1}^{-1} + \alpha_{n-1}.A_{n-1}.A_{n-1}^+ \end{bmatrix}$$

$$R_n = \begin{bmatrix} \alpha_{n-1}^{-1} + A_{n-1}^+.R_{n-1}.A_{n-1} & -A_{n-1}^+.R_{n-1} \\ -R_{n-1}.A_{n-1} & R_{n-1} \end{bmatrix}$$

$$\text{with } \alpha_n^{-1} = \left[1 - |\mu_n|^2\right]\alpha_{n-1}^{-1} \text{ and } A_n = \begin{bmatrix} A_{n-1} \\ 0 \end{bmatrix} + \mu_n.\begin{bmatrix} A_{n-1}^{(-)} \\ 1 \end{bmatrix} \text{ where } V^{(-)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \iddots & 0 \\ 1 & 0 & 0 \end{bmatrix}.V^*$$

THALES

Block matrix structure & extended eigen-values

- **Matrix** $R_n^{(1)-1/2+}.R_n^{(2)}.R_n^{(1)-1/2}$ **or** $R_n^{(1)1/2+}.R_n^{(2)-1}.R_n^{(1)1/2}$ **that is used for Rao Metric computation has same block structure**

$$\Omega_n = R_n^{(1)1/2+}.R_n^{(2)-1}.R_n^{(1)1/2} = \begin{bmatrix} \beta_{n-1} & \beta_{n-1}.W_{n-1}^{+} \\ \beta_{n-1}.W_{n-1} & \Omega_{n-1} + \beta_{n-1}.W_{n-1}.W_{n-1}^{+} \end{bmatrix}$$

$$\text{with} \quad W_{n\text{-}1} = \sqrt{\alpha_{n-1}^{(1)}}.R_{n-1}^{(1)1/2+}.\left[A_{n-1}^{(2)} - A_{n-1}^{(1)}\right] \quad \text{and} \quad \beta_{\text{n-1}} = \frac{\alpha_{n-1}^{(2)}}{\alpha_{n-1}^{(1)}}$$

- This block structure to compute extended eigen-values :

$$\begin{cases} F^{(n)}\left(\lambda_k^{(n)}\right) = \lambda_k^{(n)} - \beta_{n-1} + \beta_{n-1}.\lambda_k^{(n)}.\sum_{i=1}^{n-1} \frac{\left|W_{n-1}^{+}.X_i^{(n-1)}\right|^2}{\left(\lambda_i^{(n-1)} - \lambda_k^{(n)}\right)} = 0 \\ \dfrac{X_k^{(n)}}{X_{k,1}^{(n)}} = \begin{bmatrix} 1 \\ -\lambda_k^{(n)}.U_{n-1}.\left(\Lambda_{n-1} - \lambda_k^{(n)}.I_{n-1}\right)^{-1}.U_{n-1}^{+}.W_{n-1} \end{bmatrix} \end{cases}$$

THALES

Block matrix structure & extended eigen-values

- We can deduce an algorithm that could be parallelized according to CAR model order for each extended eigen-values :

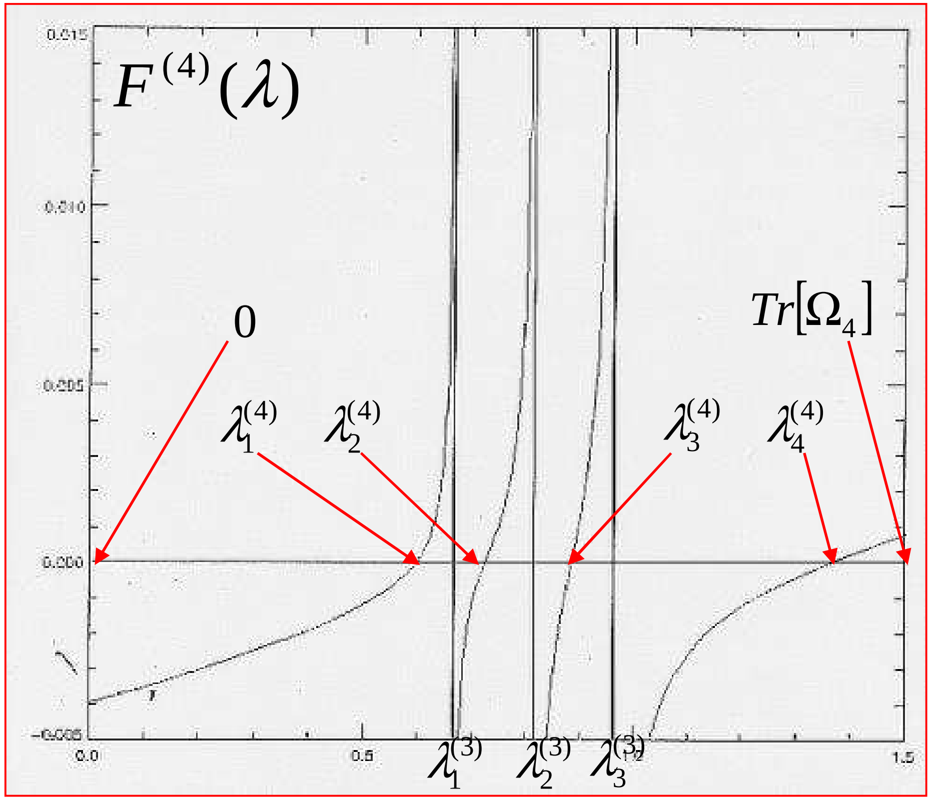
$$\begin{cases} F^{(n)}\left(\lambda_k^{(n)}\right)=\lambda_k^{(n)}-\beta_{n-1}+\beta_{n-1}.\lambda_k^{(n)}.\sum_{i=1}^{n-1}\frac{\left|W_{n-1}^{+}.X_i^{(n-1)}\right|^2}{\left(\lambda_i^{(n-1)}-\lambda_k^{(n)}\right)}=0 \\ \frac{X_k^{(n)}}{X_{k,1}^{(n)}}=\begin{bmatrix} 1 \\ -\lambda_k^{(n)}.U_{n-1}.\left(\Lambda_{n-1}-\lambda_k^{(n)}.I_{n-1}\right)^{-1}.U_{n-1}^{+}.W_{n-1}\end{bmatrix}\end{cases}$$

Renormalisation at each iteration :

$$\frac{X_k^{(n)}}{X_{k,1}^{(n)}}\rightarrow\left(1+\lambda_k^{(n)2}.\sum_{i=1}^{n-1}\frac{\left|W_{n-1}^{+}.X_i^{(n-1)}\right|^2}{\left(\lambda_i^{(n-1)}-\lambda_k^{(n)}\right)^2}\right)^{-1/2}.\frac{X_k^{(n)}}{X_{k,1}^{(n)}}$$

Interpretation in term of projection of CAR vector :

$$\begin{cases}\sum_{k=1}^{n}\left(\frac{\partial F^{(n)}\left(\lambda_k^{(n)}\right)}{\partial\eta}\right)^{-1}=1 \\ \begin{bmatrix}1\\ A_{n-1}\end{bmatrix}=\sum_{k=1}^{n}\left(\frac{\partial F^{(n)}\left(\lambda_k^{(n)}\right)}{\partial\eta}\right)^{-1}.\frac{X_k^{(n)}}{X_{k,1}^{(n)}}\end{cases}$$



$$0<\lambda_n^{(n)}<\lambda_{n-1}^{(n-1)}<\lambda_{n-1}^{(n)}<\ldots<\lambda_2^{(n)}<\lambda_1^{(n-1)}<\lambda_1^{(n)}<Tr[\Omega_n]$$

stricly increasing curve on each interval :

$$\frac{\partial F^{(n)}(\eta)}{\partial\eta}=\frac{\beta_{n-1}}{\eta_k^{(n)}.\left|X_{k,1}^{(n)}\right|^2}=1+\beta_{n-1}.\sum_{k=1}^{n-1}\frac{\eta_k^{(n-1)}.\left|W_{n-1}^{+}.X_k^{(n-1)}\right|^2}{\left(\eta_k^{(n-1)}-\eta\right)^2}>1$$

avec $Tr[\Omega_n]=Tr[\Omega_{n-1}]+\beta_{n-1}.\begin{bmatrix}1\\ W_{n-1}\end{bmatrix}^{+}\begin{bmatrix}1\\ W_{n-1}\end{bmatrix}$

THALES

Recursive Expression for Rao Metric

- **If we use block structure matrix, we can compute Rao metric at model order n from model order at n-1**
  - Recursive Expression of Rao Metric:

$$ds_n^2 = Tr\left[\left(R_n.dR_n^{-1}\right)^2\right] = ds_{n-1}^2 + \left(\frac{d\alpha_{n-1}}{\alpha_{n-1}}\right)^2 + \alpha_{n-1}.dA_{n-1}^+.R_{n-1}.dA_{n-1}$$

  - We can observe that :

$$\left(\frac{d\alpha_{n-1}}{\alpha_{n-1}}\right)^2 + \alpha_{n-1}.dA_{n-1}^+.R_{n-1}.dA_{n-1} = d\begin{bmatrix}\log\alpha_{n-1}\\ A_{n-1}\end{bmatrix}^+ . \begin{bmatrix}1 & 0\\ 0 & \alpha_{n-1}R_{n-1}\end{bmatrix} d\begin{bmatrix}\log\alpha_{n-1}\\ A_{n-1}\end{bmatrix}$$

  - This could be interpreted with Information Geometry & Fisher matrix

$$\begin{cases} R_{\hat{A}_n}^{-1} = I\left[A_{n-1}\right] = \alpha_{n-1}R_{n-1} \\ \log\alpha_{n-1} \perp A_{n-1} \end{cases} \Rightarrow Z = \begin{bmatrix}\log\alpha_{n-1}\\ A_{n-1}\end{bmatrix}\begin{bmatrix}\log\alpha_{n-1}\\ A_{n-1}\end{bmatrix}^+ + i\begin{bmatrix}1 & 0\\ 0 & \left(\alpha_{n-1}R_{n-1}\right)^{-1}\end{bmatrix}$$

THALES

Non-Symmetric Square Root of Siegel Group

- **If we consider Cholesky decomposition of covariance matrix :**
  - Cholesky decomposition (Goldberg inversion algorithm) :

$$\Omega_n = (\alpha_n . R_n)^{-1} = W_n . W_n^{+} = \left(1 - |\mu_n|^2\right) . \begin{bmatrix} 1 & A_{n-1}^{+} \\ A_{n-1} & \Omega_{n-1} + A_{n-1} . A_{n-1}^{+} \end{bmatrix}$$

$$\text{with } W_n = \sqrt{1 - |\mu_n|^2} \begin{bmatrix} 1 & 0 \\ A_{n-1} & \Omega_{n-1}^{1/2} \end{bmatrix} \text{ and } \Omega_{n-1} = \Omega_{n-1}^{1/2} . \Omega_{n-1}^{1/2+}$$

  - All distribution of n-dimensionnal variable is associated with Affine Group. It is the element such that its action on vector $Z \sim N_n(0, I_n)$

Is transformed to random vector : $X \sim N_n(A_n, \Omega_n)$

$$\begin{bmatrix} 1 & 0 \\ A_{n-1} & \Omega_{n-1}^{1/2} \end{bmatrix} . \begin{bmatrix} 1 \\ Z \end{bmatrix} = \begin{bmatrix} 1 \\ \Omega_{n-1}^{1/2} . Z + A_{n-1} \end{bmatrix} = \begin{bmatrix} 1 \\ X \end{bmatrix}$$

  - This representation of Affine Group elements could be considered as non symmetric square root of Siegel Group element :

$$\begin{bmatrix} 1 & A_{n-1}^{+} \\ A_{n-1} & \Omega_{n-1} + A_{n-1} . A_{n-1}^{+} \end{bmatrix}$$

Recursive Expression of Kullback Divergence

- **We consider expression of Kullback Divergence for Multivariante Gaussian Law of zero mean :**
  - Kullback divergence is given by : $K\left(R_n^{(1)}, R_n^{(2)}\right) = \frac{1}{2}\left[Tr[\Sigma_n] - \ln|\Sigma_n| - n\right]$

$$\Sigma_n = R_n^{(1)1/2+}.R_n^{(2)-1}.R_n^{(1)1/2} = \begin{bmatrix} \beta_{n-1} & \beta_{n-1}.V_{n-1}^{+} \\ \beta_{n-1}.V_{n-1} & \Sigma_{n-1} + \beta_{n-1}.V_{n-1}.V_{n-1}^{+} \end{bmatrix}$$

$$\text{where } \mathrm{V}_{\text{n-1}} = F_{n-1}^{(1)+}.\left[A_{n-1}^{(2)} - A_{n-1}^{(1)}\right] \text{ and } \beta_{\text{n-1}} = \frac{\alpha_{n-1}^{(2)}}{\alpha_{n-1}^{(1)}} \text{ with } \alpha_n^{(1)}.R_n^{(1)} = F_n^{(1)}.F_n^{(1)+}$$

$$F_n^{(1)} = \frac{1}{\sqrt{1-\left|\mu_n^{(1)}\right|^2}} \begin{bmatrix} 1 & -A_{n-1}^{(1)+} \\ 0 & I_{n-1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & F_{n-1}^{(1)} \end{bmatrix}$$

  - If we consider recursive relation :

$$|\Sigma_n| = \beta_{n-1}.|\Sigma_{n-1}| \qquad Tr[\Sigma_n] = Tr[\Sigma_{n-1}] + \beta_{n-1}.\left[1 + V_{n-1}^{+}.V_{n-1}\right]$$

  - Kullback Divergence is given recursively by :

$$K\left(R_n^{(1)}, R_n^{(2)}\right) = K\left(R_{n-1}^{(1)}, R_{n-1}^{(2)}\right) + \frac{1}{2}\left[\left(\beta_{n-1} - 1\right) + \beta_{n-1}.V_{n-1}^{+}.V_{n-1} - \ln\left(\beta_{n-1}\right)\right]$$

THALES

Trench/Verblunsky Theorem & Partial Iwasawa Parameterization

- **All Toeplitz Hermitian Positive Definite Matrix can be parameterized by Reflection/Verblunsky Coefficients:**
  - Block structure of covariance matrix & Verblunsky Parameterization:

$$R_n^{-1} = \begin{bmatrix} \alpha_{n-1} & \alpha_{n-1}.A_{n-1}^{+} \\ \alpha_{n-1}.A_{n-1} & R_{n-1}^{-1} + \alpha_{n-1}.A_{n-1}.A_{n-1}^{+} \end{bmatrix} \quad R_n = \begin{bmatrix} \alpha_{n-1}^{-1} + A_{n-1}^{+}.R_{n-1}.A_{n-1} & -A_{n-1}^{+}.R_{n-1} \\ -R_{n-1}.A_{n-1} & R_{n-1} \end{bmatrix}$$

$$\text{with} \quad \alpha_n^{-1} = \left[1 - \left|\mu_n\right|^2\right]\alpha_{n-1}^{-1} \quad \text{and} \quad P_0 = \alpha_0^{-1}$$

S. Verblunsky (PhD student of Littlewood)

$$A_n = \begin{bmatrix} A_{n-1} \\ 0 \end{bmatrix} + \mu_n.\begin{bmatrix} A_{n-1}^{(-)} \\ 1 \end{bmatrix} \quad \text{where}\ V^{(-)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \cdot^{\cdot^{\cdot}} & 0 \\ 1 & 0 & 0 \end{bmatrix}.V^{*}$$

Iwasawa (Lie Group Theory)

  - Verblunsky/Trench Theorem: Exitence of diffeomorphism φ

$$\begin{aligned} \varphi : THPD_n &\rightarrow R_{+}^{*} \times D^{n-1} \\ R_n &\mapsto \left(P_0, \mu_1, \ldots, \mu_{n-1}\right) \end{aligned}$$

$$\text{with} \quad D = \left\{ z \in C / \left|z\right| < 1 \right\}$$

Preservation of Toeplitz Structure by Verblunsky Coefficients

- **Conformal Information Geometry metric (metric = Hessian of Entropy):**

$$g_{ij} \equiv \frac{\partial^2 \tilde{\Phi}}{\partial \theta_i^{(n)} \partial \theta_j^{(n)*}}$$

E. Kähler

with $\theta^{(n)} = \left[P_0 \quad \mu_1 \quad \cdots \quad \mu_{n-1} \right]^T$

μ_k : Verblunsky coefficient with $|\mu_k| < 1$

- **Entropy Φ as Kähler potential:**

$$\tilde{\Phi}(R_n, P_0) = \log\left(\det R_n^{-1}\right) - \log(\pi.e) = -\sum_{k=1}^{n-1} (n-k).\ln\left[1 - |\mu_k|^2\right] - n.\ln\left[\pi.e.P_0\right]$$

- **Conformal metric on Verblunsky parameterization:**

$$ds_n^2 = d\theta^{(n)+} \left[g_{ij} \right] d\theta^{(n)} = n.\left(\frac{dP_0}{P_0}\right)^2 + \sum_{i=1}^{n-1} (n-i) \frac{|d\mu_i|^2}{\left(1 - |\mu_i|^2\right)^2}$$

Seminal Erich Kähler Paper, 1932, Hamburg

« Kähler Erich, Mathematical Works », Edited by R. Berndt and O. Riemenschneider, Berlin, Walter de Gruyter, ix, 2003

Erich Kähler

$$ds^2 = \sum \frac{\partial^2 U}{\partial x_i \, \partial \bar{x}_k} dx_i \, d\bar{x}_k$$

invariant gegenüber der Gruppe der „hyperfuchsschen" Transformationen (4). Analog ist für die „hyperabelschen" Transformationen

$$x_i' = \frac{\alpha_i x_i + \beta_i}{\gamma_i x_i + \delta_i} \qquad (i = 1, 2, \cdots, n),$$

die die Einheitskreise

$$1 - x_i \bar{x}_i = 0 \qquad (i = 1, 2, \cdots, n)$$

festlassen, die aus dem Potential

$$U = \sum_{i=1}^{n} k_i \log (1 - x_i \bar{x}_i) \qquad (k_i \text{ Konstanten})$$

ableitbare Metrik invariant, und es ist klar, daß auch alle Zwischenfälle, etwa die aus hyperfuchsschen Transformationen in r und s $(r + s = n)$ Variablen komponierten Gruppen, zu Metriken von jenem Typus führen.

Über eine bemerkenswerte Hermitesche Metrik.

Von ERICH KÄHLER in Hamburg.

1.

Bei der Untersuchung der Invarianten einer reell $2n$-dimensionalen Hermiteschen Metrik[1])

$$(1) \qquad ds^2 = \sum g_{i\bar{k}} \, dx_i \, d\bar{x}_k$$

...

Auf die Verwendung der vorliegenden formalen Entwicklungen für die Theorie der automorphen Funktionen und auf die Analogie der Gleichungen

$$D(U) = 0, \qquad D(U) = e^{kU}$$

mit den klassischen Differentialgleichungen

$$\Delta U = 0, \qquad \Delta U = e^{kU}$$

auf die bereits G. Giraud[1]) und A. Bloch[2]) hingewiesen haben, gedenke ich in einer späteren Arbeit einzugehen.

[1]) G. Giraud, Sur une équation aux dérivées partielles, non linéaire etc. Comptes Rendus 166, I (1918), p. 893.

[2]) A. Bloch, Sur une nouvelle et importante généralisation de l'équation de Laplace. L'Enseignement Mathém. 26 (1927), p. 52.

Hamburg, den 22. Oktober 1932.

Entropy U=-logdet[R] , for Complex Autoregressive model, can be considered as a Kähler Potential that depends on reflection coefficient. This expression is exactly the same than the first potential proposed by E. Kähler

Regularized Verblunsky Coefficients: Regularized Burg Algorithm

- **Regularized Burg Algorithm (THALES Patent)**

. Initialisation :

$$f_0(k) = b_0(k) = z(k) \quad , \quad k=1,\dots,N \quad (\mathrm{N: nb. ech.})$$

$$P_0 = \frac{1}{N}.\sum_{k=1}^{N} |z(k)|^2$$

$$a_0^{(0)} = 1$$

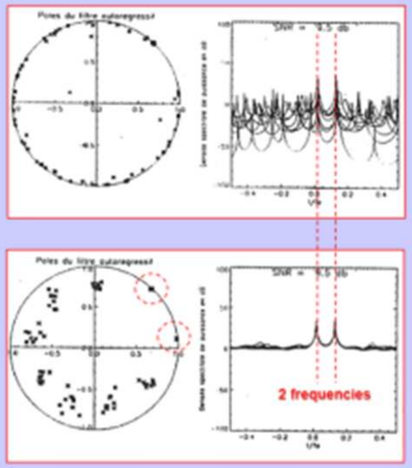
. Step (n) : For $n = 1$ to M

Regularization

$$\mu_n = -\frac{\dfrac{2}{N-n}\displaystyle\sum_{k=n+1}^{N} f_{n-1}(k).b_{n-1}^*(k-1) + 2.\sum_{k=1}^{n-1}\beta_k^{(n)}.a_k^{(n-1)}.a_{n-k}^{(n-1)}}{\dfrac{1}{N-n}\displaystyle\sum_{k=n+1}^{N}\left|f_{n-1}(k)\right|^2 + \left|b_{n-1}(k-1)\right|^2 + 2.\sum_{k=0}^{n-1}\beta_k^{(n)}.\left|a_k^{(n-1)}\right|^2} \quad \text{with} \quad \beta_k^{(n)} = \gamma_1.(2\pi)^2.(k-n)^2$$

$$\begin{cases} a_0^{(n)} = 1 \\ a_k^{(n)} = a_k^{(n-1)} + \mu_n.a_{n-k}^{(n-1)*} \quad , \quad k=1,\dots,n\text{-}1 \\ a_n^{(n)} = \mu_n \end{cases}$$

$$\begin{cases} f_n(k) = f_{n-1}(k) + \mu_n.b_{n-1}(k-1) \\ b_n(k) = b_{n-1}(k-1) + \mu_n^*.f_{n-1}(k) \end{cases}$$



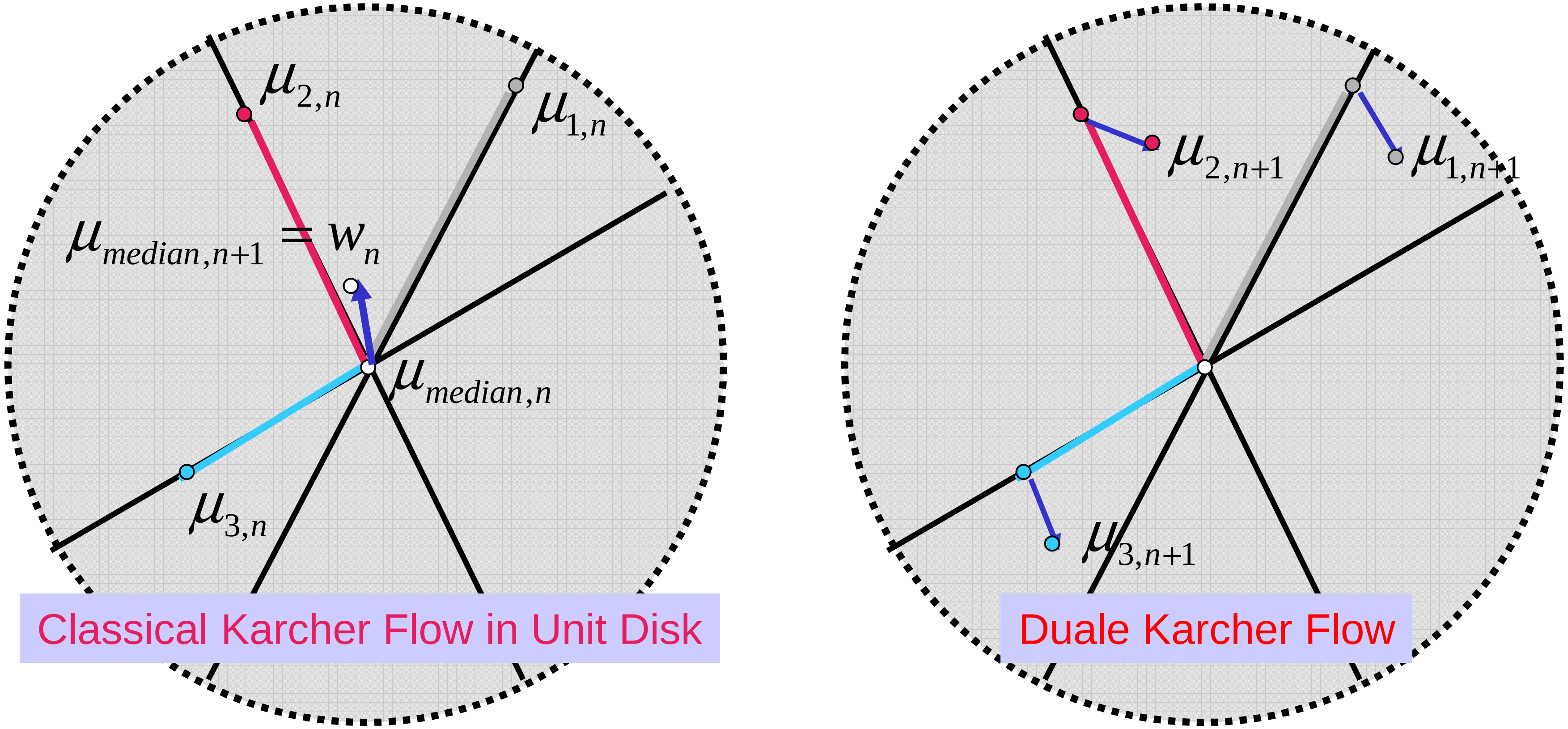
Non-regularised Burg Algorithm (model order : 9)

Regularised Burg Algorithm (model order : 9)

Median by Median Reflection/Verblunsky Coefficients μ_k

$$w_n = \gamma_n \sum_{\substack{k=1 \\ k \neq l}}^{m} \frac{\mu_{k,n}}{|\mu_{k,n}|} \quad avec \quad \{l / |\mu_{l,n}| < \varepsilon\}$$

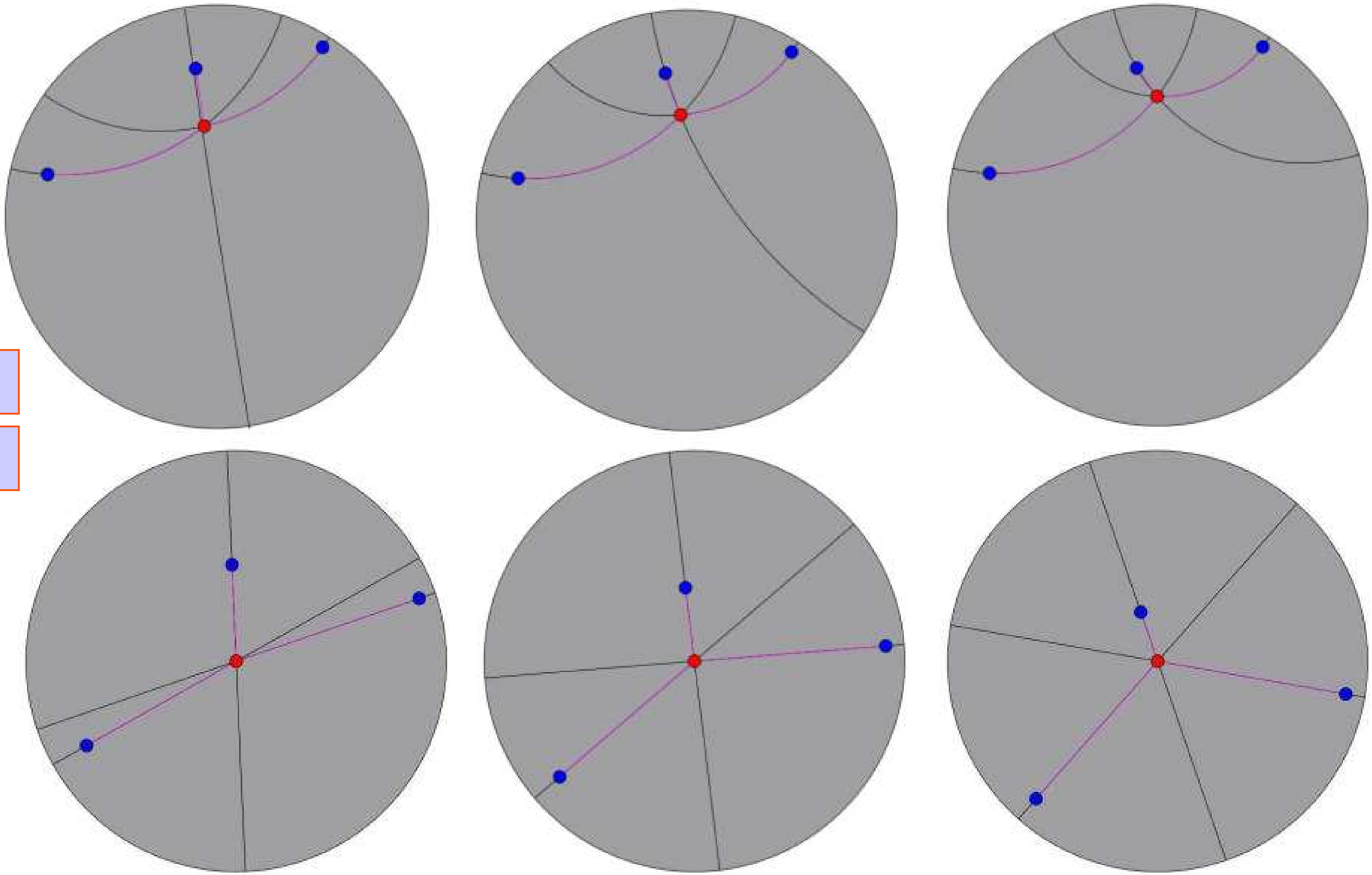
$$\mu_{k,n+1} = \frac{\mu_{k,n} - w_n}{1 - \mu_{k,n}.w_n^*}$$



$$\mu_{median,n+1} = \frac{\mu_{median,n} + w_n}{1 + \mu_{median,n} w_n^*}$$

THALES

Median by Median Reflection/Verblunsky Coefficients μ_k



Inductive Mean (K.T. Sturm)

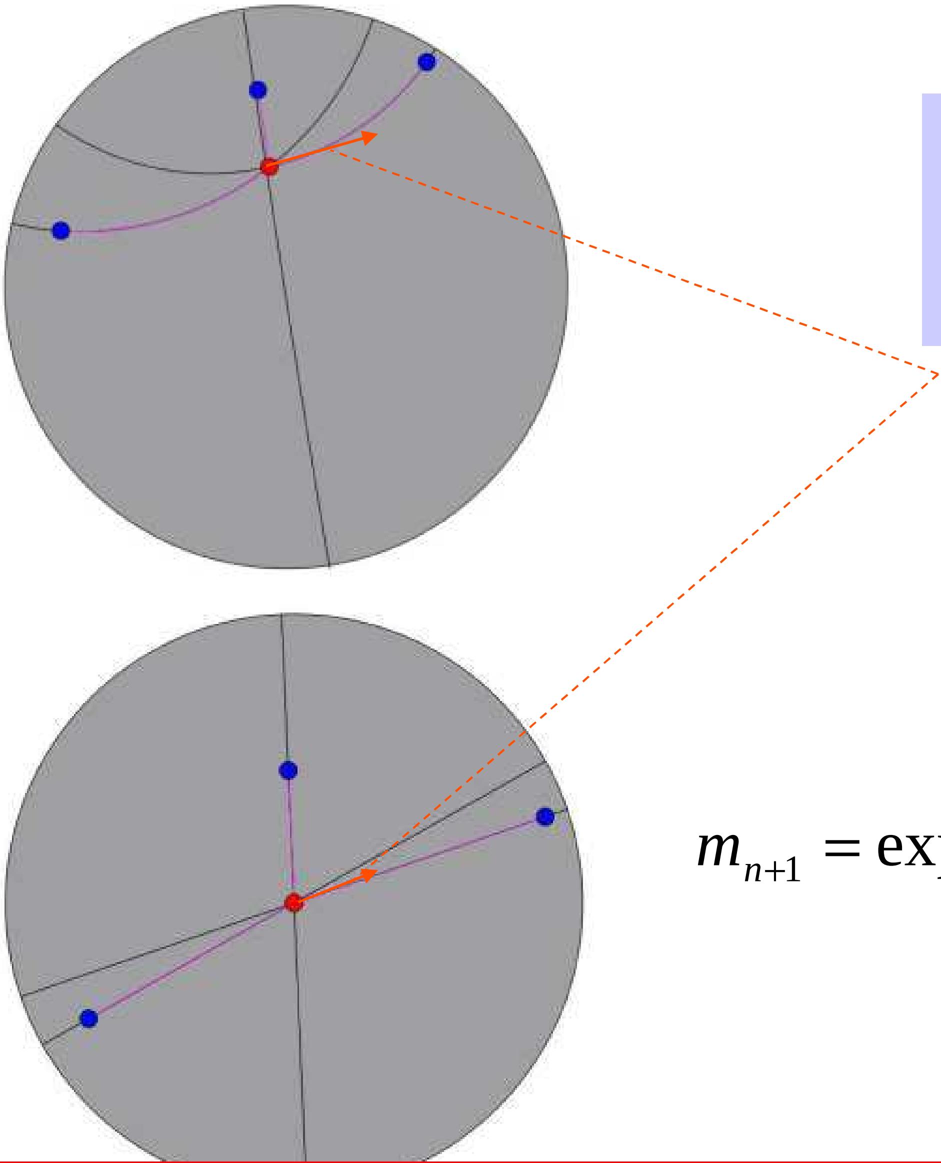
- A natural way to define the « expectation » E(Y) of a random variable Y is to use generalizations of the law of large numbers.
- Given any sequence $\{B_k\}_{k=1}^{n}$ of points, we define a new sequence $\{S_k\}_{k=1}^{n}$ of points by induction on n as follow (inductive mean) :

$$S_1 = B_1 \text{ and } S_n = S_{n-1} \bullet_{1/n} B_n = S_{n-1}^{1/2}\left(S_{n-1}^{-1/2} B_n S_{n-1}^{-1/2}\right)^{1/n} S_{n-1}^{1/2}$$

- strongly depend on permutations but $E\left(\delta^2\left(E(B_1), S_n\right)\right) \leq \frac{1}{n} V\left(B_1\right)$
- Extension of scalar case :

$$s_1 = b_1 \text{ and } s_n = s_{n-1} \bullet_{1/n} b_n = \left(1 - \frac{1}{n}\right) s_{n-1} + \frac{1}{n} b_n$$

Stochastic Approach: Arnaudon Flow



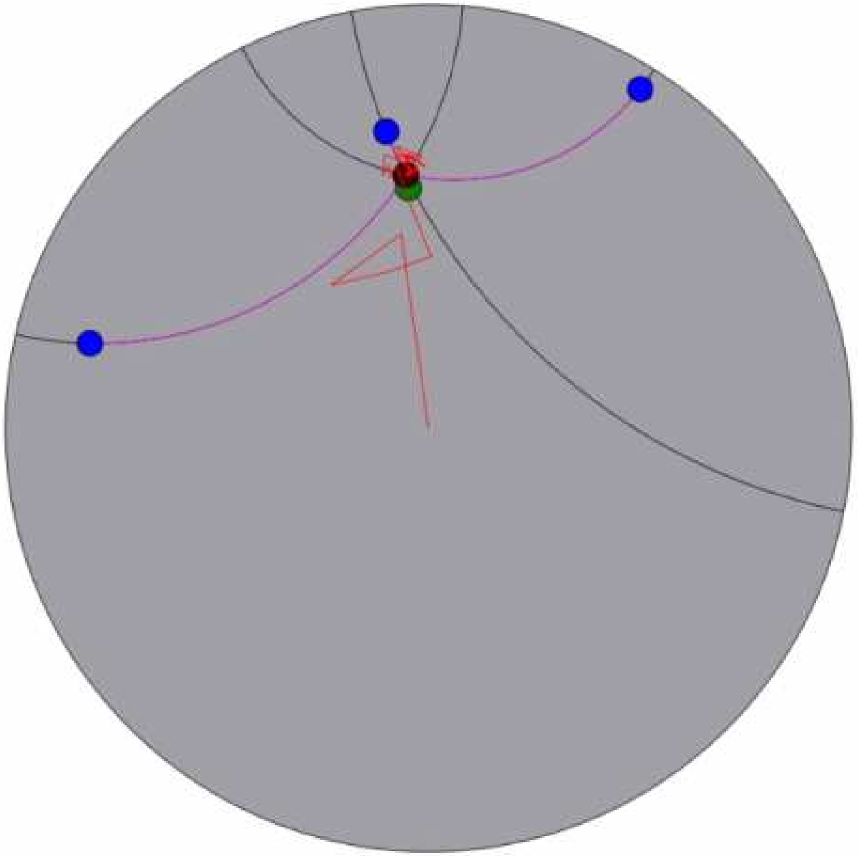
Random selection at each iteration of 1 point in the disk + driven of the flow along the geodesic

$$w_n = \gamma_n \frac{\mu_{rand(n),n}}{\left|\mu_{rand(n),n}\right|}$$

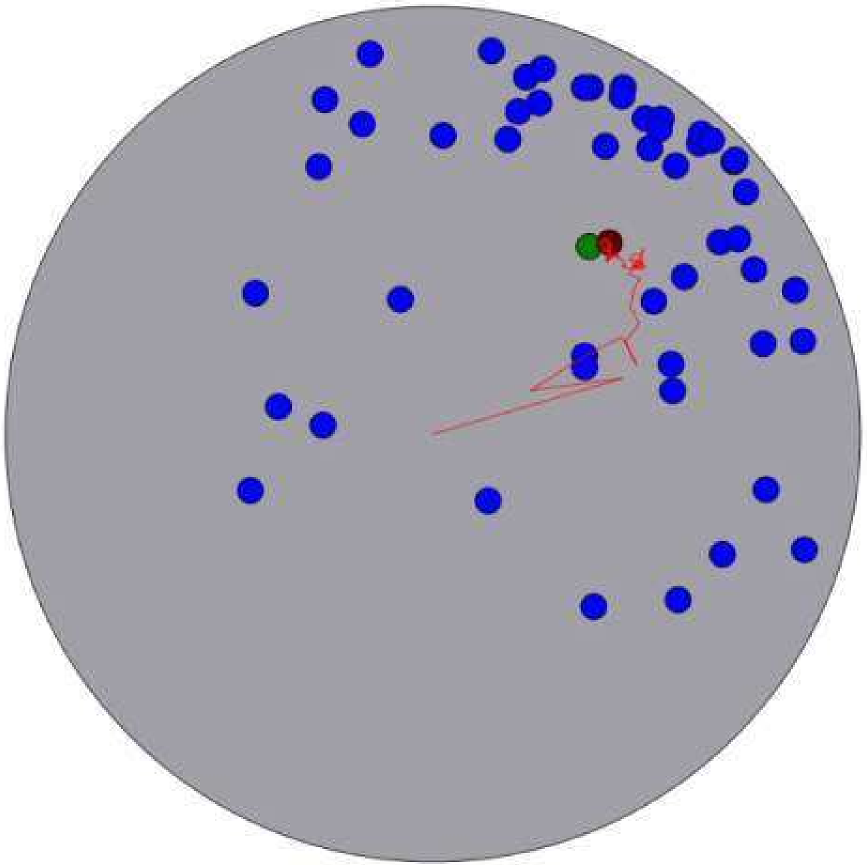
$$m_{n+1} = \exp_{m_n}\left(t_n \cdot \frac{\exp_{m_n}^{-1}\left(x_{rand(n)}\right)}{\left\|\exp_{m_n}^{-1}\left(x_{rand(n)}\right)\right\|} \right)$$

M. Arnaudon, C. Dombry, A.Phan, L.Yang, "Stochastic algorithms for computing means of probability measures", http://hal.archives-ouvertes.fr/hal-00540623

Arnaudon Stochastic Flow



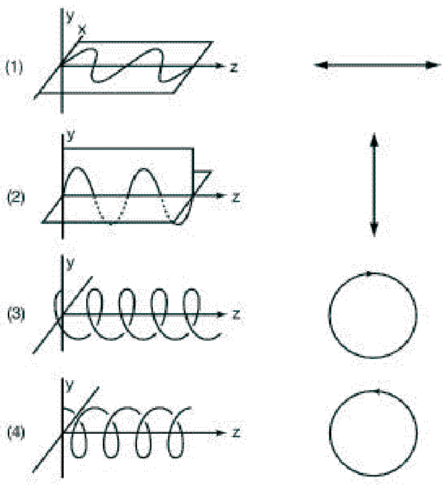
$$w_n = \gamma_n \frac{\mu_{rand(n),n}}{\left|\mu_{rand(n),n}\right|}$$

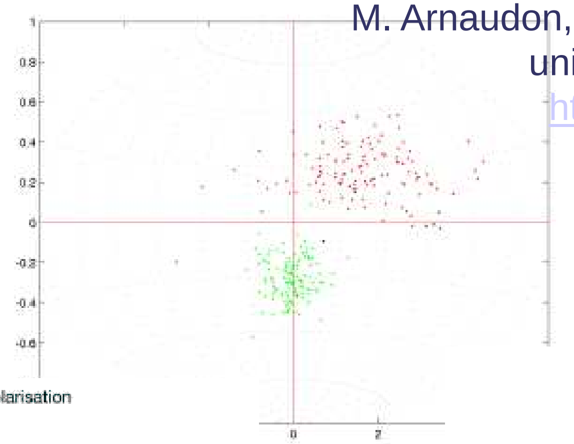


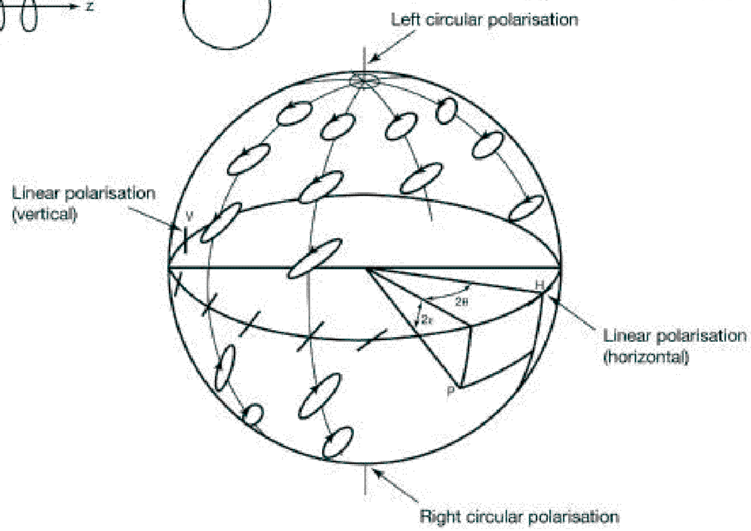
$$m_{n+1} = \exp_{m_n}\left(t_n \cdot \frac{\exp_{m_n}^{-1}\left(x_{rand(n)}\right)}{\left\|\exp_{m_n}^{-1}\left(x_{rand(n)}\right)\right\|} \right)$$

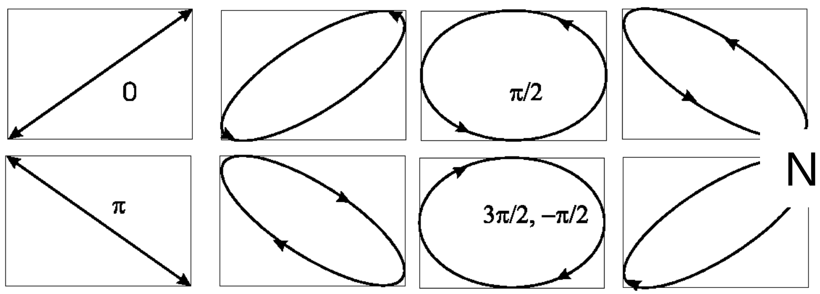
Arnaudon Flow extended on compact space

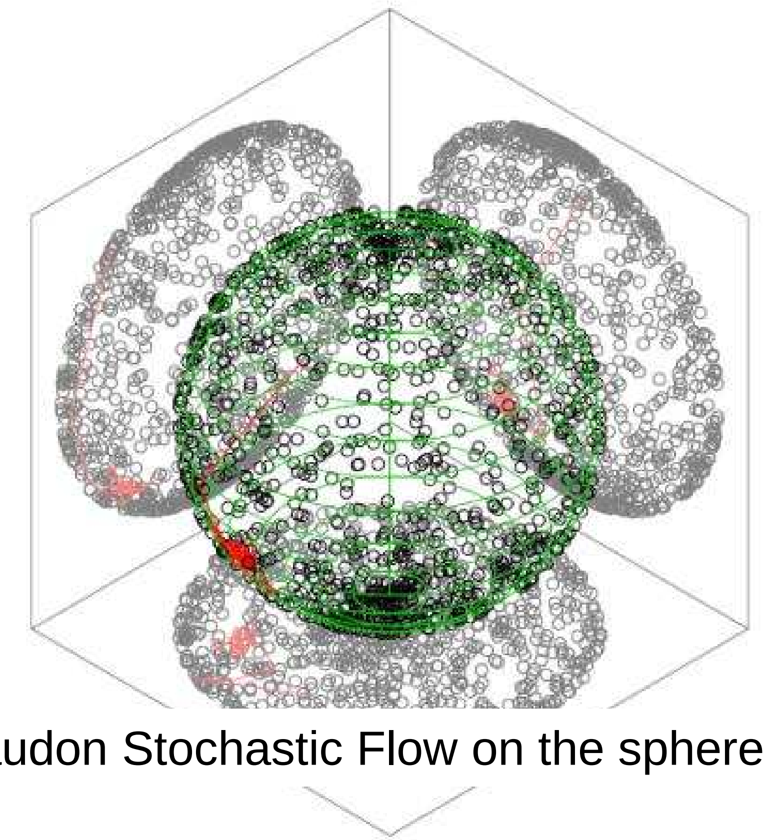
M. Arnaudon, L. Miclo, "Means in complete manifolds: uniqueness and approximation", http://arxiv.org/abs/1207.3232











New Arnaudon Stochastic Flow on the sphere

Conformal Busemann Barycenter (Douady-Earle,…)

- **Case of Poincaré unit disk:** $D = \{z \in C \,/\, |z| < 1\}$
  - G+ sub-group of automorphisms preserving orientation

    $$z \mapsto \lambda \frac{z-a}{1-a^*z} \quad \text{avec } |z| = 1 \text{ et } |a| < 1$$

  - G operate on $P(S^1)$ of probability measure of $S^1 = \{z \in C \,/\, |z| = 1\}$
  - Principe : assign to all probability measure μ on S^1 a point $B(\mu) \in D$ such that $\mu \mapsto B(\mu)$ is conformal and verify:

    $$B(\mu) = 0 \Leftrightarrow \int_{S^1} \zeta . d\mu(\zeta) = 0$$

  - There is only one conformal way to assign to each probability measure μ on S^1 one vectors fields ξ_μ on *D* such that: $\xi_\mu(0) = \int_{S^1} \zeta . d\mu(\zeta) = 0$

$$\xi_\mu(w) = \frac{1}{(g_w)'(w)} \xi_{(g_w)^*(\mu)}(0) = \left(1 - |w|^2\right) \int_{S^1} \left(\frac{\zeta - w}{1 - w^* \zeta} \right) d\mu(\zeta) = 0$$

with $g_w(z) = \dfrac{z-w}{1-w^*z}$

THALES

Median Fréchet Barycenter = Busemann Barycenter

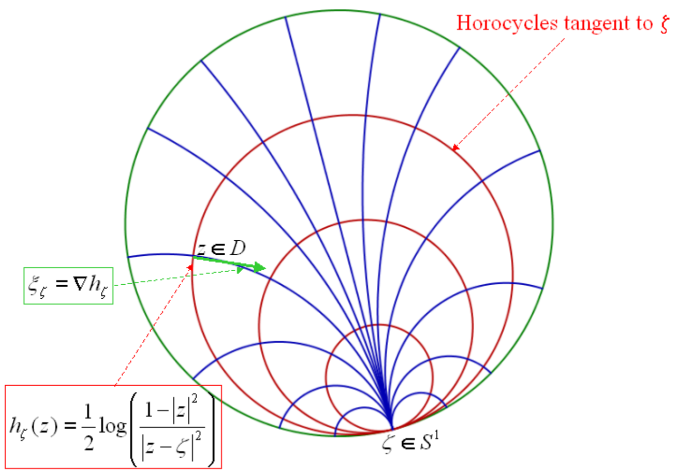
- **Case of Poincaré Disk**
  - $\xi_\mu(z)$ could be described according to $\xi_\zeta(z)$ that is the tangent vector of the geodesic in $z \in D$ pointing to $\zeta \in S^1$:

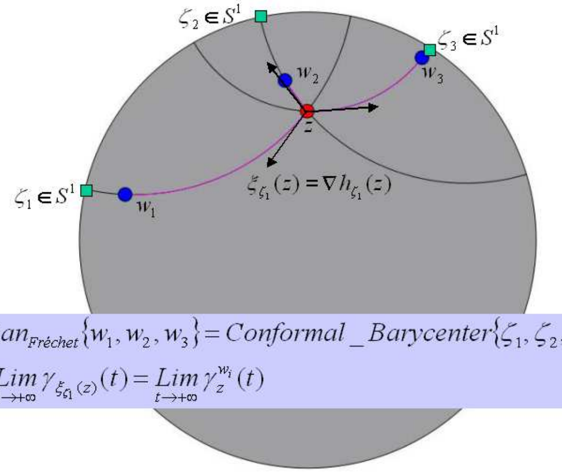
$$\xi_\mu(z) = \int_{S^1} \xi_\zeta(z)\,d\mu(\zeta)$$

  - In Poincaré Geometry of unit disk, vectors field ξ_ζ is the gradient of a function h_ζ whose the level sets are the horocycles tangents to S^1 in $\zeta \in S^1$:

$$\xi_\mu = \nabla h_\mu \quad \text{avec} \quad h_\mu : z \mapsto \int_{S^1} h_\zeta(z)\,d\mu(\zeta)$$

$$h_\mu(z) == \int_{S^1} \frac{1}{2}\log\left(\frac{1-|z|^2}{|z-\zeta|^2}\right) d\mu(\zeta) = \int_{S^1} \lim_{r \to 1^-} \left[d(0,r) - d(z, r\zeta)\right] d\mu(\zeta)$$



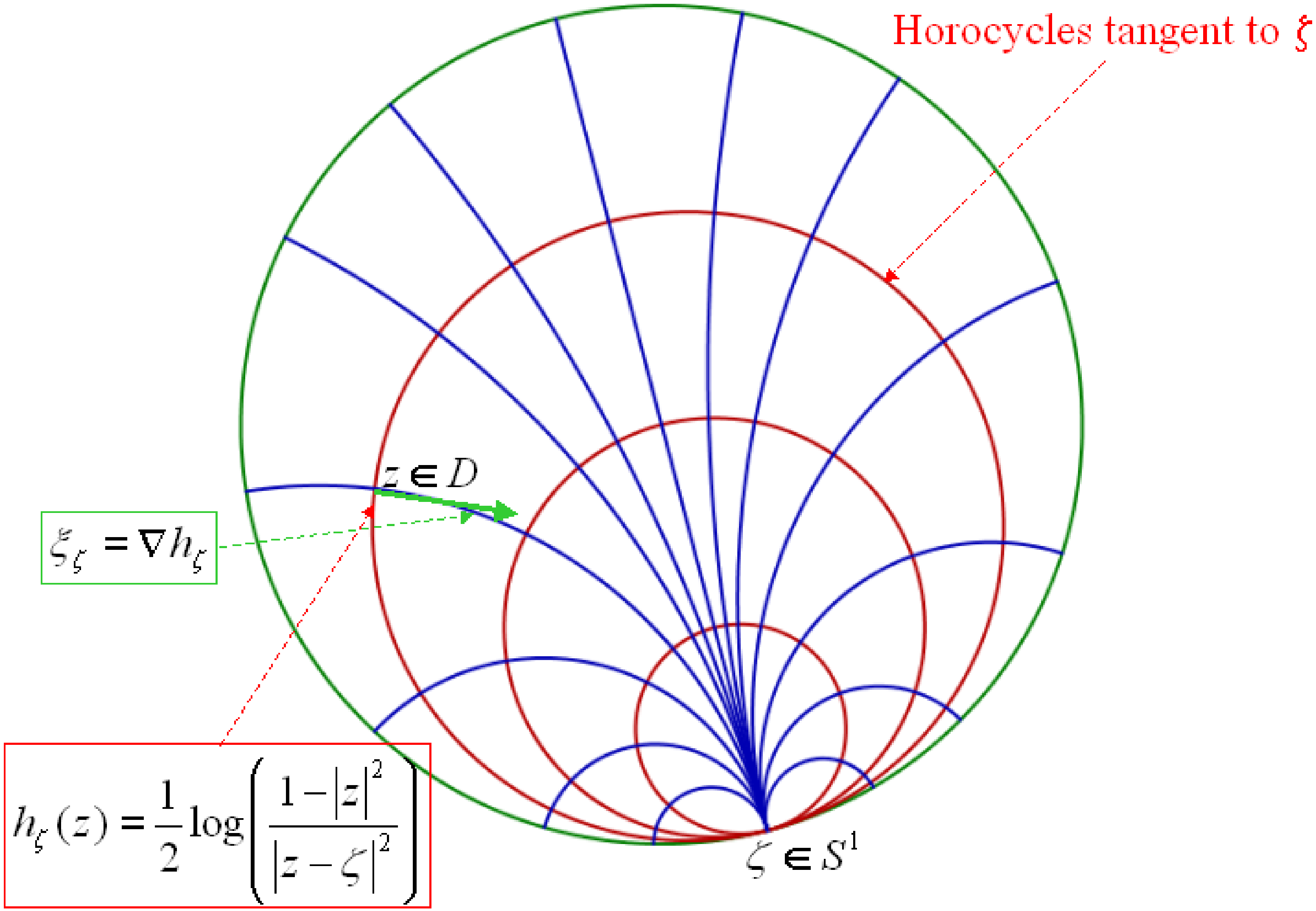


$$Median_{Fréchet}\{w_1, w_2, w_3\} = Conformal\_Barycenter\{\zeta_1, \zeta_2, \zeta_3\}$$

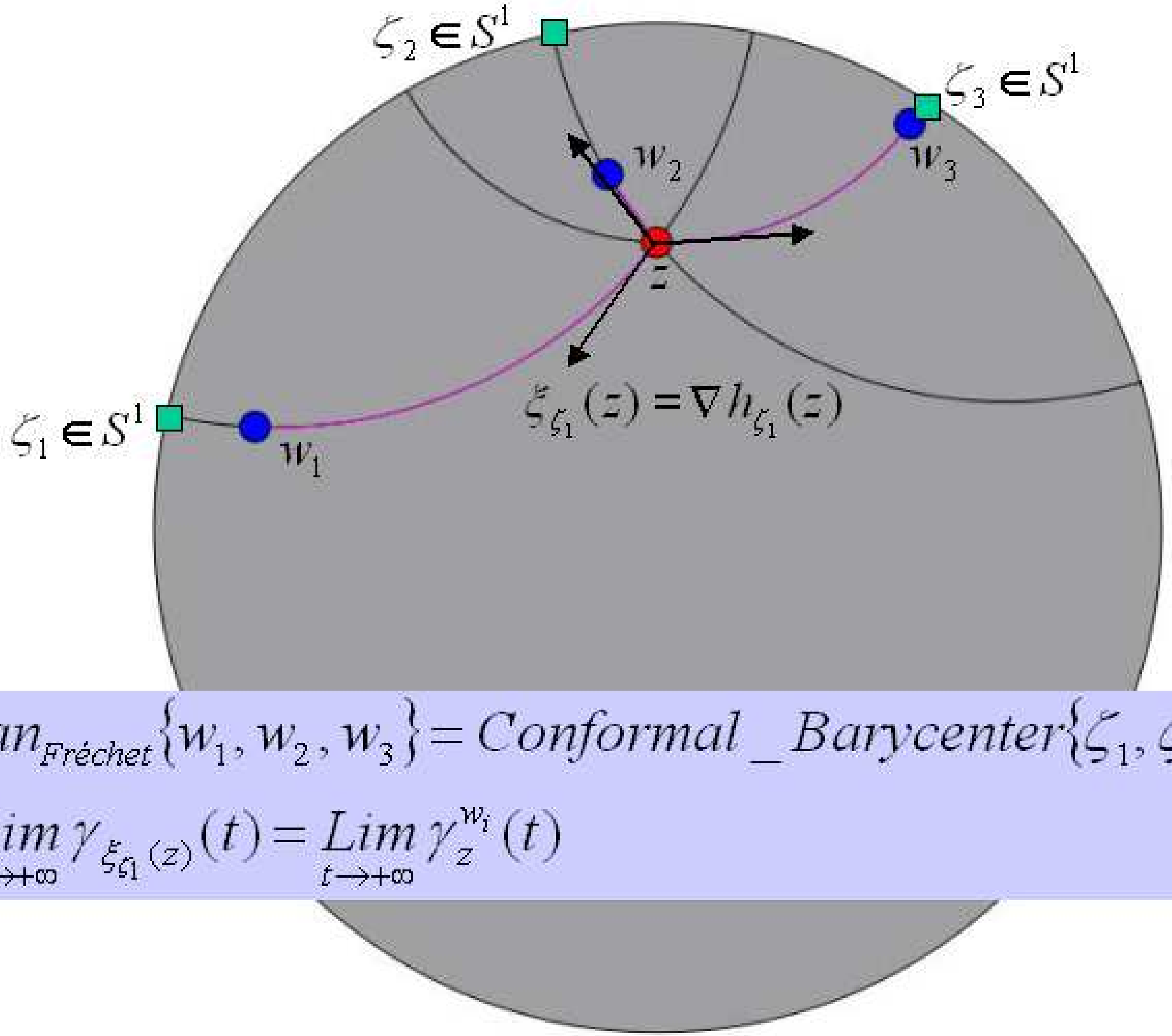
$$\zeta_i = \lim_{t \to +\infty} \gamma_{\xi_{\zeta_i}(z)}(t) = \lim_{t \to +\infty} \gamma_z^{w_i}(t)$$

THALES

Median Fréchet Barycenter = Busemann Barycenter



Median Fréchet Barycenter = Busemann Barycenter



$$Median_{Fréchet}\{w_1, w_2, w_3\} = Conformal\_Barycenter\{\zeta_1, \zeta_2, \zeta_3\}$$

$$\zeta_i = \underset{t \to +\infty}{Lim}\, \gamma_{\xi_{\zeta_1}(z)}(t) = \underset{t \to +\infty}{Lim}\, \gamma_z^{w_i}(t)$$

Busemann Barycenter on Cartan-Hadamard Manifold

- **From Cartan Center of Mass to Busemann Barycenter**
  - Elie Cartan has proved for simply connected riemannian manifolds with negative or null curvature, existence and unicity of the minimum of:

$$x \mapsto \int_{H^n} d(x,z)^2 \, d\nu(z)$$

  - Following function, is simply connexe and reachs its minimum:

$$x \mapsto \int_{H^n} d(x,z) \, d\nu(z)$$

  - when z goes to infinity converging to a point θ of ∂H^n, the normalized distance function to z converges : Busemann function

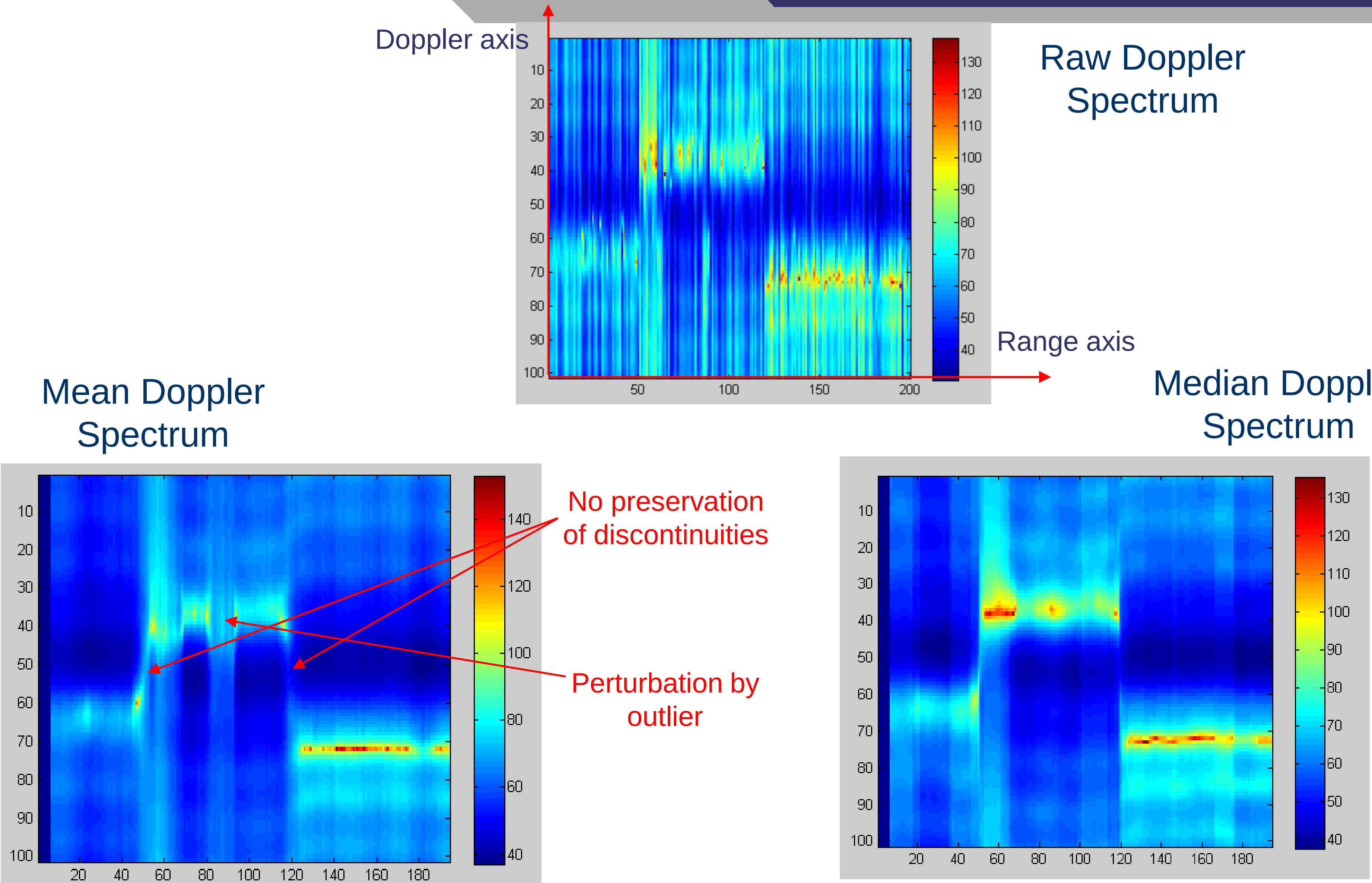
$$B_\theta(x) = \lim_{z \to \theta} d(x,z) - d(O,z)$$

  - Fixing an origine $O \in H^n$. Let ν finite measure on ∂H^n with a support of more than 2 points. The following strictly convexe function reachs its minimum at a unique point, independant of O choice and is called Busemann Barycenter :

$$\beta_\nu(x) = \int_{\partial H^n} B_\theta \, d\nu(\theta)$$

THALES

Comparison of Mean & Median Doppler Spectrum



THALES

Other approach to smooth spectrum : Fourier Heat Equation

- **Diffusion Fourier Equation on 1D graph (scalar case)**
  - Approximation par un Laplacien discret :

J. Fourier

$$\frac{\partial u}{\partial t}=\frac{\partial^2 u}{\partial x^2}\Rightarrow\frac{\partial u}{\partial t}=\frac{1}{\nabla x}\left(\frac{u_{n+1}-u_n}{\nabla x}-\frac{u_n-u_{n-1}}{\nabla x}\right)=\frac{2}{\nabla x^2}\left(\hat{u}_n-u_n\right)$$

with arithmetic mean of adjacent points : $\hat{u}_n=\left(u_{n+1}+u_{n-1}\right)/2$

  - Dicrete Fourier Heat Equation for Scalar Values in 1D :

$$u_{n,t+1}=(1-\rho).u_{n,t}+\rho.\hat{u}_{n,t}\quad\text{with}\quad\rho=\frac{2.\nabla t}{\nabla x^2}$$

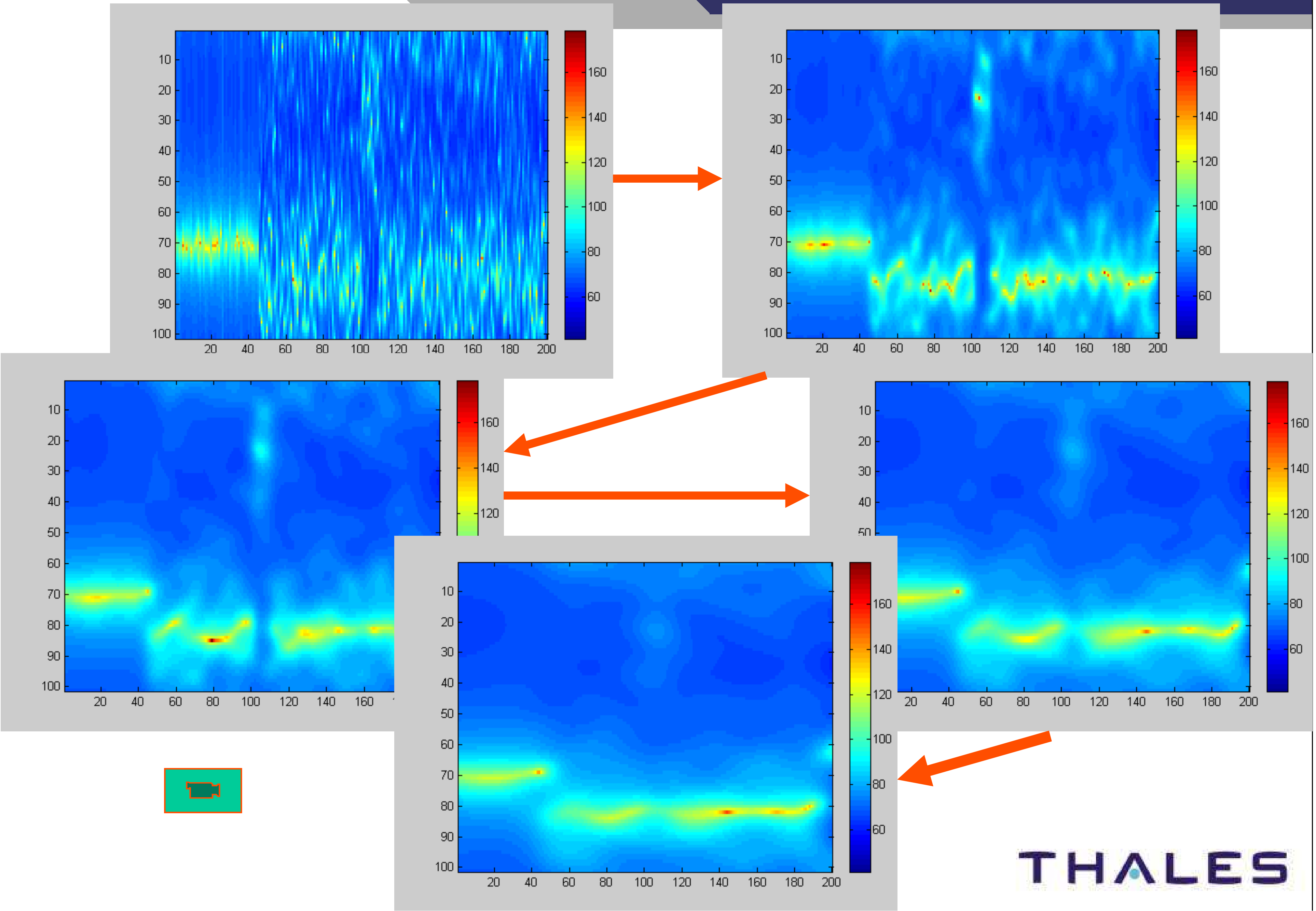
- **By Analogy, we can define Fourier Heat Equation in 1D grapf of HDP(n) matrices :**

$$X_{n,t+1}=X_{n,t}^{1/2}e^{\rho\log\left(X_{n,t}^{-1/2}\hat{X}_{n,t}X_{n,t}^{-1/2}\right)}X_{n,t}^{1/2}=X_{n,t}^{1/2}\left(X_{n,t}^{-1/2}\hat{X}_{n,t}X_{n,t}^{-1/2}\right)^{\rho}X_{n,t}^{1/2}$$

$$\text{with}\quad\hat{X}_{n,t}=X_{n+1,t}^{1/2}\left(X_{n+1,t}^{-1/2}X_{n-1,t}X_{n+1,t}^{-1/2}\right)^{1/2}X_{n+1,t}^{1/2}=X_{n+1,t}\circ_{1/2}X_{n-1,t}$$

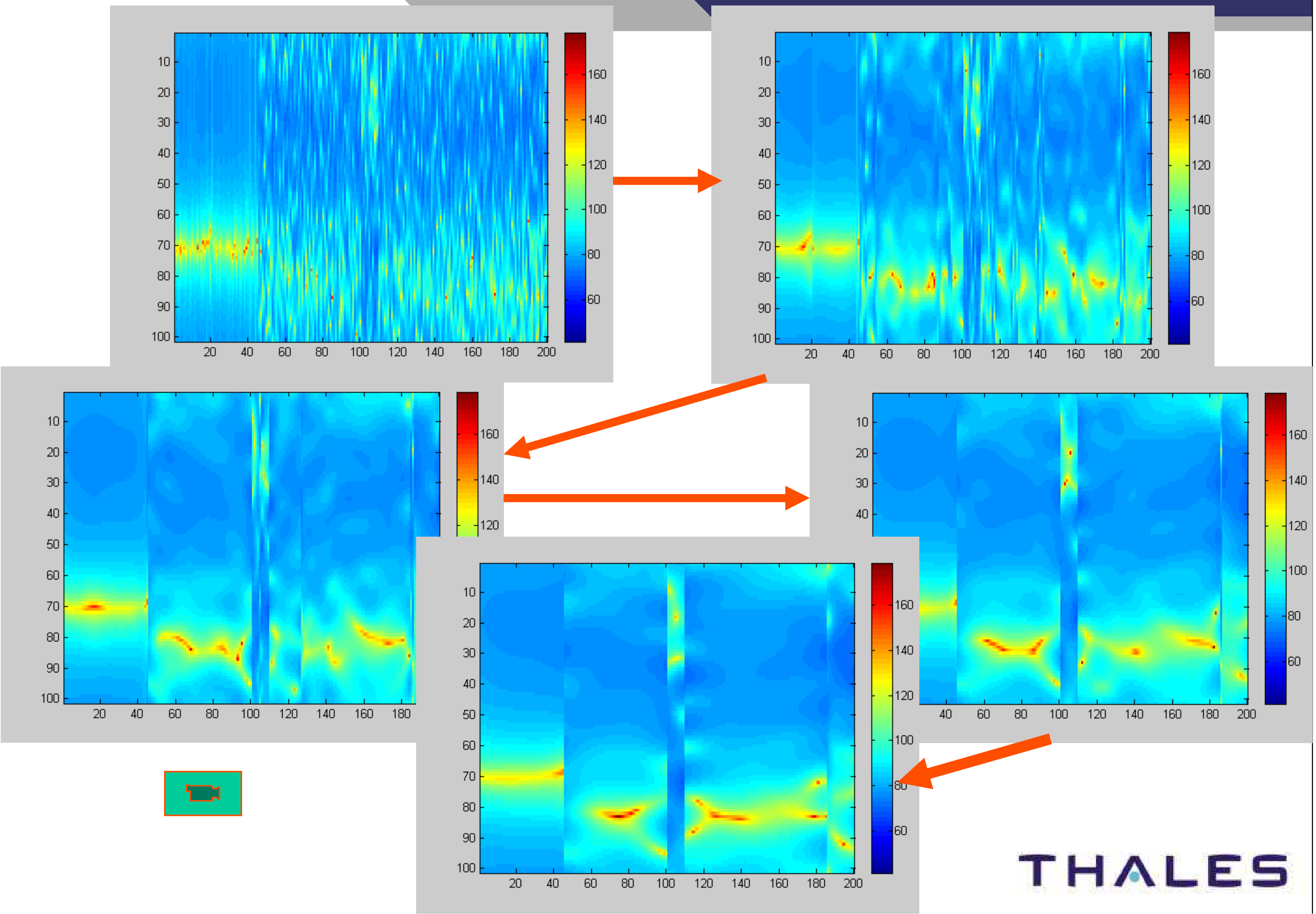
THALES

Isotropic Diffusion of Doppler Spectrum



THALES

Anisotropic Diffusion of Doppler Spectrum



THALES

Gauss-Laplace Law

- Fisher Information Matrix for Gauss-Laplace Gaussian Law:

$$I(\theta)=\sigma^{-2}.\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix} \quad \text{with} \quad E\left[\left(\theta-\hat{\theta}\right)\left(\theta-\hat{\theta}\right)^{T}\right]\geq I(\theta)^{-1} \quad \text{and} \quad \theta=\begin{pmatrix}m\\ \sigma\end{pmatrix}$$

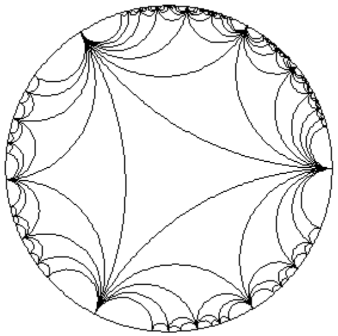
- Rao-Chentsov Metric of Information Geometry

$$ds^2=d\theta^T.I(\theta).d\theta=\frac{dm^2}{\sigma^2}+2.\frac{d\sigma^2}{\sigma^2}=2.\sigma^{-2}\left[\left(\frac{dm}{\sqrt{2}}\right)^2+(d\sigma)^2\right]$$

H. Poincaré

- Fisher Metric is equal to Poincaré Metric for Gaussian Laws:

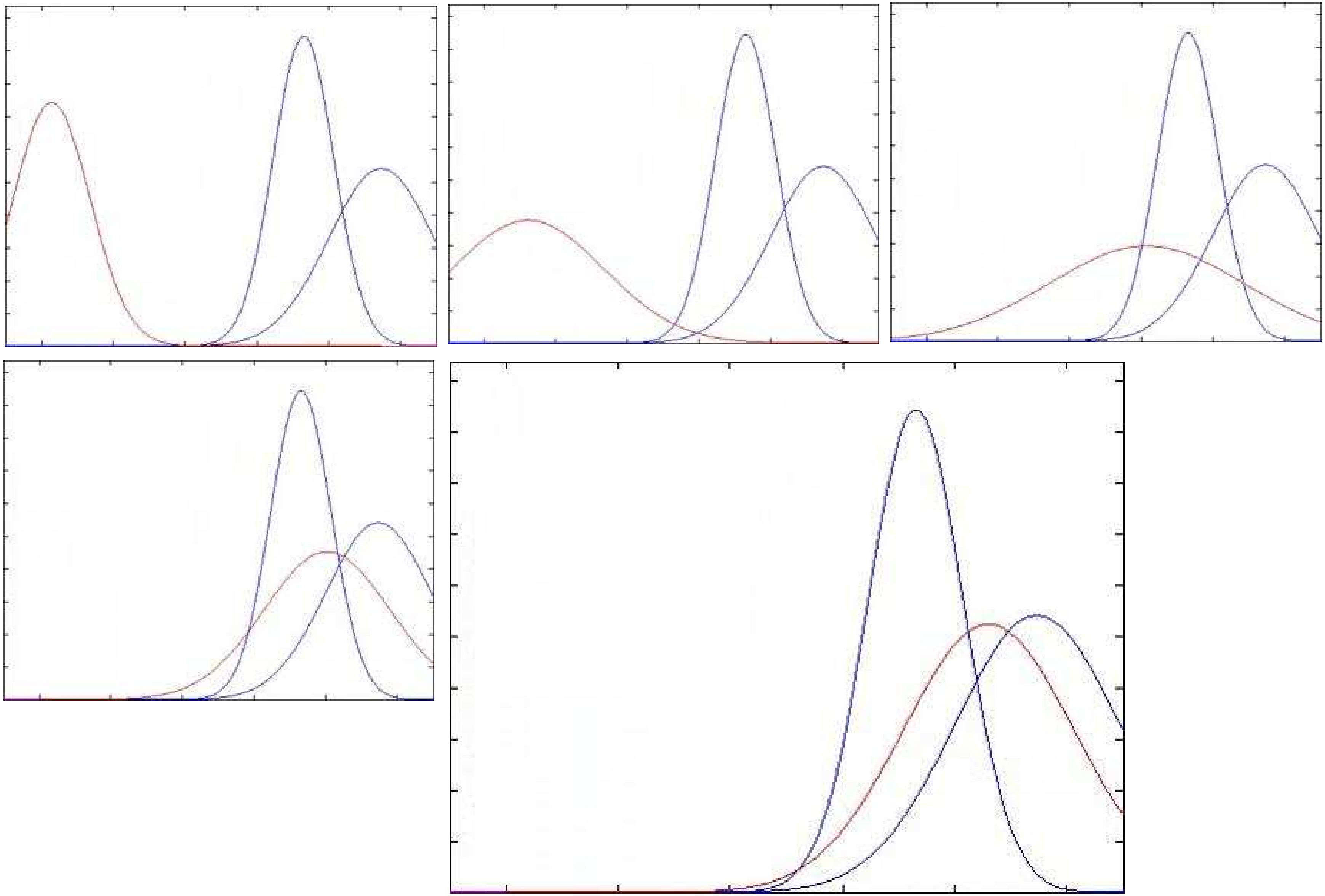
$$z=\frac{m}{\sqrt{2}}+i.\sigma \qquad \omega=\frac{z-i}{z+i} \quad (|\omega|<1) \qquad \Rightarrow ds^2=8.\frac{|d\omega|^2}{\left(1-|\omega|^2\right)^2}$$



$$d^2(\{m_1,\sigma_1\},\{m_2,\sigma_2\})=2.\left(\log\frac{1+\delta(\omega^{(1)},\omega^{(2)})}{1-\delta(\omega^{(1)},\omega^{(2)})}\right)^2$$

$$\text{with} \quad \delta(\omega^{(1)},\omega^{(2)})=\left|\frac{\omega^{(1)}-\omega^{(2)}}{1-\omega^{(1)}\omega^{(2)*}}\right|$$

Median of Gaussian laws



Lien entre géométrie Hessienne et de l'information

Travaux de Jacob Burbea (1984)

- On considère une densité de probabilité pour une variable de C^n

$$g_{ij}(z) = \int p^{-1}(t/z)\frac{\partial p(t/z)}{\partial z_i}\frac{\partial p(t/z)}{\partial z_j^*}d\mu(t) = \int p(t/z)\frac{\partial \log p(t/z)}{\partial z_i}\frac{\partial \log p(t/z)}{\partial z_j^*}d\mu(t)$$

$$ds^2(z) = \sum_{i=1}^{N} g_{ij}dz_i dz_j^* = \left\| p^{1/2}\partial \log p \right\|_\mu^2$$

- Identification avec le carré du module d'une fonction normalisé:

$$p(t/z) = \left|\psi(t/z)\right|^2 \quad \text{avec} \quad \left\|\psi(./z)\right\|_\mu^2 = 1$$

$$g(t/z) = \sqrt{K(z,z^*)}.\psi(t/z) \qquad K(z,w^*) = \langle g(t/z), g(t/w)\rangle_\mu$$

- On considère la pseudo-distance de Skwarczynski :

$$\lambda(z,w) = \sqrt{1 - \left|\int \psi(t/z)\psi(t/w)d\mu(t)\right|} = \sqrt{1 - \left|\langle \psi(t/z), \psi(t/w)\rangle_\mu\right|}$$

$$ds^2_{Bergman} = d^2\lambda(z,w)\Big|_{w=z} = \left\|d\psi\right\|_\mu^2 - \left|\langle \psi, d\psi\rangle_\mu\right|^2 = K^{-2}\left[K\partial\bar{\partial}K - \left|\partial K\right|^2\right] = \partial\bar{\partial}\log K = \sum_{i,j=1}^{N}\frac{\partial \log K(z,z^*)}{\partial z_i \partial z_k}dz_i dz_k^*$$

- On en déduit l'équivalence entre métrique de Fisher et métrique Hessienne :

$$\begin{cases} p(t/z) = \left|\psi(t/z)\right|^2 \\ g(t/z) = \sqrt{K(z,z^*)}.\psi(t/z) \end{cases} \Rightarrow \log p(t/z) = \log g(t/z) + \log g^*(t/z) - \log K(z,z^*)$$

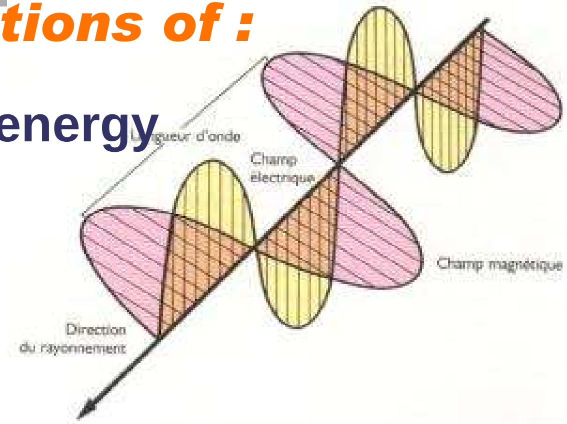
$$\frac{\partial^2 \log p(t/z)}{\partial z_i \partial z_j^*} = -\frac{\partial^2 \log K(z,z^*)}{\partial z_i \partial z_j^*} \Rightarrow g_{ij} = -E\left[\frac{\partial^2 \log p(t/z)}{\partial z_i \partial z_j^*}\right] = \frac{\partial^2 \log K(z,z^*)}{\partial z_i \partial z_j^*}$$

www.thalesgroup.com

New « Ordered Statistic High Doppler Resolution CFAR » (OS-HDR-CFAR)

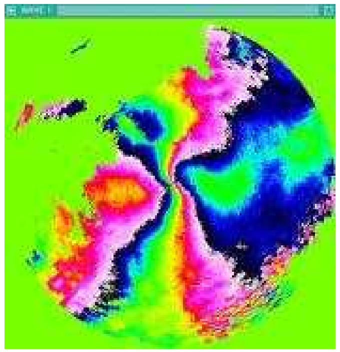
THALES

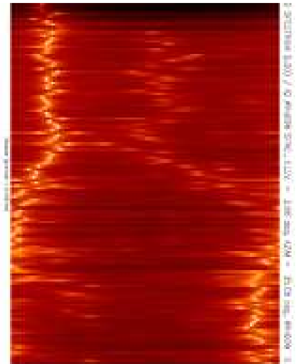
Radar estimates Electromagnetic fluctuations of :

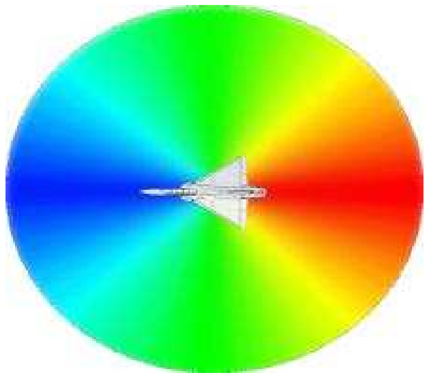


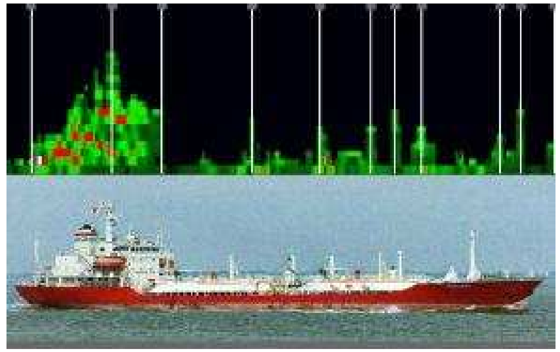
- Reflectivity : Amplitude of EM Wave reflected energy

- Phases variations due to Doppler-Fizeau

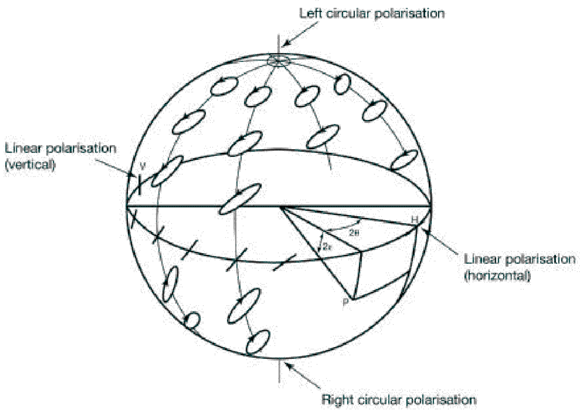


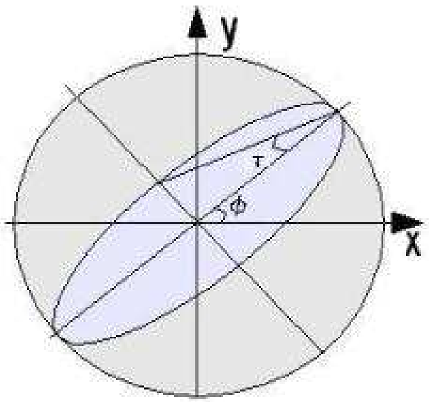


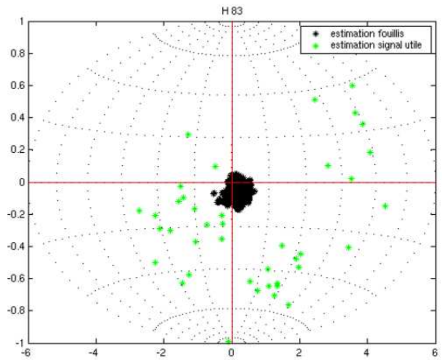




- Polarization of EM Wave







Radar estimates Electromagnetic fluctuations of :

- **Reflectivity : Amplitude of EM Wave reflected energy**

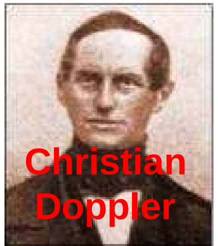
$$Z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \Rightarrow R_{reflectivité} = \frac{1}{n}\sum_{k=0}^{n} |z_k|^2$$

- **Phases variations due to Doppler-Fizeau**

$$f_{Doppler} = 2\frac{V_r / c}{1 - V_r / c} f_0 \underset{V_r/c \ll 1}{\approx} 2.\frac{V_r}{\lambda}$$

$$\text{Transmit}: s(t) = u(t).e^{2\pi f_0 t}$$

$$\text{Receive}: z(t) \approx A.u\left(t - \frac{2R}{c}\right) e^{i\left(2\pi f_0 + 2\pi f_{Doppler}\right)t} e^{i\varphi}$$



Toeplitz Hermitian Positive Definite Covariance matrix

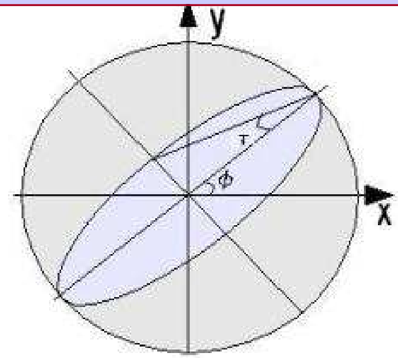
$$R = E\left[\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}^+\right] = \begin{bmatrix} c_0 & \overline{c}_1 & \cdots & \overline{c}_{n-1} \\ c_1 & c_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \overline{c}_1 \\ c_{n-1} & \cdots & c_1 & c_0 \end{bmatrix}$$

- **Polarization of EM Wave**

$$\vec{S} = \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} |z_1|^2 + |z_2|^2 \\ |z_1|^2 - |z_2|^2 \\ 2\,\mathrm{Re}(z_2 z_1^*) \\ 2\,\mathrm{Im}(z_2 z_1^*) \end{bmatrix} = \begin{bmatrix} s_0 \\ s_0 \cos 2\tau \cos 2\phi \\ s_0 \cos 2\tau \sin 2\phi \\ s_0 \sin 2\tau \end{bmatrix}$$

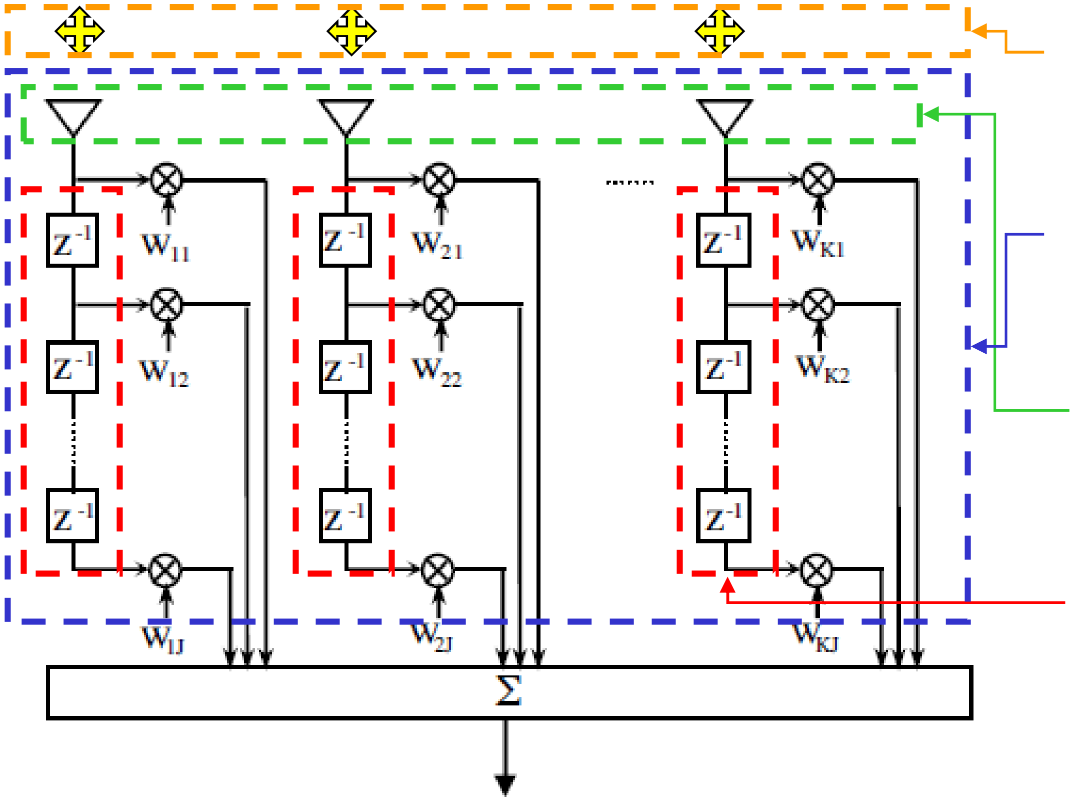
$$\phi = \frac{\arctan\left(\frac{s_2}{s_1}\right)}{2}$$

$$\tau = \frac{\arctan\left(\frac{s_3}{\sqrt{s_1^2 + s_2^2}}\right)}{2}$$



THALES

Radar Processing based on Covariance Matrix

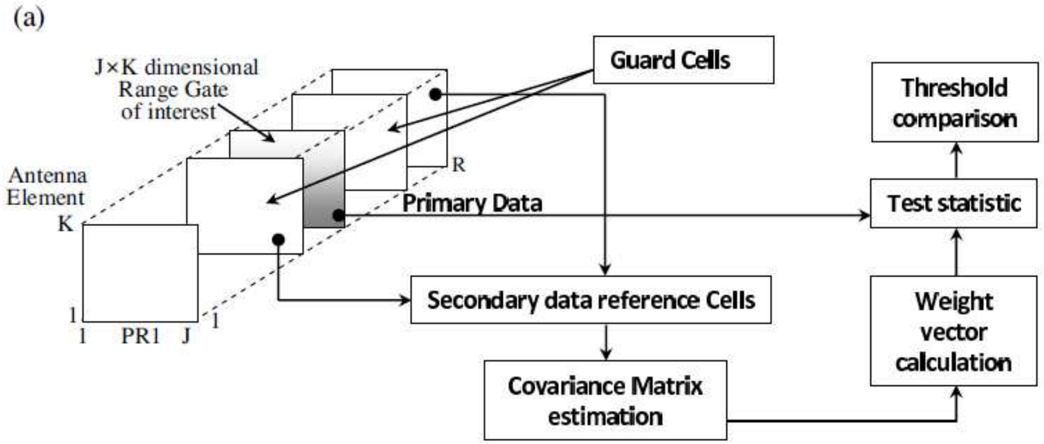


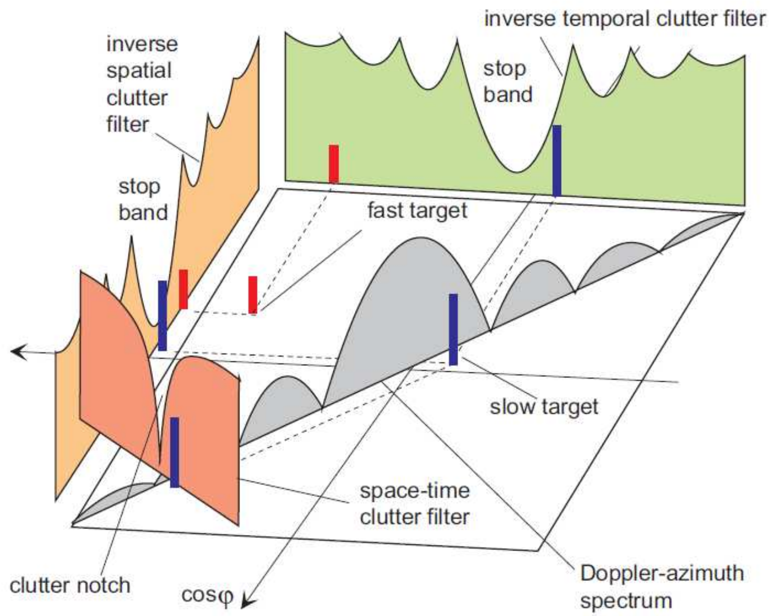
Polarimetric Radar Processing (Polarimetry Covariance Matrix)

STAP Processing (Space-Time Covariance Matrix)

Antenna Processing (Space Covariance Matrix)

Doppler Processing (Time Covariance Matrix)





Military Air Defense Application

Detection of slow/stealth targets in inhomogeneous Clutter

- **New requirements in Air Defense to detect low altitude or surface targets at low elevation. Target Doppler is very close to fluctuating Clutter Doppler (Ground & Sea Clutters)**
- **Detection of asymmetric & stealth targets in Ground Clutter**
  - Microlight Airplane
  - General Aviation
  - UAV & Micro UAV
  - Micro Helicopter
- **Detection of small targets in Sea Clutter**
  - Wooden & inflatable canoe
  - Jetski
  - Unmanned Boat
  - Naval Micro Helicopter
  - Periscope



Detection of low RCS targets
Detection of teneous Doppler Signal
Increase Range & Reactivity

Monitoring of turbulences: New Requirement in ATC (SESAR)

- **In Turbulences, signal is no longer characterized by Doppler velocity Mean but by Doppler Spectrum Width & "shape"**

- **Atmospheric Air turbulences**
  - Eddy Dissipation Rate
  - Turbulent Kinetic Energy
- **Airplane Wake-Vortex turbulences (A380, B747-8)**
  - Circulation
  - Decay Rate

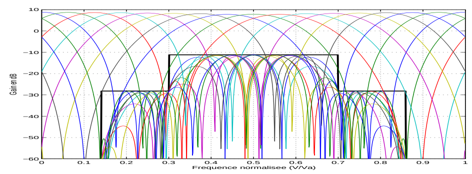
- **Windshear in Final Approach**
  - Headwind
  - Crosswind

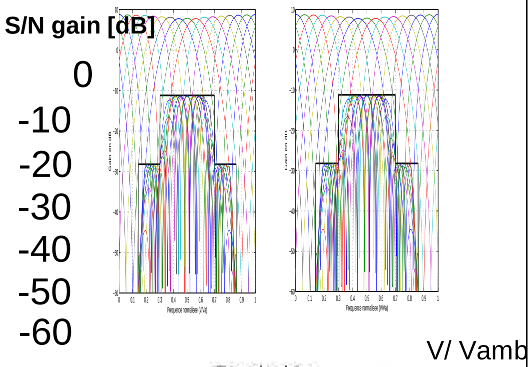
Improve Safety by mitigating weather hazards
Increase Capacity by reducing safe separations

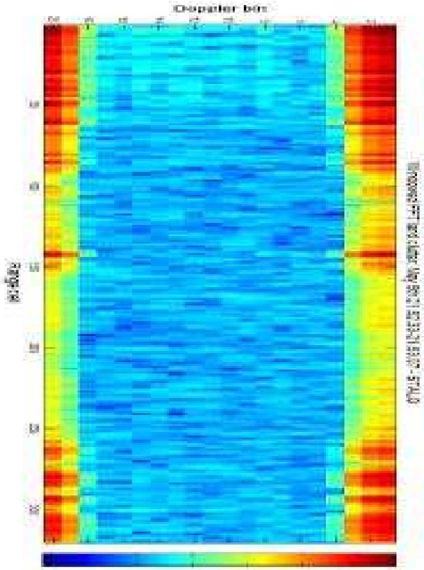
Challenges of Doppler Radar Processing: Detection on Ground

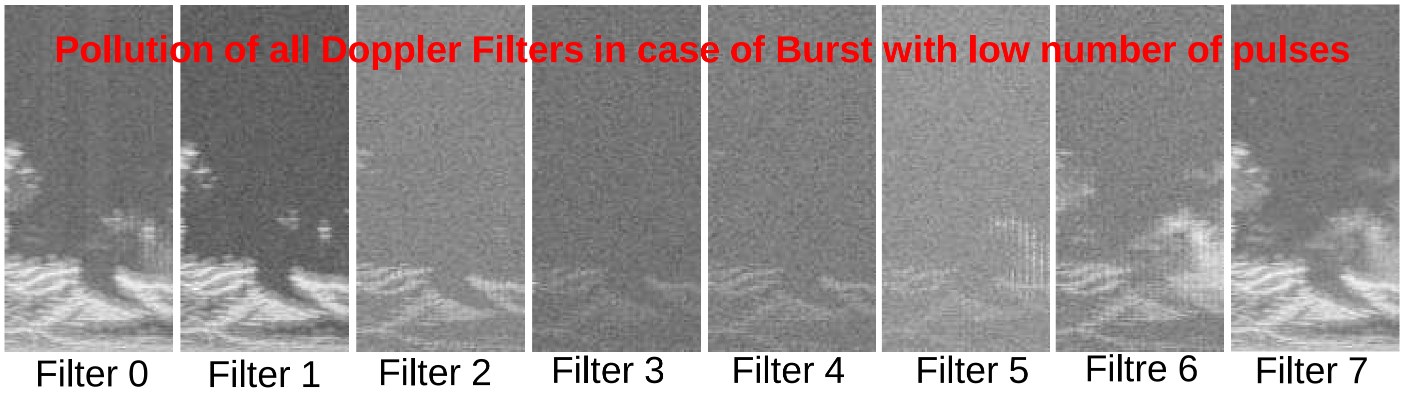
Detection of asymmetric & stealth targets in Inhomogenesous Ground Clutter

- **Classical Doppler Filter Banks (or FFT) are not efficient with very short bursts (<16 pulses) :**
  - Low Resolution of Doppler Filters with short Bursts (Low sidelobes / high loss, wide filter)
  - If Target Doppler is between two Doppler filters, energy is spread on adjacent filters. Gain between 2 filters is lower than gain at filter center ("Straddling loss")
  - Ground Clutter Energy is not limited to zero-Doppler filter but pollution is spread over all filters due to poor Filter-Banks Resolution & Doppler side lobes in case of very short Bursts.







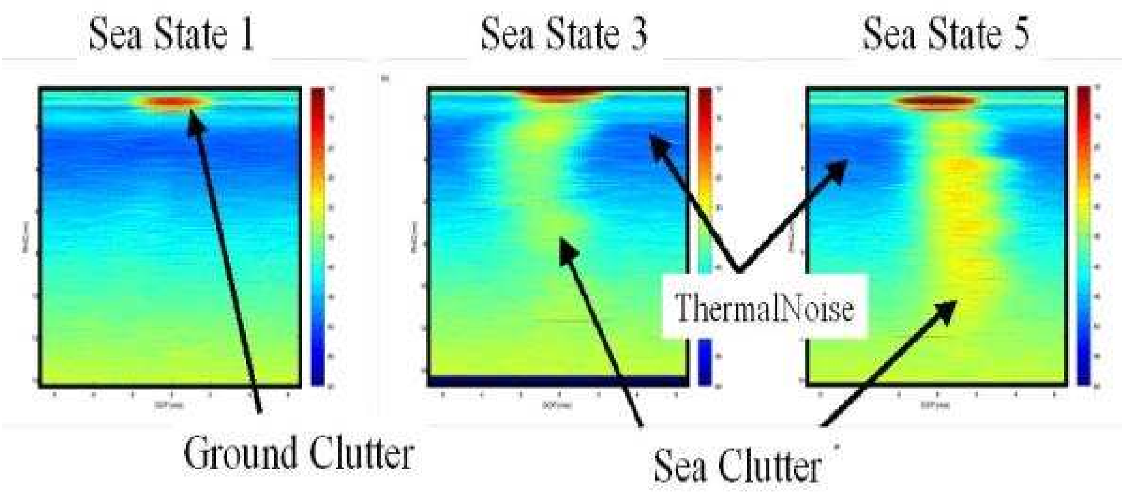


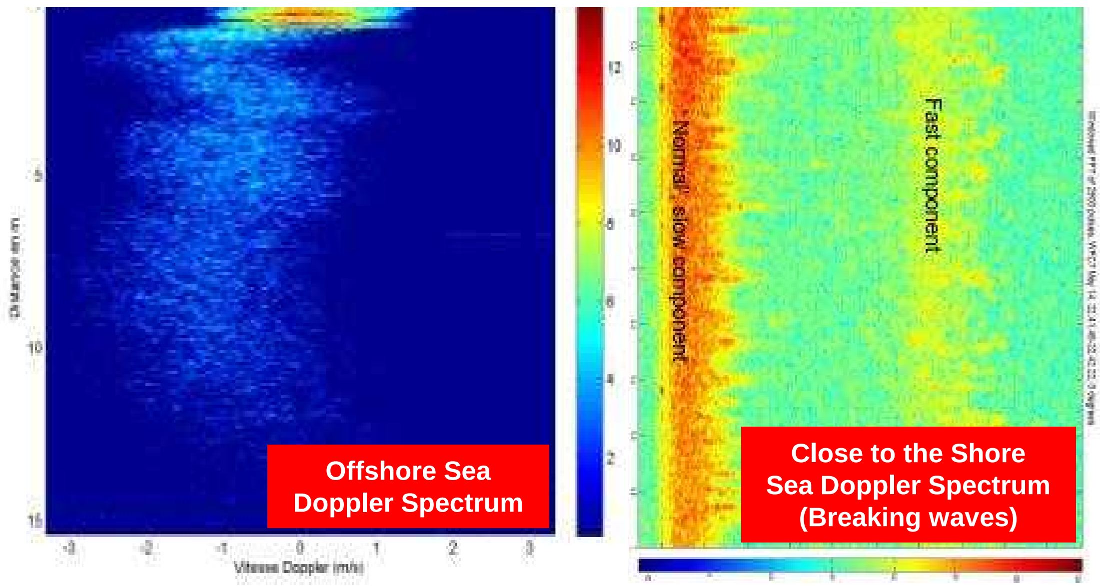
THALES

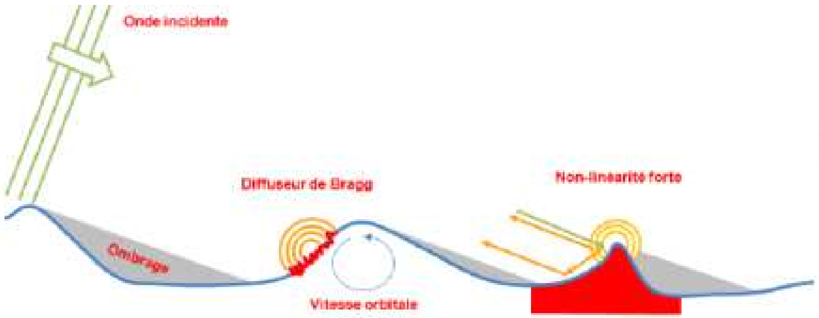
Challenges of Doppler Radar Processing:Detection in Sea

Detection of slow targets in Sea Clutter

- **Sea Clutter is highly inhomogeneous**
  - Doppler fluctuation
  - Time/space Fluctuation
- **Sea Clutter is dependant of**
  - Sea current
  - Surface wind
  - fetch
  - Bathymetry
- **Sea Clutter is corrupted by**
  - Spikes due to breaking waves
  - "Moutonement"





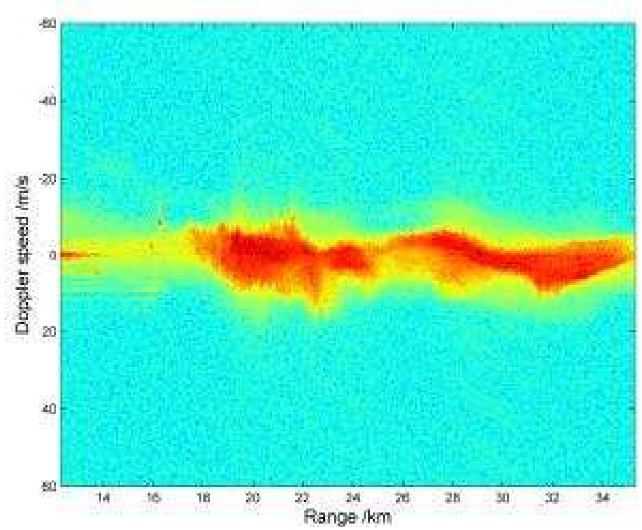


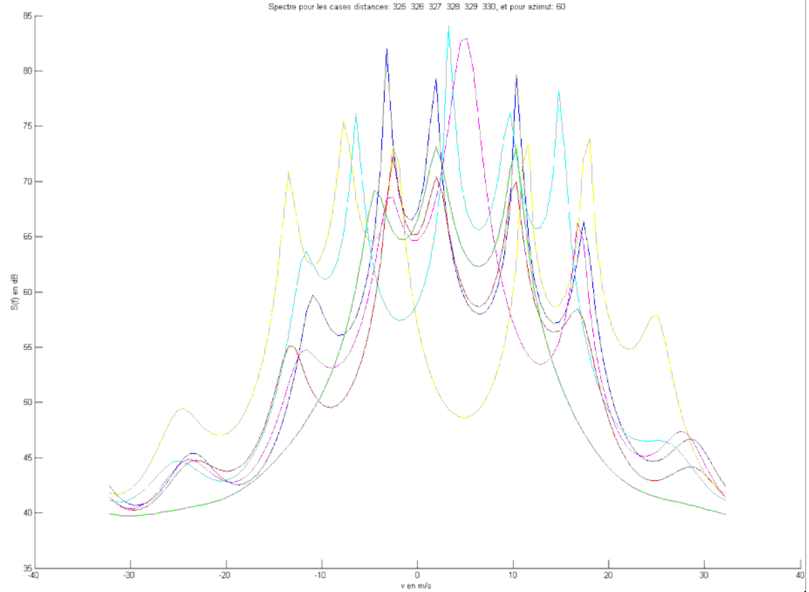
THALES

Challenges of Doppler Radar Processing: Atm. turbulences

Monitoring of atmospheric turbulences

- **Air turbulence is characterized by Spread of fluctuating speeds**
- **This composition of different Doppler speeds in Radar cells generates a Widen Doppler Spectrum**
- **Speed variance of Doppler Spectrum Width are related to 2 measures of turbulence :**
  - EDR: Eddy Dissipation Rate
  - TBE: Turbulent Kinetic Energy





THALES

Challenges of Doppler Radar Processing: Wake Turbulences

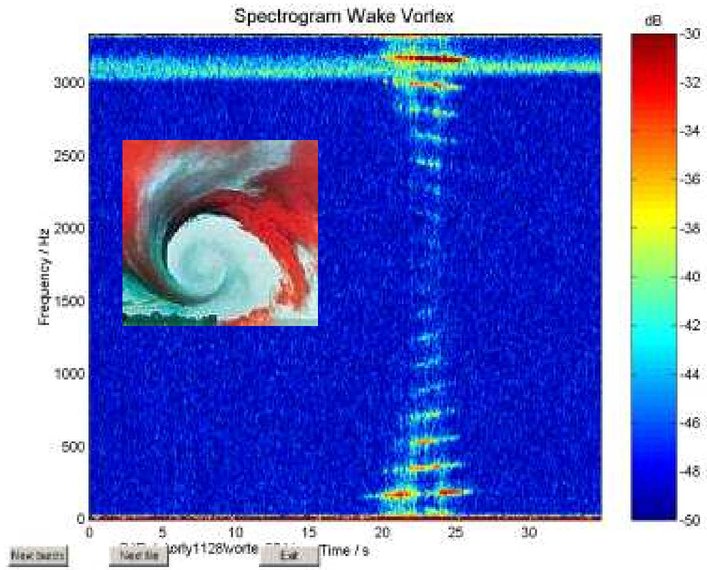
Monitoring of Wake-Vortex Turbulences

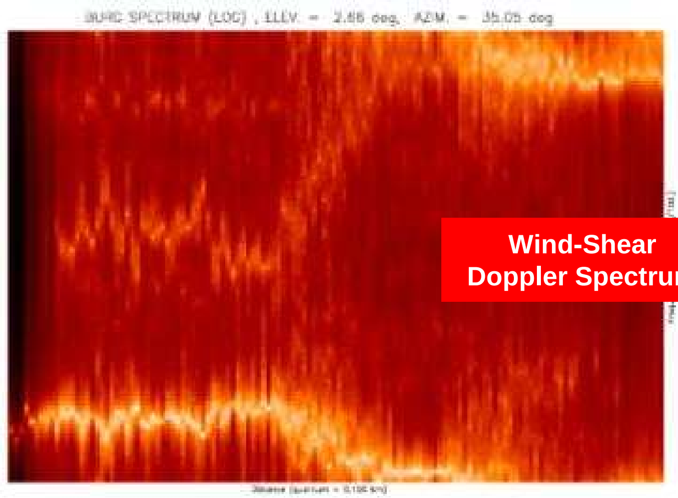
- **Wake vortex generate to contra-rotative roll-up spirals**
- **Mean speed depends on cross-wind**
- **Wake-Vortex has spiral geometry with increasing speed in the core and decreasing speed outside the core**
- **Wake-Vortex Speed and structure depend on Wake-Vortex age/decay phase**
- **Wake-Vortex Strength is characterized by Circulation in m²/s**

Monitoring of Windshear

- **Inversion of speed in range or in altitude**
- **Microburst in the same radar cell**

Wake-Vortex Doppler Spectrum





Wind-Shear Doppler Spectrum

THALES

OS-HD-Doppler-CFAR

For detection of Slow & Stealth/Small Target in inhomogeneous clutter, we need simultaneously :

- **High Doppler Resolution with short Bursts**
- **Robust CFAR in inhomogeneous clutter & closely separated targets**

Proposed Solution : OS-HR-Doppler-CFAR

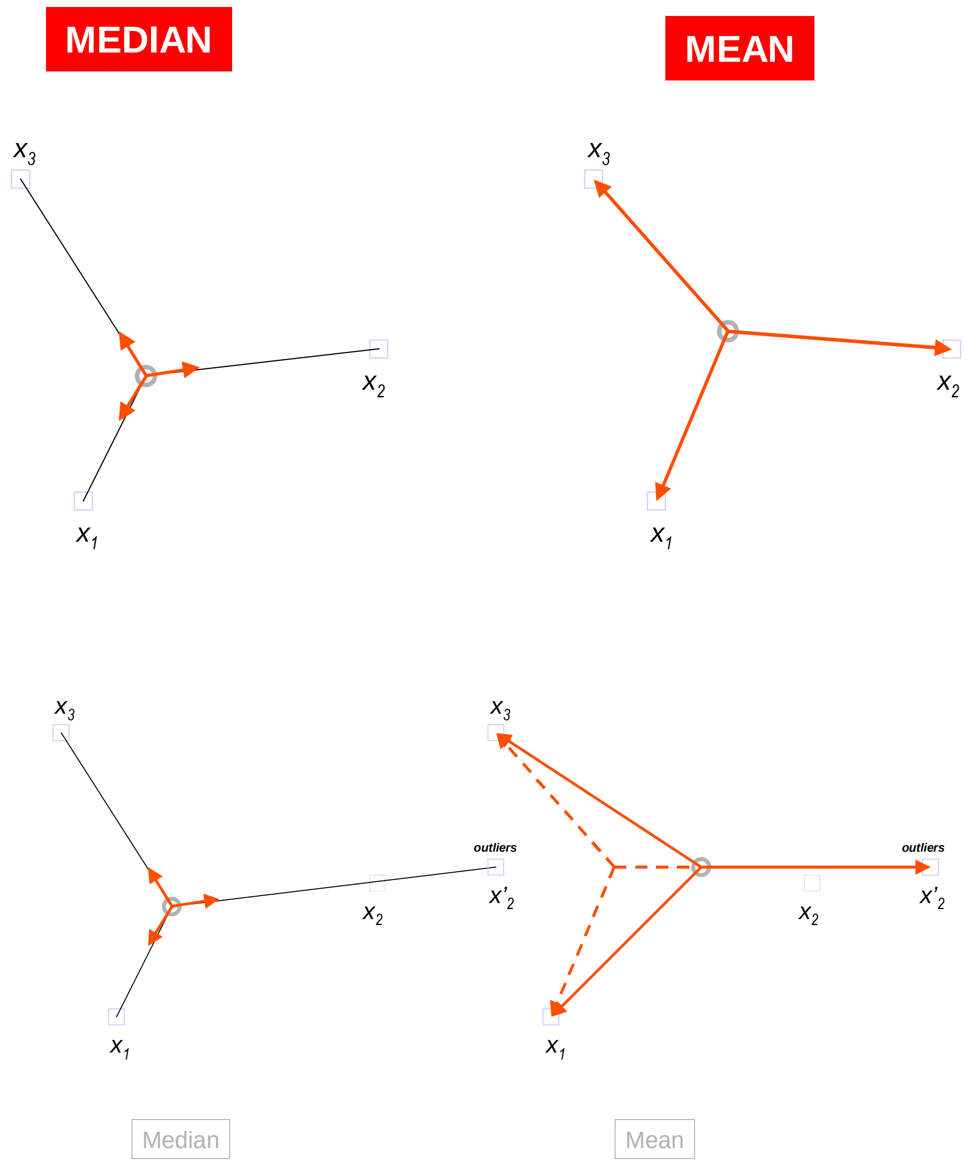
- **Avoid drawbacks of Doppler Filters / FFT in case of short bursts**
- **Take advantages of Robust Ordered Statistic of OS-CFAR (Ordered Statistic CFAR, Median-CFAR)**

Challenges to define OS-HR-Doppler-CFAR :

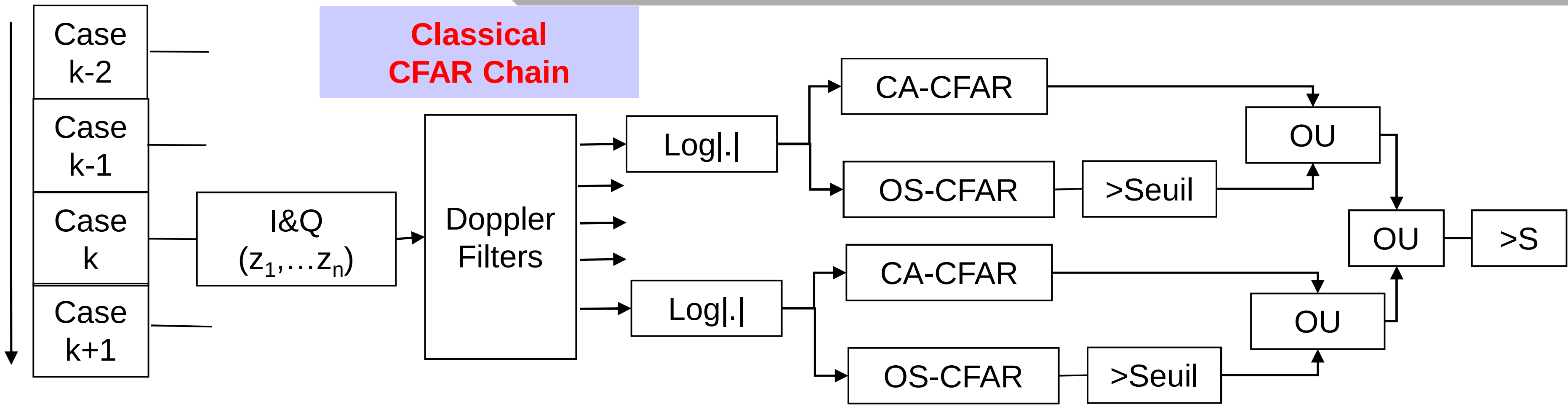
- **Can we order « Doppler Spectrums » :** NO
  - there is no total order of covariance matrices $R_1>R_2>\ldots>R_n$
  - There is only Partial « Lowner » order : $R_1>R_2$ if and only if R1-R2 Positive Definite
- **Can we define « Median » of « Doppler Spectrums » :** YES !!!
  - In a « Metric Space », the median is defined as the point that minimizes the « geodesic » distance to each point (compared to the mean that minimizes the square distance to each point)
  - We can define a deterministic or stochastic gradient flow that converges fastly to « median spectrum » (Modified Karcher Flow : **THALES Patent**)

THALES

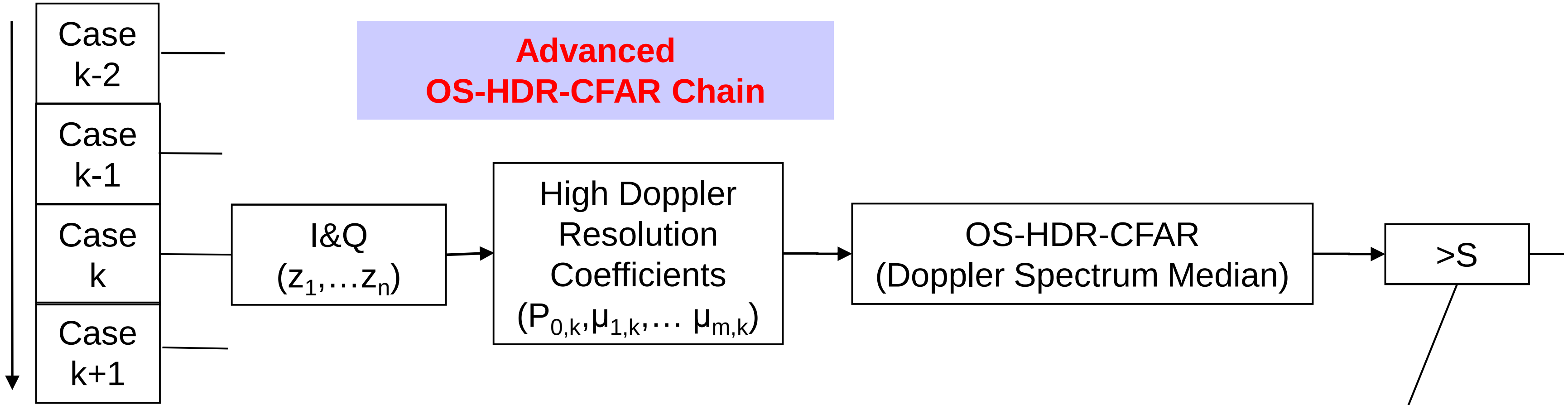
Sensitivity to outliers : Median versus Mean



Classical CFAR Chain versus OS-HDR-CFAR



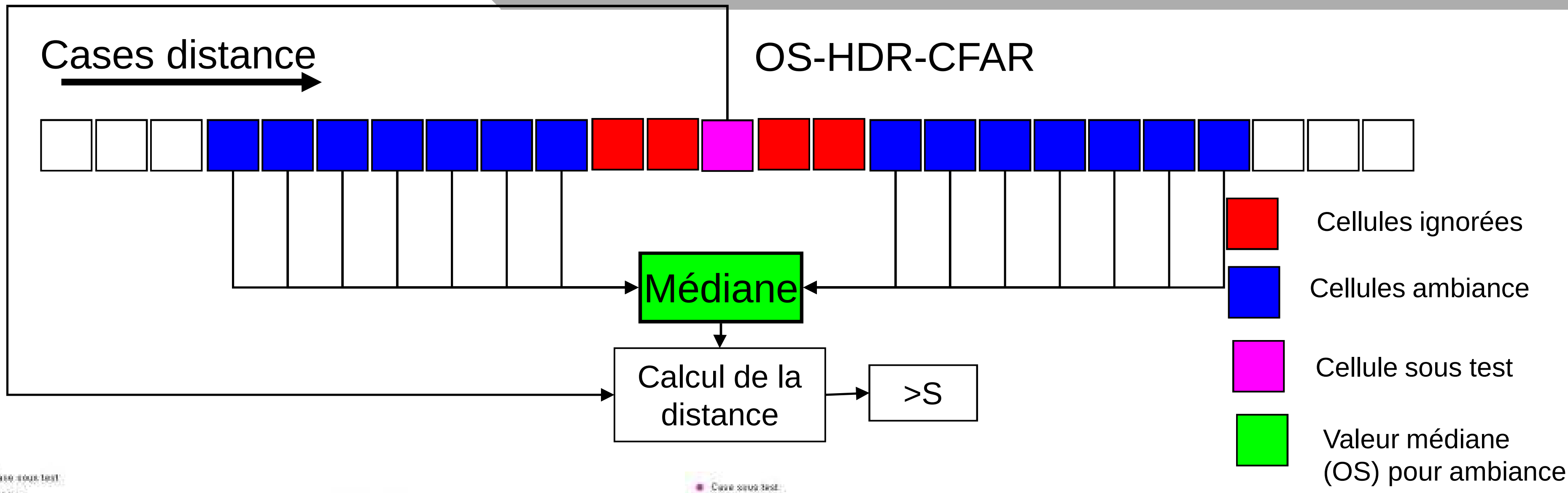
Cases distances

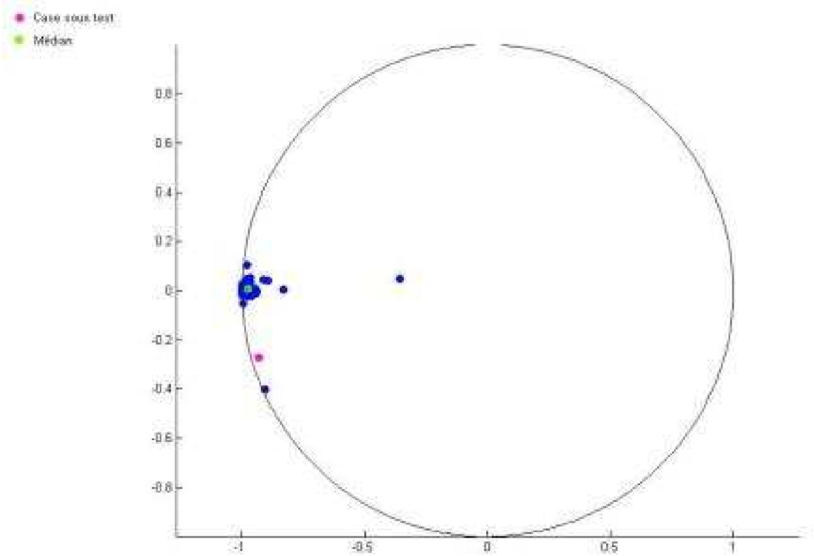


$$dist\left((P_{0,k},\mu_{1,k},\ldots,\mu_{m,k}),(P_{0,median},\mu_{1,median},\ldots,\mu_{m,median})\right)=\sqrt{m^2\ln\left(\frac{P_{0,median}}{P_{0,k}}\right)^2+\sum_{n=1}^{m}\left[(m-n)\,\mathrm{arg\,th}\left(\left|\frac{\mu_{n,k}-\mu_{n,median}}{1-\mu_{n,k}\mu^*_{n,median}}\right|\right)\right]^2}$$

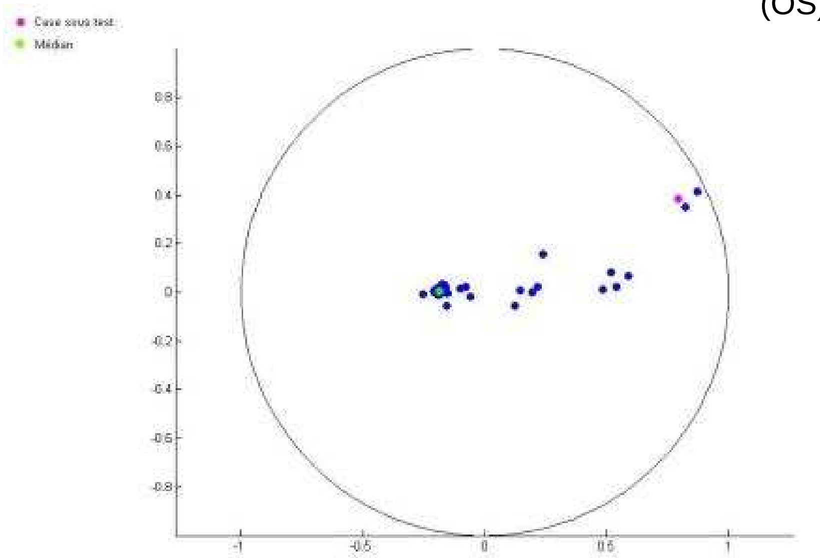
THALES

OS-HDR-CFAR on reflexion coefficients





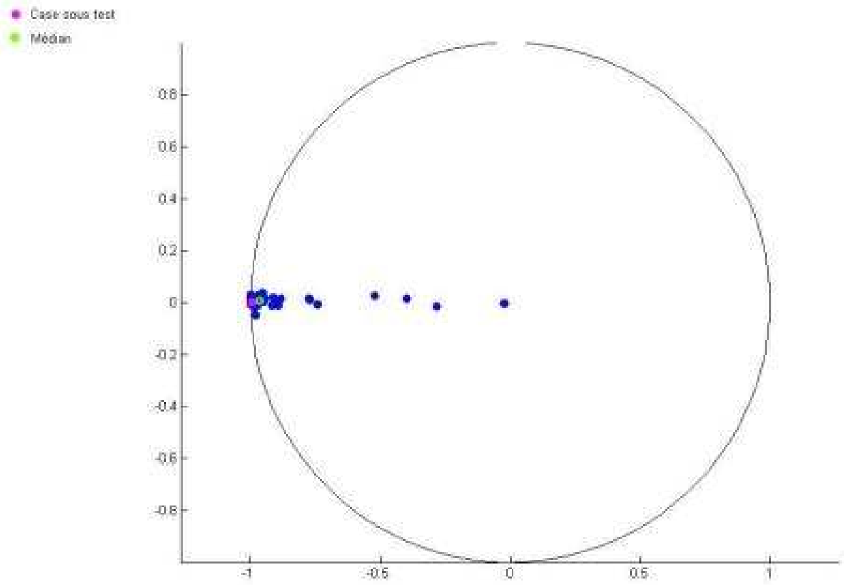
μ_1



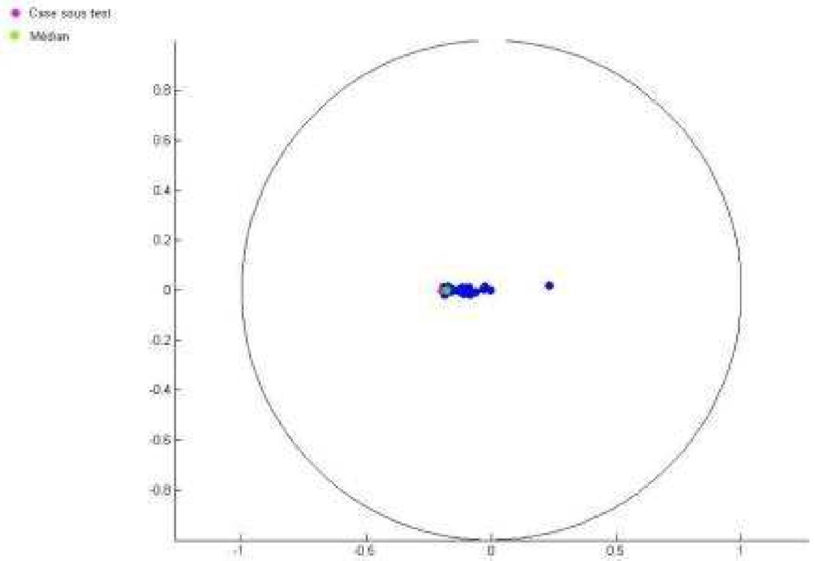
μ_2

THALES

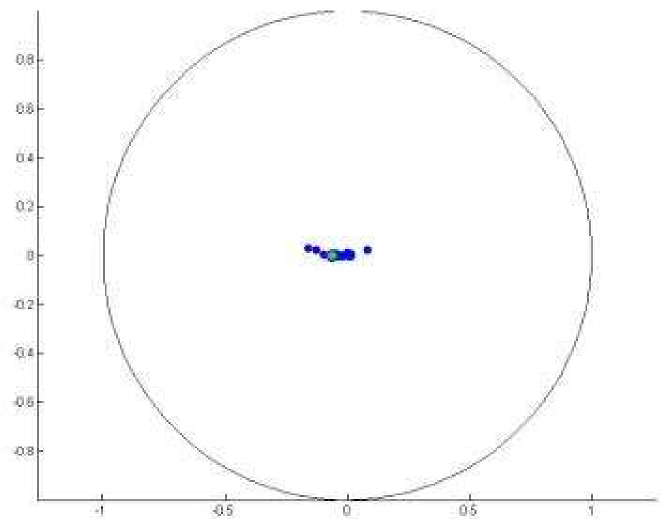
Reflection coefficients of only Ground Clutter



μ_1



μ_2

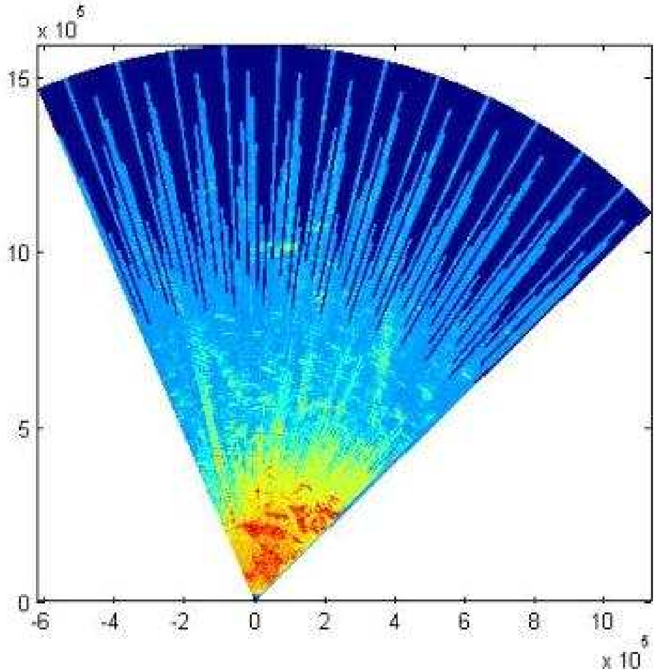


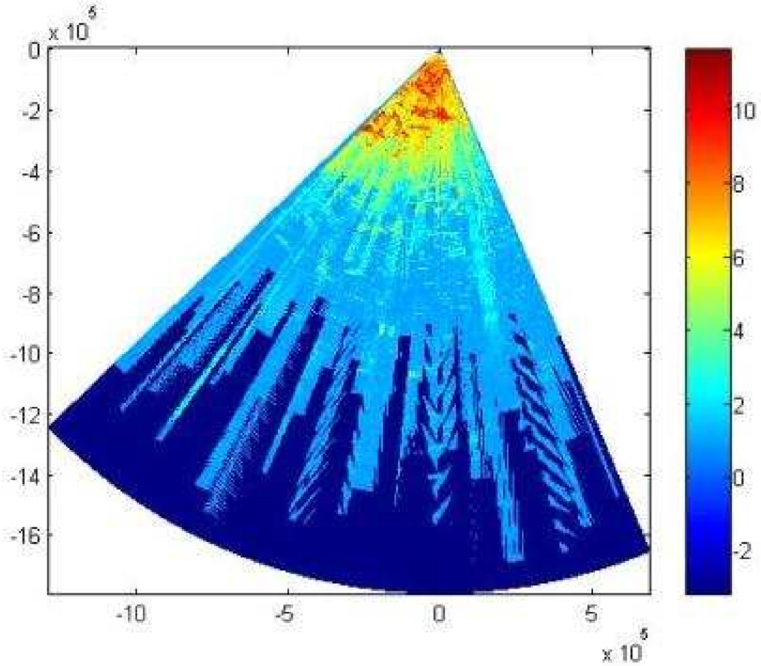
μ_3

Real Data Records of Ground Clutter

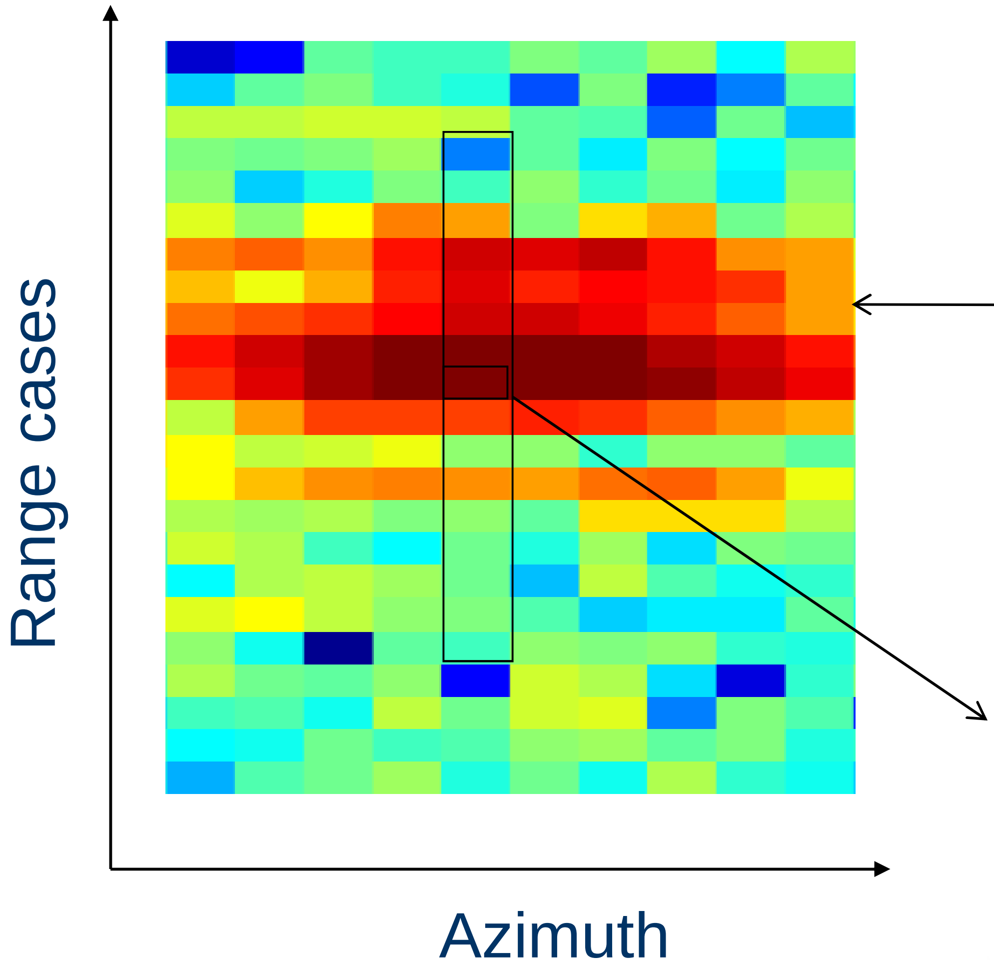
I&Q Data Records of Ground Clutter

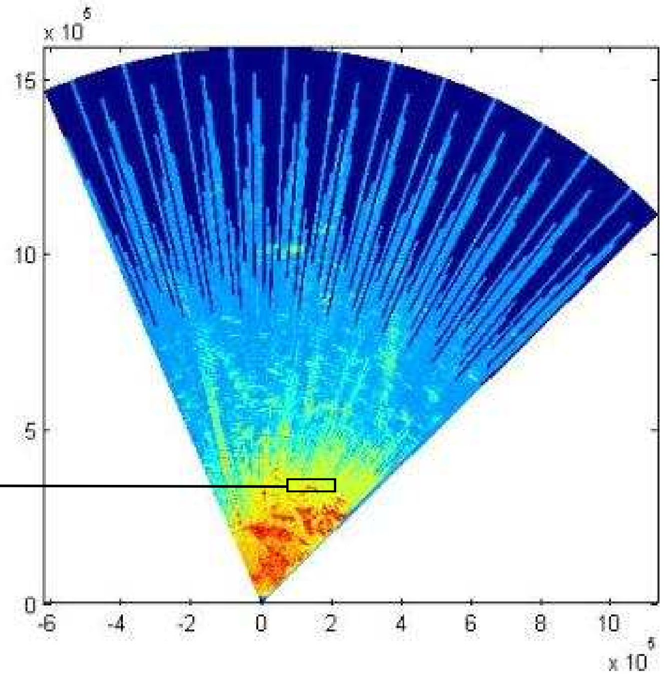
- 12 and 10 scans





Context of the detection problem





- **For each space cell, the signal is formed by d pulses** $\boldsymbol{z} = (z_1, \dots, z_d)$

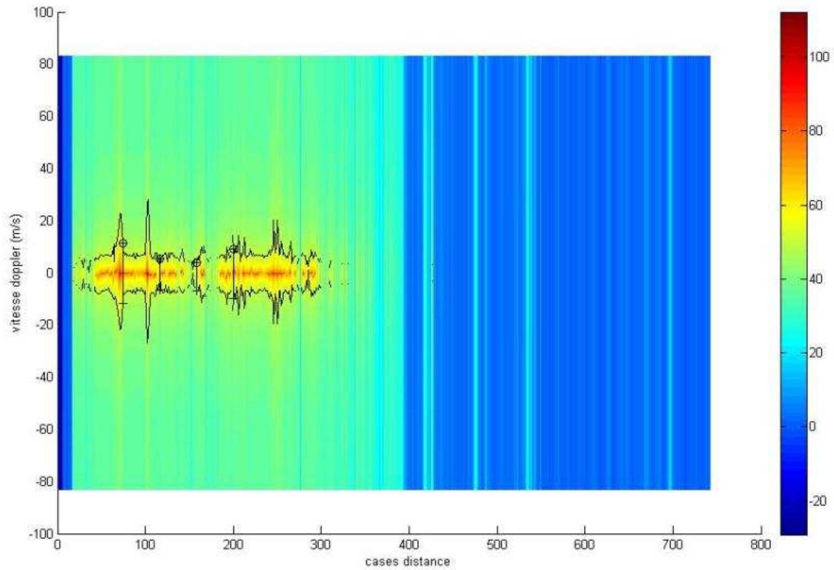
$$\begin{cases} (H_0) & \boldsymbol{z} = \boldsymbol{c} \quad : \text{clutter} \\ (H_1) & \boldsymbol{z} = \boldsymbol{c} + \boldsymbol{s} \quad : \text{clutter} + \text{target} \end{cases}$$

THALES

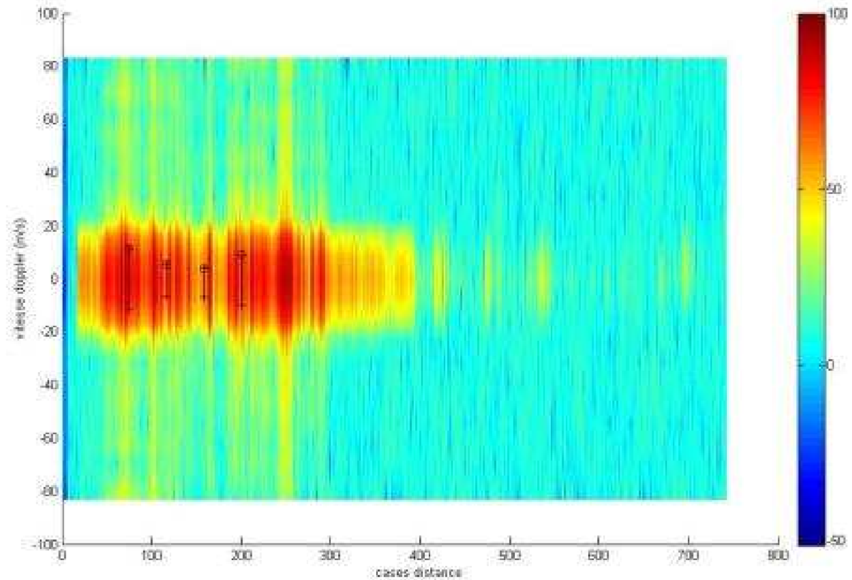
OS-HDR-CFAR Performance Analysis Methodology

Insertion of synthetic targets in real ground clutter

- Normalized Doppler/speed distance from target to clutter : α*Variance with Variance = Capon Spectrum Interval > X dB
- Power of inserted targets : P_{target}



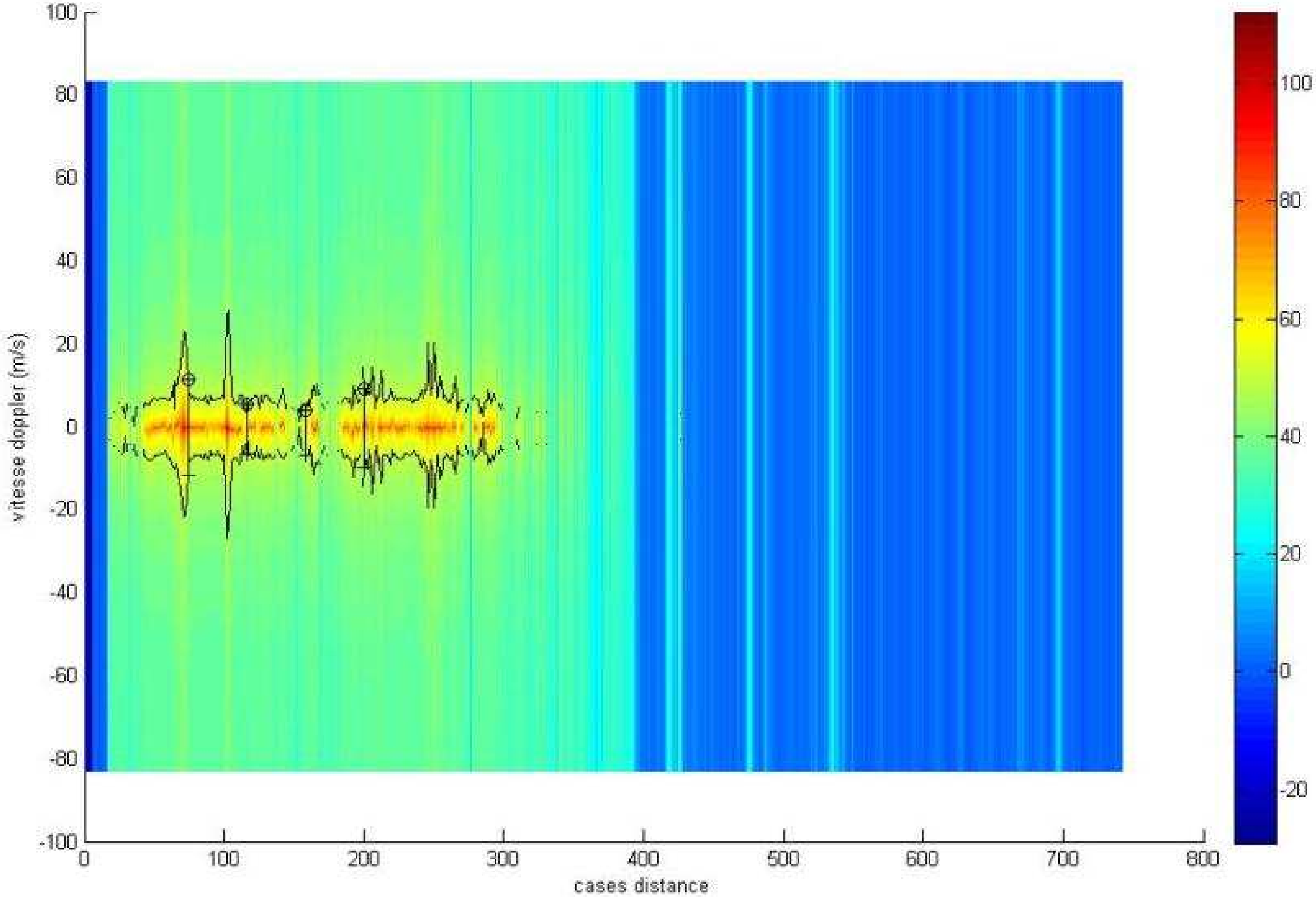
Capon Spectrum before target insertion (α=1; SNR=13dB)



Capon Spectrum before target insertion

THALES

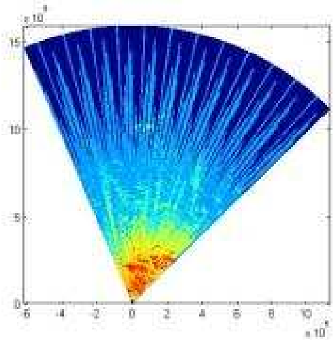
Methodology

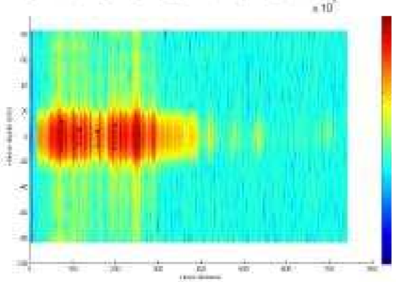


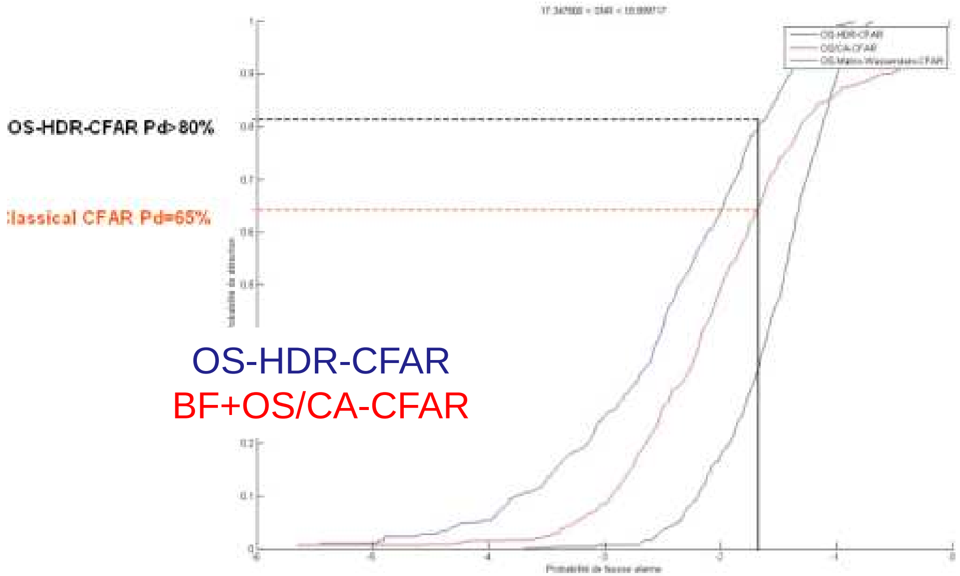
Capon Spectrum insertion before target insertion (α=1; SNR=13dB)

Tests on real recorded I&Q data of Ground Clutter

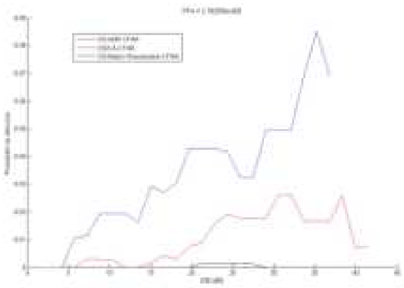
Low Elevation Beam Recording



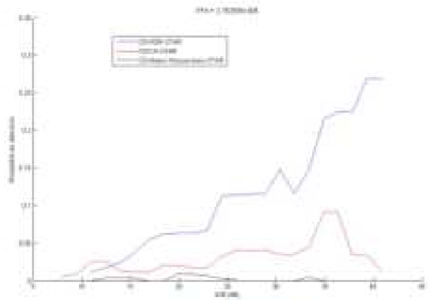




OS-HDR-CFAR
BF+OS/CA-CFAR

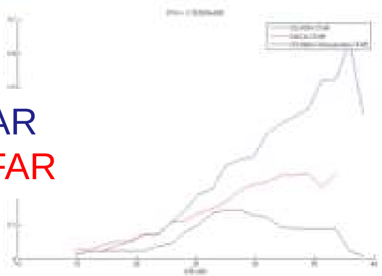


$\alpha = 0.25$

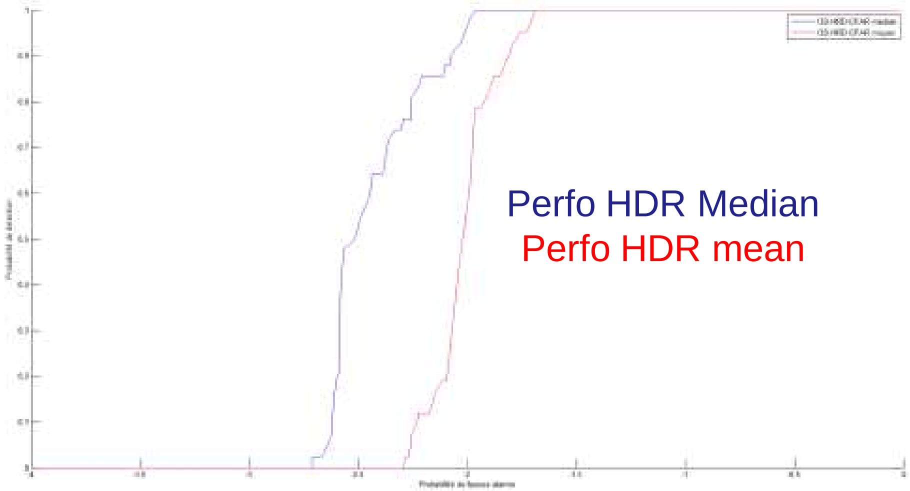


$\alpha = 0.5$

OS-HDR-CFAR
BF+OS/CA-CFAR

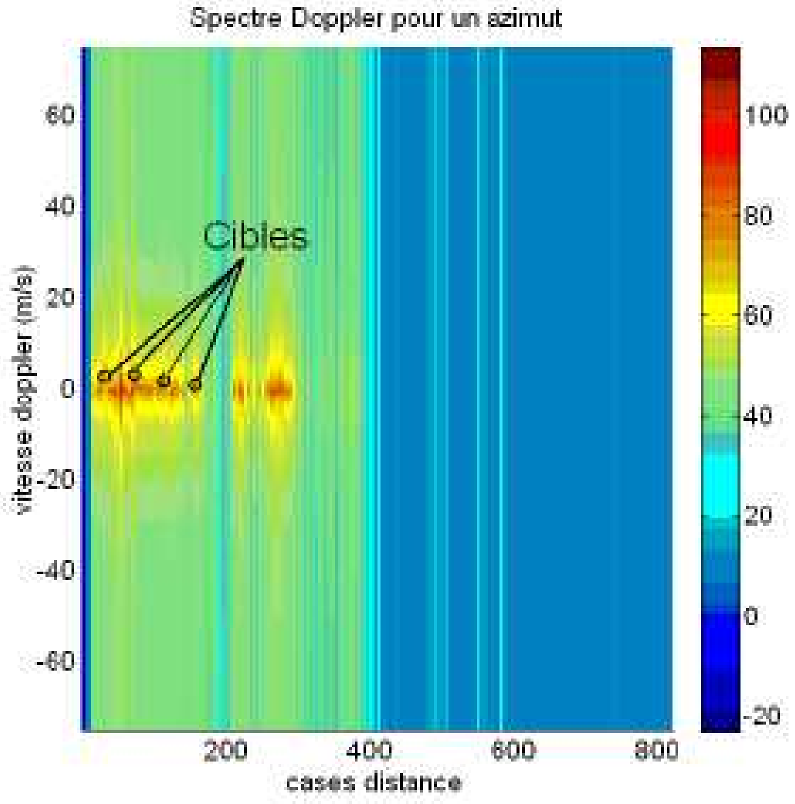


$\alpha = 1$

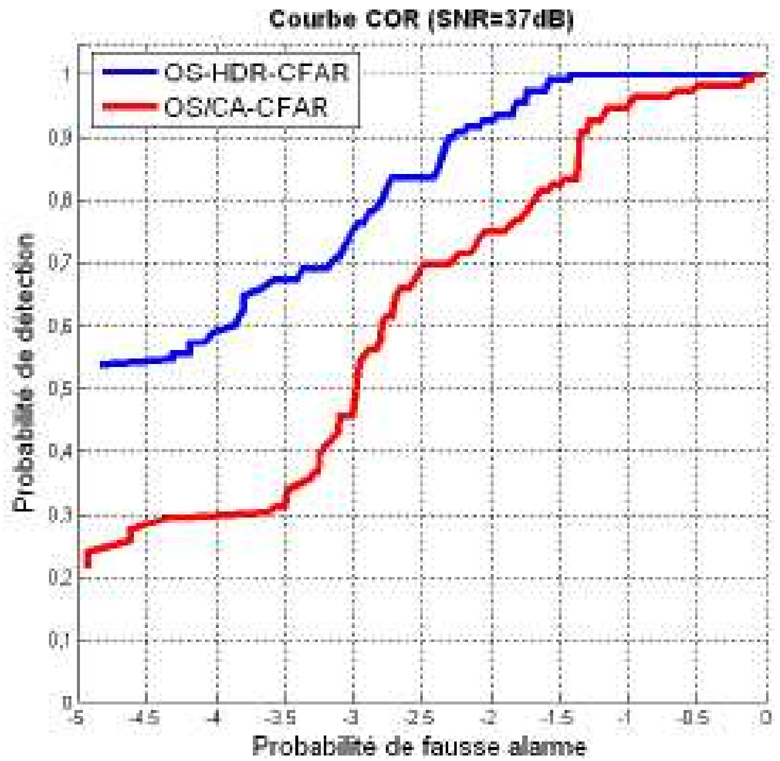


Perfo HDR Median
Perfo HDR mean

COR Curve

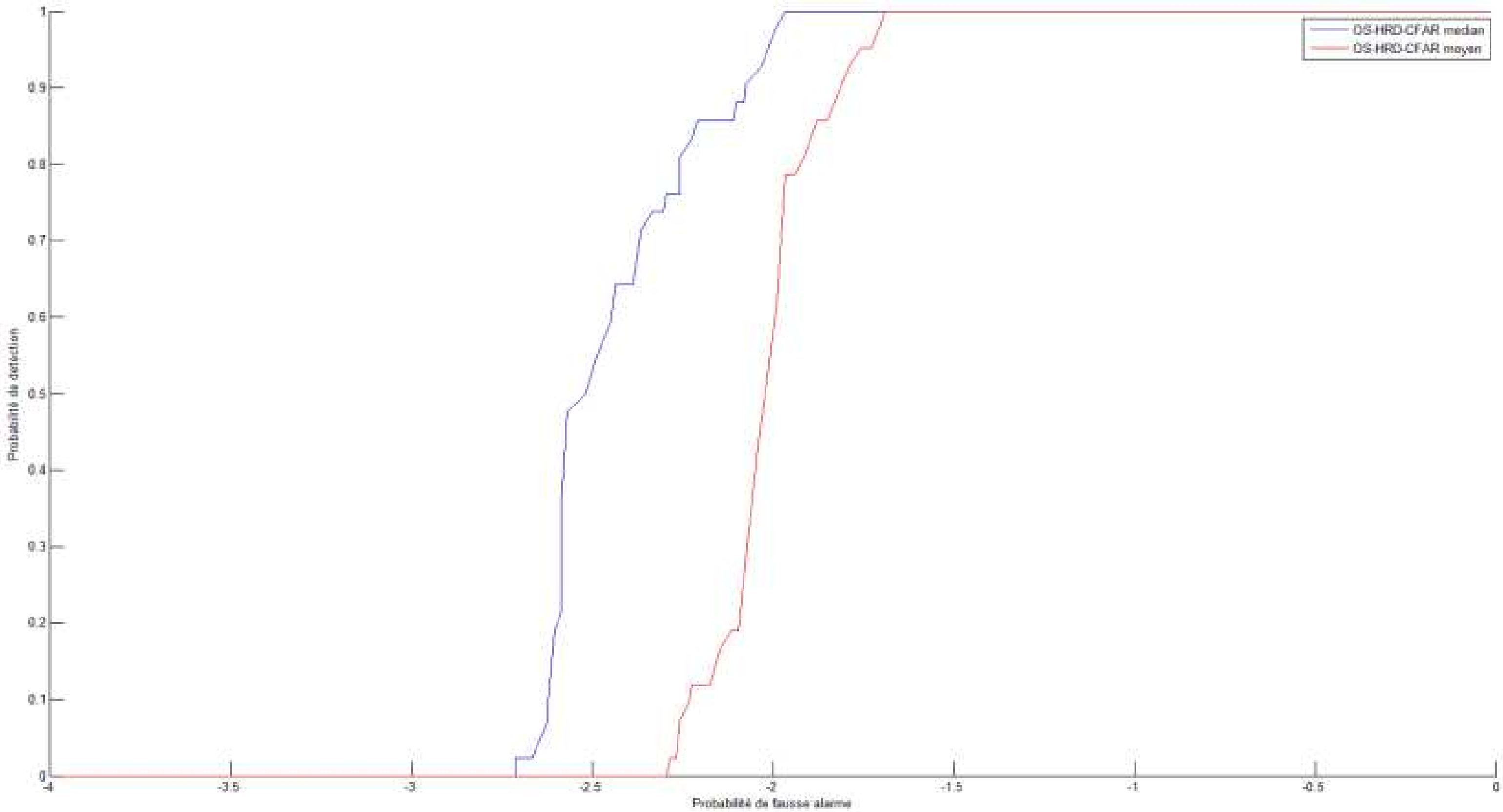


Spectre Doppler pour un azimut
Cibles
vitesse doppler (m/s)
60
40
20
0
-20
-40
-60
200
400
600
800
cases distance
100
80
60
40
20
0
-20

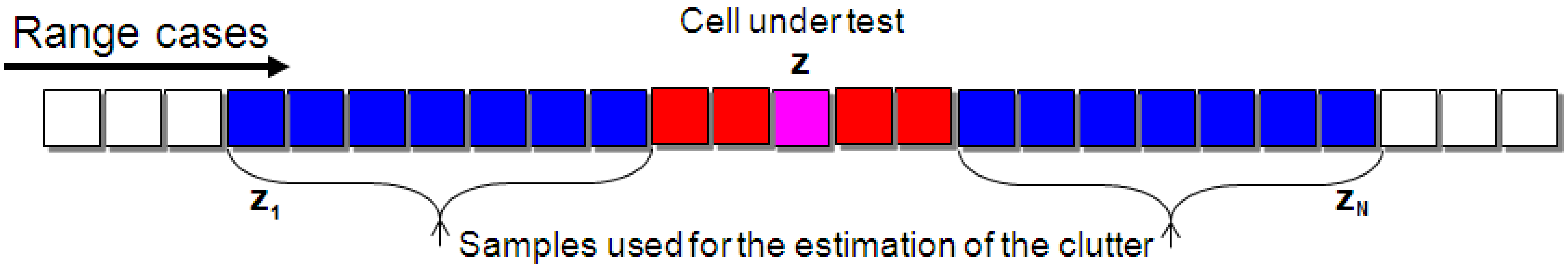


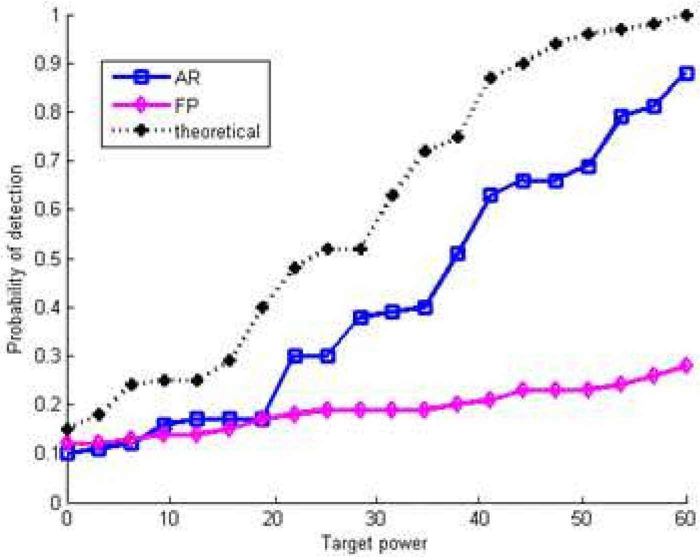
Courbe COR (SNR=37dB)
OS-HDR-CFAR
OS/CA-CFAR
Probabilité de détection
Probabilité de fausse alarme
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-5
-4.5
-4
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0

Comparison Frechet Mean / Frechet Median



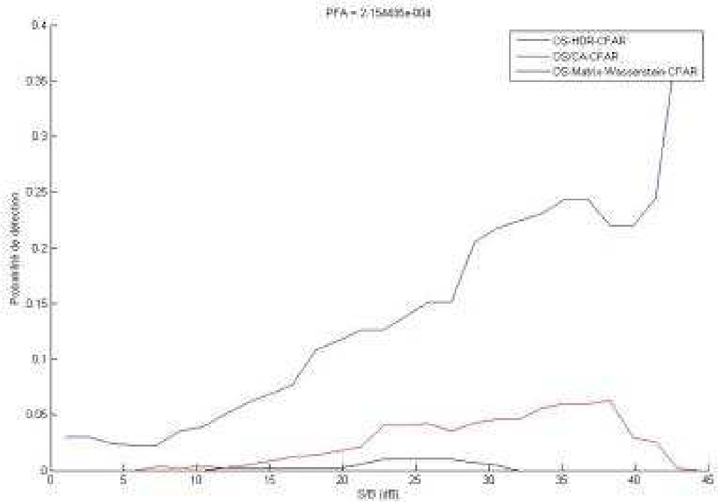
Robustness to multi-target case for a Weibull texture with a shape parameter θ=0.1



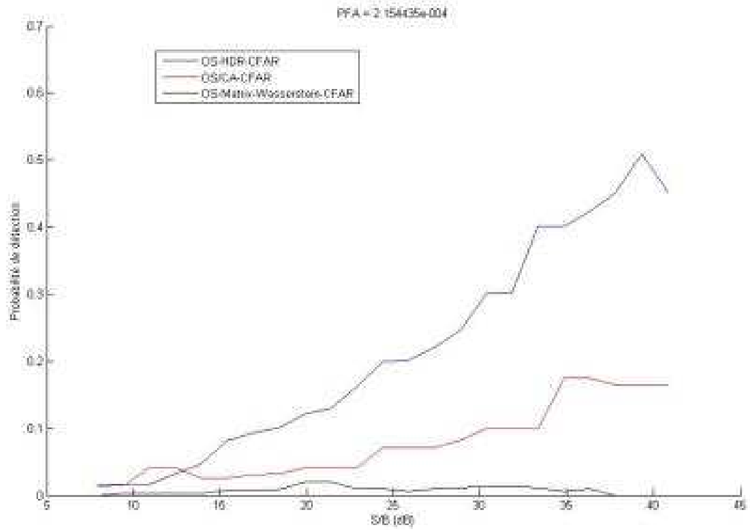


30 target samples in the reference cells

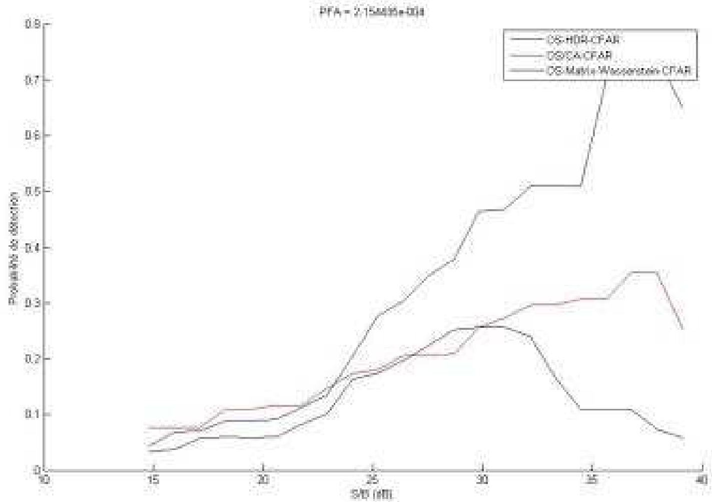
Pd versus SNR for $P_{fa} = 2.1*10^{-4}$



α = 0.25



α = 0.5



α = 1

THALES

Rao's L_1 approach

- Application to the detection of targets : let $\boldsymbol{x} = \left(\left|\sum_{i=1}^{d} e^{-\frac{2i\pi jk}{d}} z_i\right|\right)_{1 \leq j \leq d}$
- For a sample $\mathbf{x}_1,\ldots,\mathbf{x}_N$, we estimate the covariance matrix Γ by

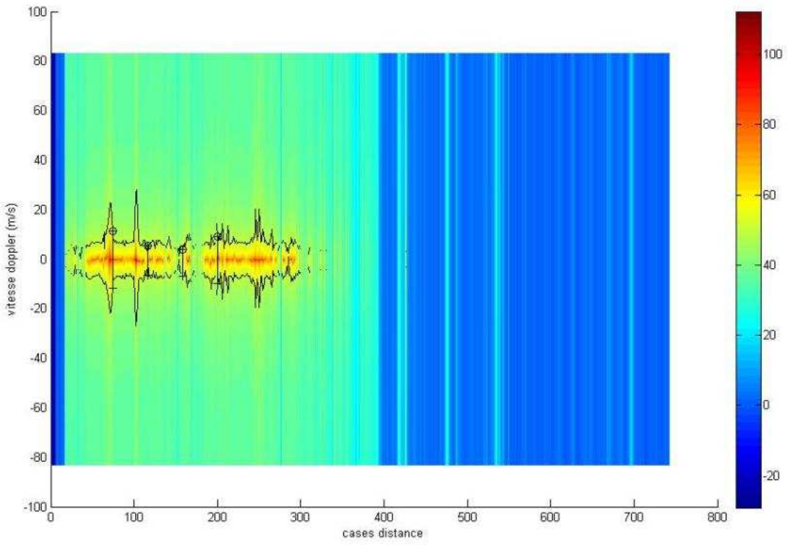
$$\hat{\Gamma} = \begin{pmatrix} \frac{\hat{\gamma}_{11}}{\hat{f}_1^2} & \cdots & \frac{\hat{\gamma}_{1d}}{\hat{f}_1\hat{f}_d} \\ \vdots & \ddots & \vdots \\ \frac{\hat{\gamma}_{d1}}{\hat{f}_d\hat{f}_1} & \cdots & \frac{\hat{\gamma}_{dd}}{\hat{f}_d^2} \end{pmatrix}$$

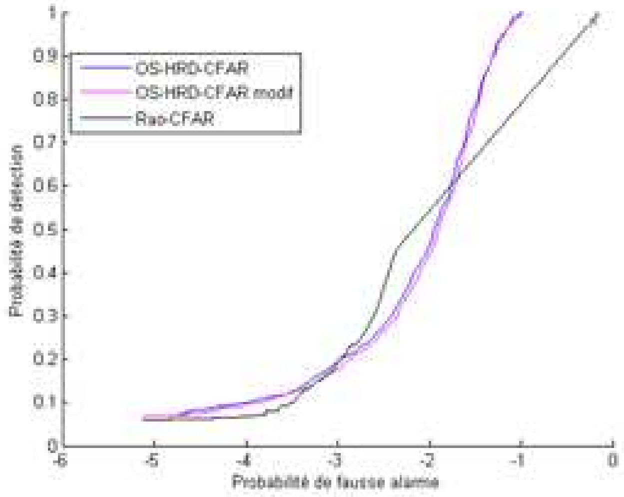
- $\hat{\gamma}_{ij}=\frac{1}{N}\sum_{n=1}^{N} \mathbb{1}(x_i \leq m_i, x_j \leq m_j) - \frac{1}{4}$
- For f_i, we choose a kernel estimator :
  - $\frac{1}{\hat{f}_i} = N^{\delta} \int_0^{\log(N)} \left[\widehat{F_N^{-1}}\left(\frac{1}{2} + uN^{-\delta}\right) - \widehat{F_N^{-1}}\left(\frac{1}{2}\right)\right] e^{-y}(1 + u - 0{,}5u^2)dy$
  - We choose $\delta = \frac{1}{2d-1}$
- The test statistics is inspired from Hotelling's test

$$(\boldsymbol{m} - \boldsymbol{x})^T \hat{\Gamma}^{-1} (\boldsymbol{m} - \boldsymbol{x}) \lessgtr T$$

THALES

- **Result on ground clutter :**





- **The estimator of Γ is non-parametric : we should incorporate the information of stationarity**

www.thalesgroup.com

Karcher Flows for Toeplitz-Block-Toeplitz HPD matrices (TBTHPD)

THALES

Extension to Toeplitz Block Toeplitz Hermitian PD Matrices

- **Previous results can be extented to Block-Toeplitz Matrices :**

$$R_{p,n+1} = \begin{bmatrix} R_0 & R_1 & \cdots & R_n \\ R_1^+ & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1 \\ R_n^+ & \cdots & R_1^+ & R_0 \end{bmatrix} = \begin{bmatrix} R_{p,n} & \tilde{R}_n \\ \tilde{R}_n^+ & R_0 \end{bmatrix}$$

$$\tilde{R}_n = V \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix}^* \quad \text{with} \quad V = \begin{bmatrix} 0 & \cdots & 0 & J_p \\ \vdots & \iddots & \iddots & 0 \\ 0 & J_p & \iddots & \vdots \\ J_p & 0 & \cdots & 0 \end{bmatrix}$$

Extension of Trench/Verblunsky Theorem for TBTHPD Matrices

- **Every Toeplitz-Block-Toeplitz HPD matrix can be parametrized by Matrix Verblunsky Coefficients:**

$$R_{p,n+1}^{-1}=\begin{bmatrix}\alpha_n & \alpha_n.\widehat{A}_n^{+}\\ \alpha_n.\widehat{A}_n & R_{p,n}^{-1}+\alpha_n.\widehat{A}_n.\widehat{A}_n^{+}\end{bmatrix}$$

$$R_{p,n+1}=\begin{bmatrix}\alpha_n^{-1}+\widehat{A}_n^{+}.R_{p,n}.\widehat{A}_n & -\widehat{A}_n^{+}.R_{p,n}\\ -R_{p,n}.\widehat{A}_n & R_{p,n}\end{bmatrix}$$

$$\text{with}\quad \alpha_n^{-1}=\left[1-A_n^nA_n^{n+}\right].\alpha_{n-1}^{-1},\quad \alpha_0^{-1}=R_0$$

$$\text{and}\quad \widehat{A}_n=\begin{bmatrix}A_1^1\\ \vdots\\ A_n^n\end{bmatrix}=\begin{bmatrix}\widehat{A}_{n-1}\\ 0_p\end{bmatrix}+A_n^n.\begin{bmatrix}J_pA_{n-1}^{n-1*}J_p\\ \vdots\\ J_pA_1^{n-1*}J_p\\ I_p\end{bmatrix}$$

- **Extension of Trench/Verblunsky Theorem: Existence of Diffeomorphism φ**

$$\varphi : TBTHPD_{n\times n}\rightarrow THPD_n\times SD^{n-1}$$

$$R\mapsto\left(R_0,A_1^1,\ldots,A_{n-1}^{n-1}\right)$$

$$\text{with}\quad SD=\left\{Z\in Herm(n)\,/\,ZZ^{+}<I_n\right\}$$

THALES

Information Geometry Metric : Entropic Kähler Potential

- **Kähler potential defined by Hessian of multi-channel/Multi-variate entropy :**

$$\tilde{\Phi}(R_{p,n}) = -\log(\det R_{p,n}) + cste = -Tr(\log R_{p,n})\mu + cste$$

$$\Rightarrow g_{i\bar{j}} = Hess[\phi(R_{p,n})]$$

$$\tilde{\Phi}(R_{p,n}) = \sum_{k=1}^{n-1}(n-k).\log\det\left[I_n - A_k^k A_k^{k+}\right] + n.\log[\pi.e.\det R_0]$$

$$ds^2 = n.Tr\left[\left(R_0^{-1}dR_0\right)^2\right] + \sum_{k=1}^{n-1}(n-k)Tr\left[\left(I_n - A_k^k A_k^{k+}\right)^{-1} dA_k^k \left(I_n - A_k^{k+} A_k^k\right)^{-1} dA_k^{k+}\right]$$

We recover the previous metric for THPD matrix !!

Multi-Channel Burg Algorithm: MOPUC

- **Multi-Channel Burg Algorithm :**

$$\left[A_1^n, A_2^n, \ldots, A_n^n\right] = \left[A_1^{n-1}, A_2^{n-1}, \ldots, A_{n-1}^{n-1}, 0\right] + A_n^n\left[JA_{n-1}^{n-1*}J, JA_{n-2}^{n-1*}J, \ldots, JA_1^{n-1*}J, I\right]$$

$$\begin{cases} \varepsilon_n^f(k) = \sum_{l=0}^{n} A_l^n(k) Z(k-l) \\ \varepsilon_n^b(k) = \sum_{l=0}^{n} JA_l^n(k)^* JZ(k-n+l) \end{cases} \quad \text{with } J = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \iddots & 1 & 0 \\ 0 & \iddots & \iddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} \text{ and } A_0^n = I$$

$$\begin{cases} \varepsilon_{n+1}^f(k) = \varepsilon_n^f(k) + A_{n+1}^{n+1}\varepsilon_n^b(k-1) \\ \varepsilon_{n+1}^f(k) = \varepsilon_n^b(k-1) + JA_{n+1}^{n+1*}J\varepsilon_n^f(k) \end{cases} \quad \text{with } \varepsilon_0^f(k) = \varepsilon_0^b(k) = Z(k)$$

$$\underset{A_{n+1}^{n+1}}{Min}\, Tr\left[P_n^f + JP_n^{b*}J\right] \Rightarrow A_{n+1}^{n+1} = -2\left[P_n^{fb} + JP_n^{fbT}J\right]\left[P_n^b + JP_n^{f*}J\right]^{-1}$$

$$A_{n+1}^{n+1} = -2\left[\sum_{k=1}^{N+n}\varepsilon_n^f(k)\varepsilon_n^b(k-1)^+\right]\left[\sum_{k=1}^{N+n}\varepsilon_n^f(k)\varepsilon_n^f(k)^+ + \sum_{k=1}^{N+n}\varepsilon_n^b(k)\varepsilon_n^b(k)^+\right]^{-1}$$

$$\text{with } \begin{cases} P_n^{fb} = E\left[\varepsilon_n^f(k)\varepsilon_n^b(k-1)^+\right] \\ P_n^f = E\left[\varepsilon_n^f(k)\varepsilon_n^f(k)^+\right] \\ P_n^b = E\left[\varepsilon_n^b(k)\varepsilon_n^b(k)^+\right] \end{cases}$$

$$A_{n+1}^{n+1}.\left(A_{n+1}^{n+1}\right)^+ < I_{n+1} \Rightarrow A_{n+1}^{n+1} \in Disk_{Siegel}$$

THALES

Mostow Decomposition (& Berger Fibration)

- **Mostow Theorem :**

Every matrix M of $GL(n,C)$ can be decomposed :

$$M = Ue^{iA}e^{S}$$

where

U is unitary

A is real antisymmetric

S is real symmetric

Can be deduce from

- **Lemma :** Let A and B two positive definite hermitian matrices, there exist a unique positive definite hermitian matrix X such that:

$$XAX = B$$

- **Corollary :** if M is Hermitian Positive Definite, there exist a unique real symmetric matrix S such that :

$$M^{*} = e^{S}M^{-1}e^{S}$$

Mostow Theorem :

All matrix M of $GL(n,C)$ can be decomposed in :

$$M = Ue^{iA}e^{S}$$

U is unitary, A is real antisymmetric réelle and S is real symmetric

$$M = Ue^{iA}e^{S} \Rightarrow P = M^{+}M = e^{S}e^{2iA}e^{S}$$

Proof :

$$\Rightarrow P^{*} = e^{S}e^{-2iA}e^{S} = e^{2S}\left(e^{-S}e^{-2iA}e^{-S}\right)e^{2S}$$

$$\Rightarrow P^{*} = e^{2S}P^{-1}e^{2S}$$

$$\text{Lemma} + \text{corrolary}: P^{*} = e^{2S}P^{-1}e^{2S} \Rightarrow e^{2S} = P^{1/2}\left(P^{-1/2}P^{*}P^{-1/2}\right)^{1/2}P^{1/2}$$

$$\Rightarrow S = \frac{1}{2}.\log\left(P^{1/2}\left(P^{-1/2}P^{*}P^{-1/2}\right)^{1/2}P^{1/2}\right) \quad \text{with} \quad P = M^{+}M$$

$$\text{exponentielle injectivity}: e^{2iA} = e^{-S}Pe^{-S}$$

$$\Rightarrow A = \frac{1}{2i}\log\left(e^{-S}Pe^{-S}\right) \quad \text{with} \quad P = M^{+}M$$

$$U = Me^{-S}e^{-iA}$$

Siegel Disc Automorphism

- Automorphism of Siegel Disc SD_n given by :

$$\Sigma = \Phi_{Z_0}(Z) = \left(I - Z_0 Z_0^+\right)^{-1/2}(Z - Z_0)\left(I - Z_0^+ Z\right)^{-1}\left(I - Z_0^+ Z_0\right)^{1/2}$$

- All automorphisms given by :

$$\forall \Psi \in Aut(SD_n), \exists U \in U(n,C) / \Psi(Z) = U \Phi_{Z_0}(Z) U^t$$

- Distance given by:

$$\forall Z, W \in SD_n, d(Z,W) = \frac{1}{2}\log\left(\frac{1 + \left\|\Phi_Z(W)\right\|}{1 - \left\|\Phi_Z(W)\right\|}\right)$$

- Inverse automorphism given by :

$$G = \left(I - Z_0 Z_0^+\right)^{1/2} \Sigma \left(I - Z_0^+ Z_0\right)^{-1/2} = (Z - Z_0)\left(I - Z_0^+ Z\right)^{-1}$$

$$\Rightarrow \begin{cases} Z = \Phi_{Z_0}^{-1}(\Sigma) = (G Z_0^+ + I)^{-1}\left(G + Z_0^+\right) \\ \text{with } G = \left(I - Z_0 Z_0^+\right)^{1/2} \Sigma \left(I - Z_0^+ Z_0\right)^{-1/2} \end{cases}$$

THALES

Iterated Computation of Median in Siegel disk

$$Initialisation : W_{median,0} = 0 \quad and \quad \{W_{1,0},...,W_{m,0}\} = \{W_1,...,W_m\}$$

$$\text{Iterate on } n \text{ until } \|G_n\|_F < \varepsilon$$

$$W_{k,n} = U_{k,n} e^{iA_{k,n}} e^{S_{k,n}} \Rightarrow H_{k,n} = U_{k,n} e^{iA_{k,n}} = W_{k,n} e^{-S_{k,n}} = e^{-\frac{S_{k,n}}{2}} W_{k,n} e^{-\frac{S_{k,n}}{2}} \quad \text{with:}$$

$$S_{k,n} = 1/2.\log\left(P_{k,n}^{1/2}\left(P_{k,n}^{-1/2} P_{k,n}^{*} P_{k,n}^{-1/2}\right)^{1/2} P_{k,n}^{1/2}\right) \quad \text{with } P_{k,n} = W_{k,n}^{+} W_{k,n}$$

$$G_n = \gamma_n \sum_{\substack{k=1 \\ k \neq l}}^{m} H_{k,n} \quad \text{with} \quad \left\{ l / \|H_{k,n}\|_F < \varepsilon \right\}$$

$$\text{For } k = 1,...,m \quad \text{then} \quad W_{k,n+1} = \Phi_{G_n}\left(W_{k,n}\right)$$

$$W_{k,n+1} = \left(I - G_n G_n^{+}\right)^{-1/2}\left(W_{k,n} - G_n\right)\left(I - G_n^{+} W_{k,n}\right)^{-1}\left(I - G_n^{+} G_n\right)^{1/2}$$

$$W_{median,n+1} = \left(G G_n^{+} + I\right)^{-1}\left(G + G_n\right) \quad \text{with} \quad G = \left(I - G_n G_n^{+}\right)^{1/2} W_{median,n}\left(I - G_n^{+} G_n\right)^{-1/2}$$

THALES

www.thalesgroup.com

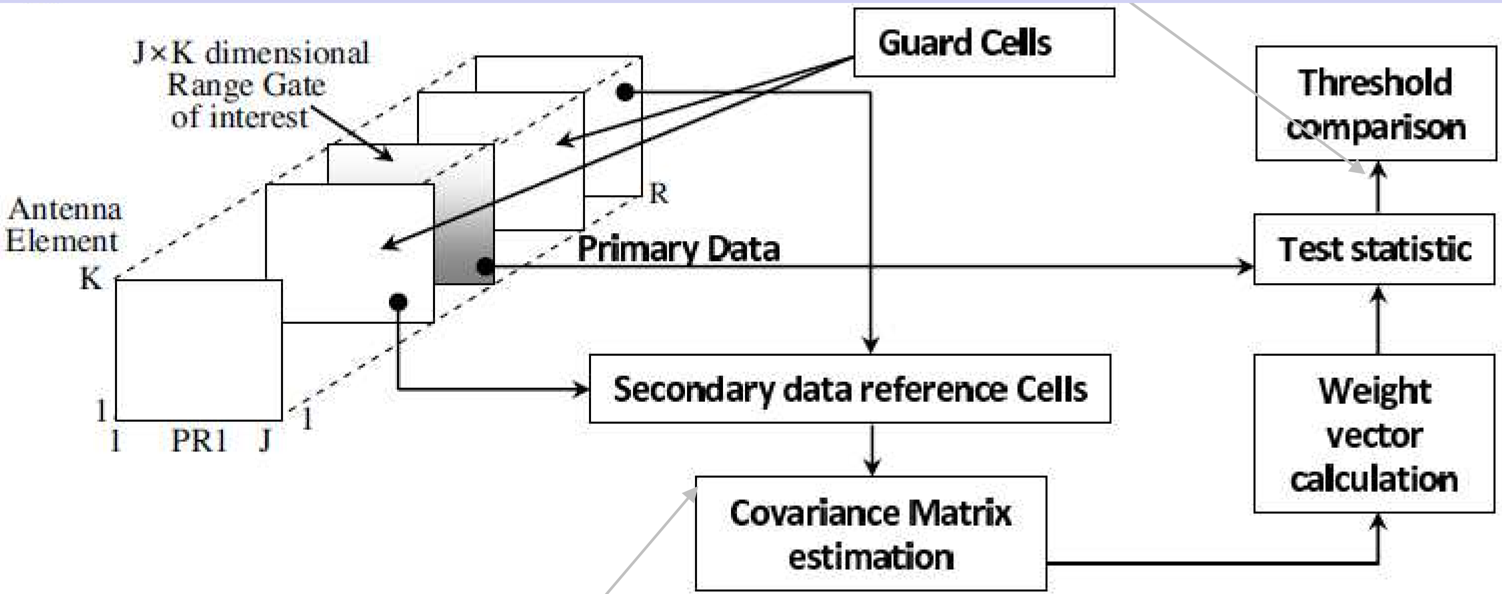
New « Ordered Statistic Space-Time Adaptive Processing » (OS-STAP)

THALES

Secondary Data Covariance Matrix Estimation

$$\Phi_Z(W) = \left(I - ZZ^+\right)^{-1/2}(W - Z)\left(I - Z^+W\right)^{-1}\left(I - Z^+Z\right)^{1/2}$$

$$d\left(R, R_{median}\right) = \left\|\log\left(R_0^{-1/2} R_{0,median} R_0^{-1/2}\right)\right\|^2 + \sum_{k=1}^{n-1} \log^2\left(\frac{1 + \left\|\Phi_{A_k^k}\left(A_{k,median}^k\right)\right\|}{1 - \left\|\Phi_{A_k^k}\left(A_{k,median}^k\right)\right\|}\right)$$



$$\left\{R_{0,p}, A_{1,p}^1, \ldots, A_{n-1,p}^{n-1}\right\}_{p=1}^{M} \underset{Mostow/Berger}{\longrightarrow} \left\{R_{0,median}, A_{1,median}^1, \ldots, A_{n-1,median}^{n-1}\right\}$$

A Riemannian approach for training data selection in STAP processing

J-Fr. DEGURSE[1,2], L. SAVY[1], J-Ph. MOLINIE, Prof. S. MARCOS[2]

[1]*ONERA, Electromagnetism and Radar Department (DEMR)*
[2]*Laboratoire des Signaux et des Systèmes (L2S SUPELEC-CNRS-Univ-Paris-Sud)*

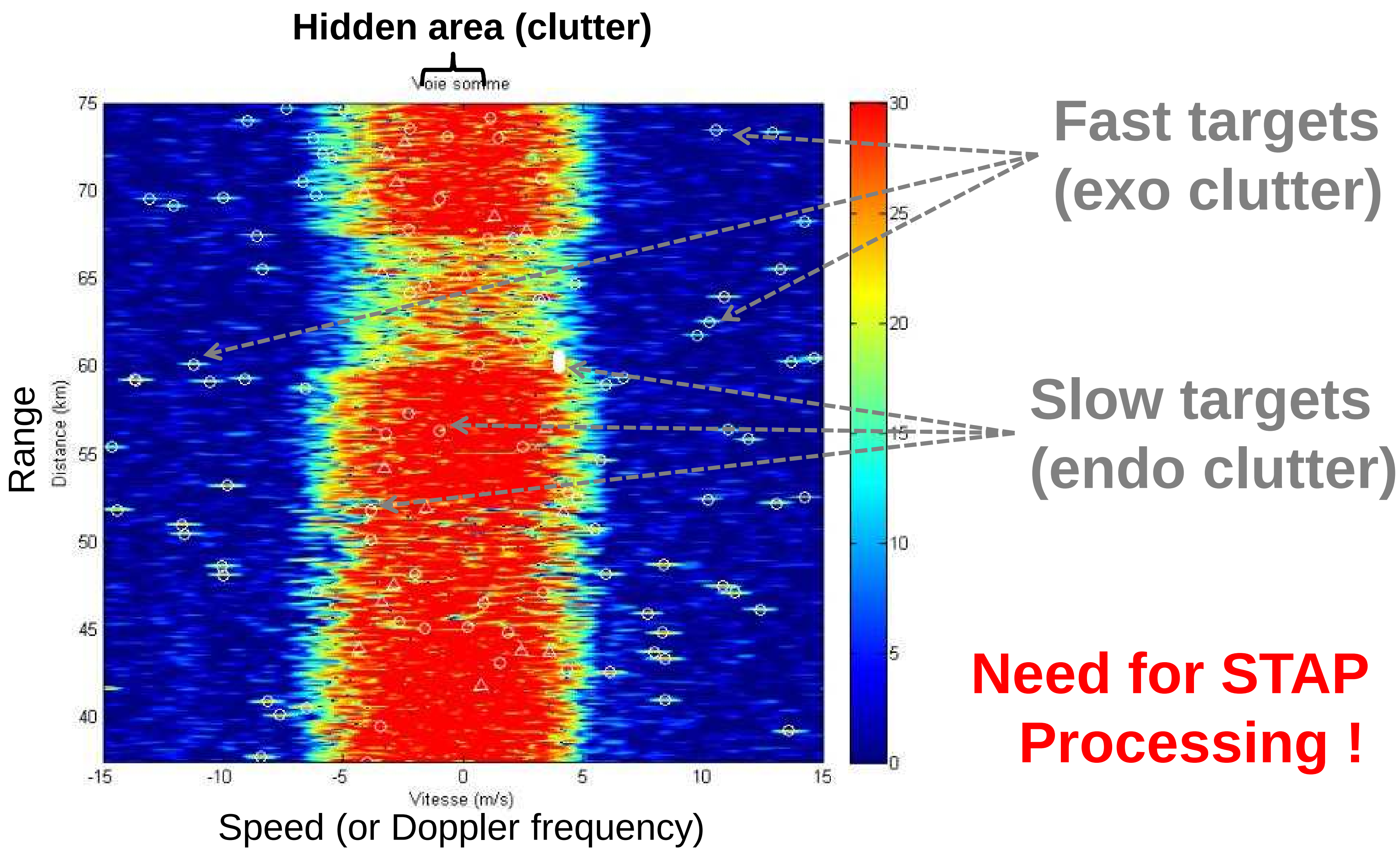
DGA

ONERA
THE FRENCH AEROSPACE LAB

retour sur innovation

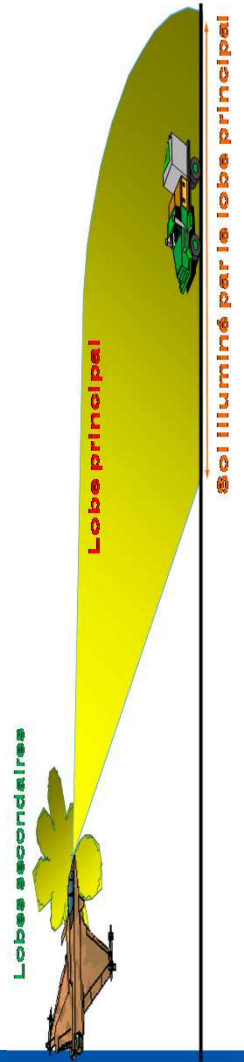
L2S

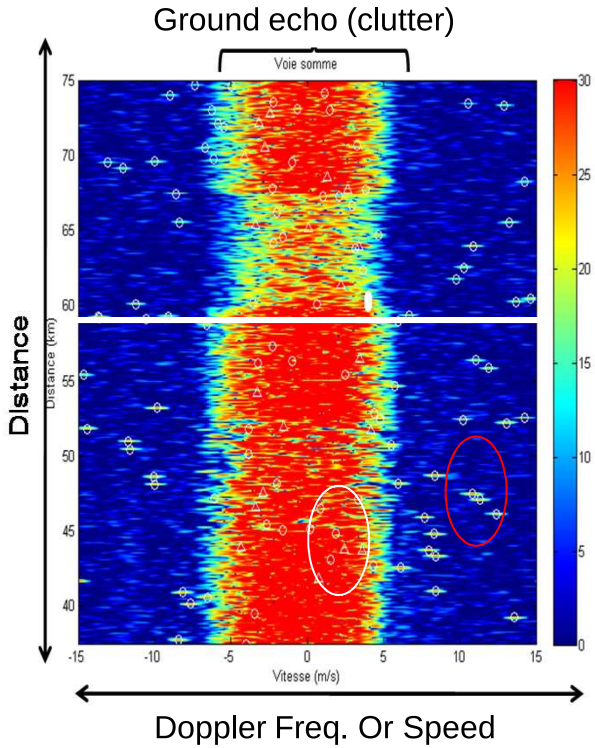
Airborne detection of slow moving targets in non stationary environments

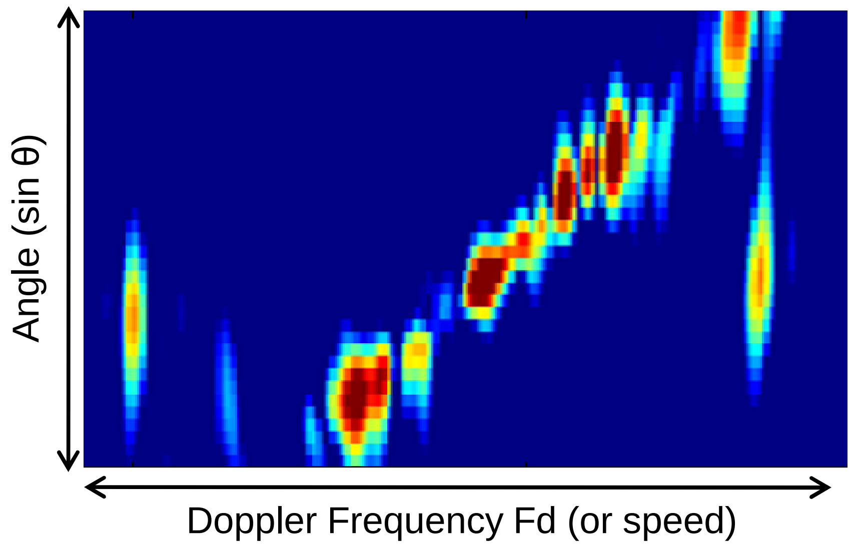


Need for STAP Processing !

Space Time Adaptive Processing (STAP)



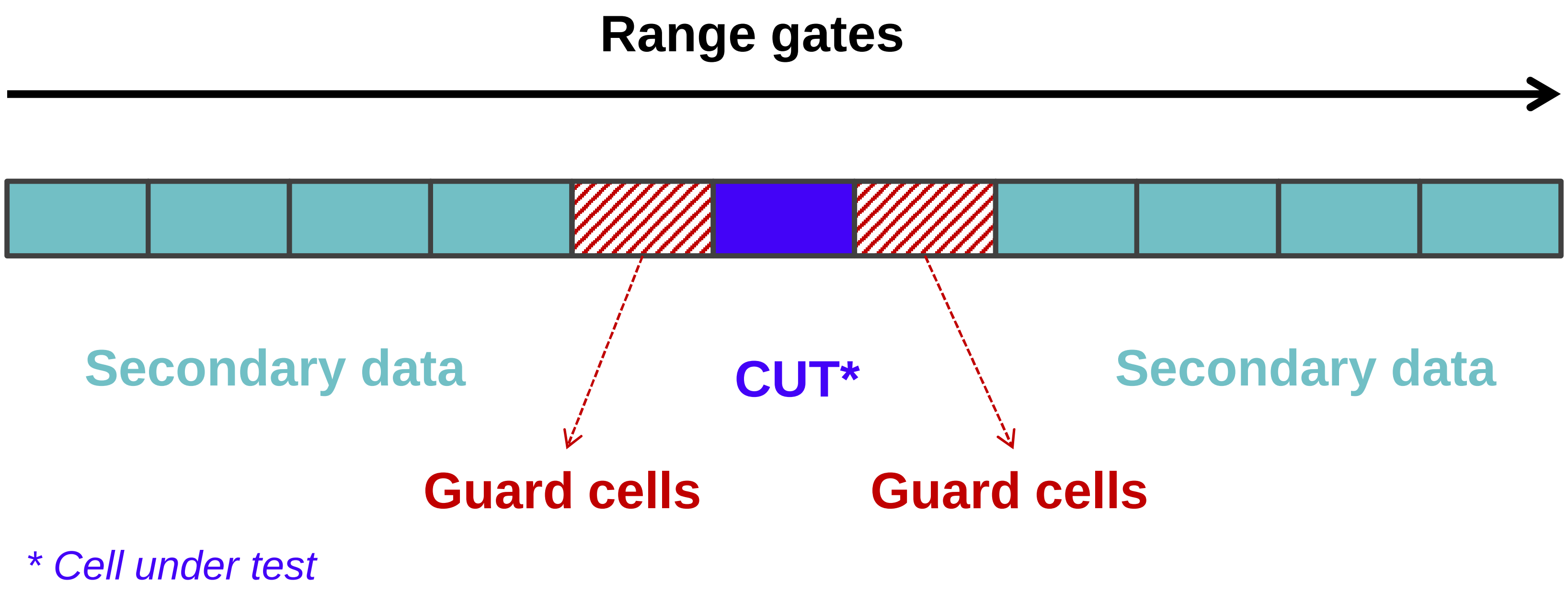




$$F_d = \frac{2V}{\lambda} \sin(\theta)$$

- Doppler and receiving angles are coupled for clutter echoes
- Monodimensionnal clutter in a bidimensionnal space

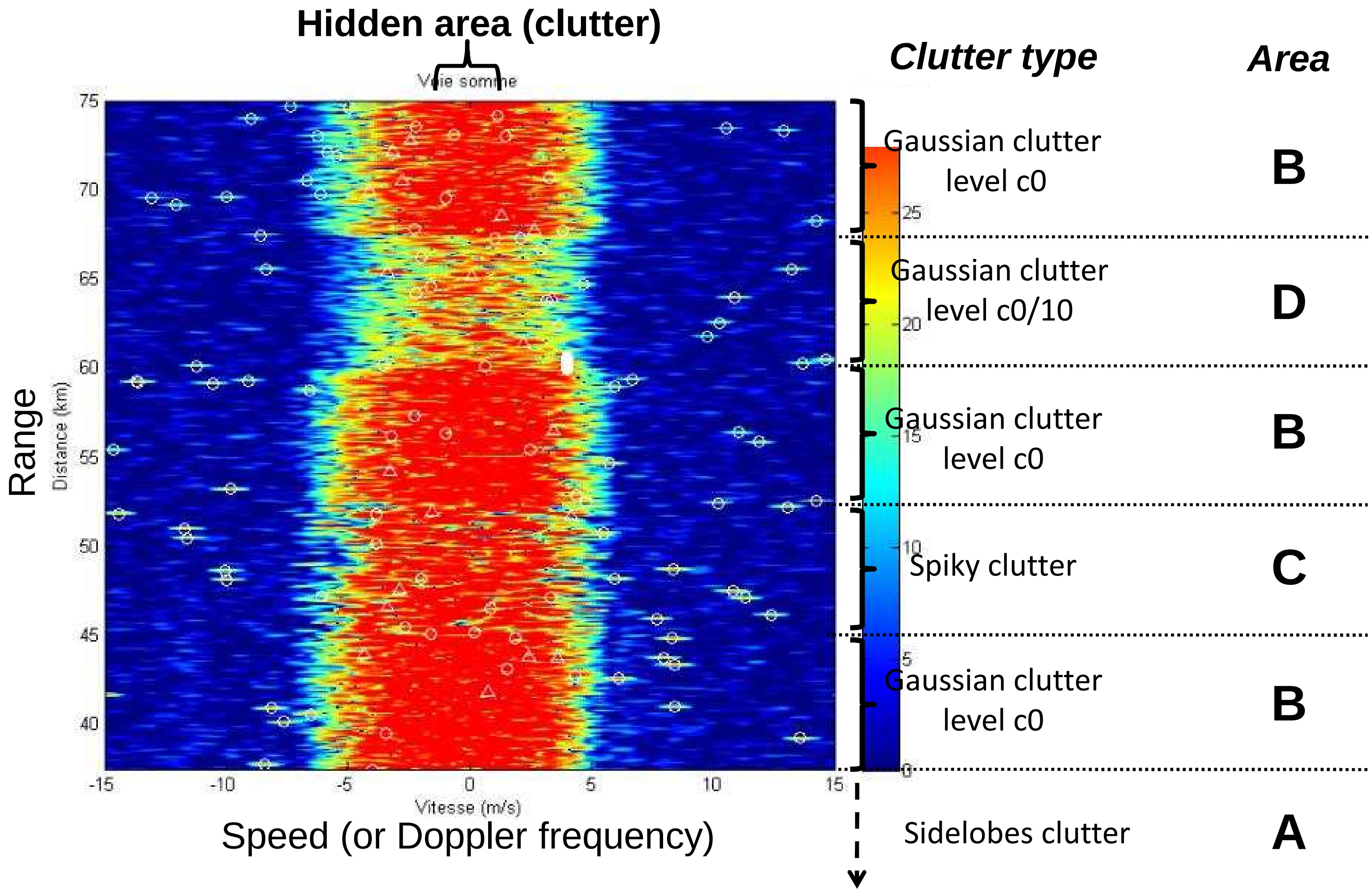
Objective



- One covariance matrix is estimated for each range gate
- Clutter from secondary data must be homogenous with clutter from CUT

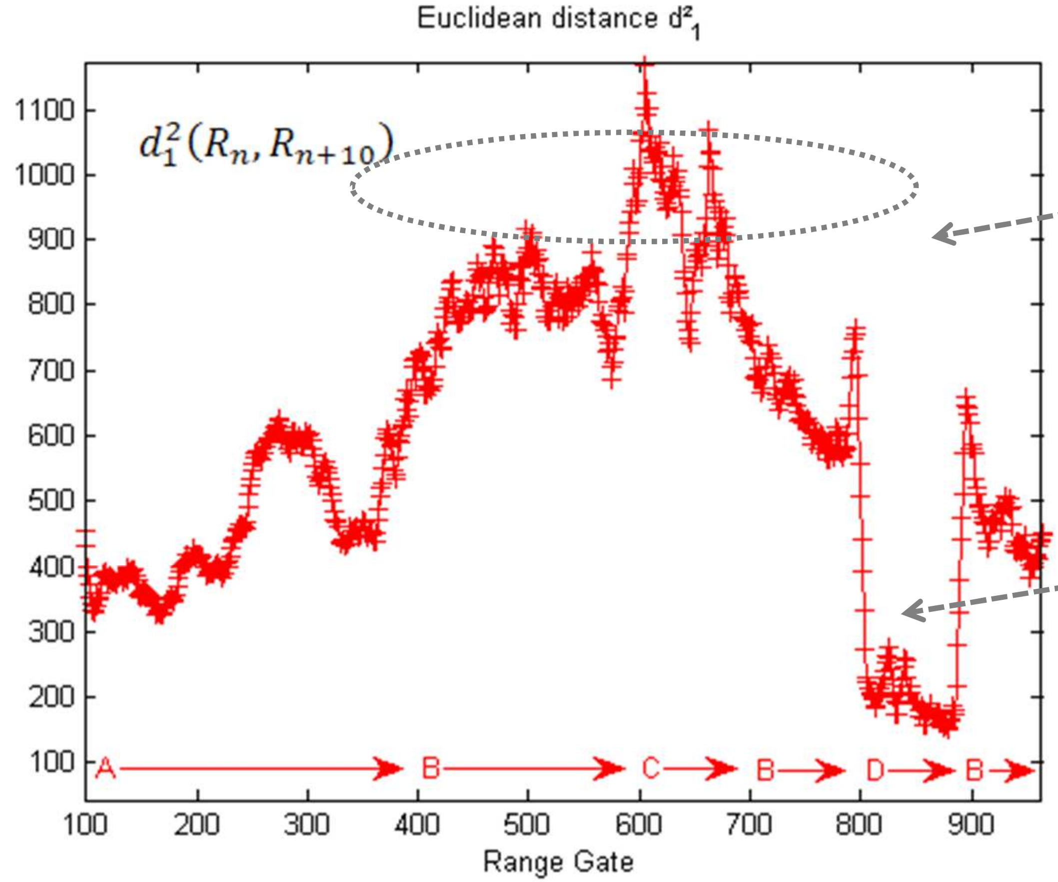
Which secondary data to choose?

Detection of slow moving targets in non stationnary environements



Use Euclidean distance between covariance matrices:

$$d_1^2(R_l, R_0) = \|R_l - R_0\|_F^2 = Tr[(R_l - R_0)(R_l - R_0)^H]$$



Main lobe clutter energy

Less powerful clutter area

Only detects power changing regions

Physical criterion for secondary data selection

$w_l^H = s_s^H R_l^{-1}$ is the weights vector obtained with R_l from the l range cell

then applied to X_0: $\quad z = w_l^H X_0$

Output $\longrightarrow$ $z = s_s^H \underbrace{R_l^{-1/2}}_{\text{matched filter}} \underbrace{R_l^{-1/2}}_{\text{whiten I+N}} X_0 = s_s^H R_l^{-1/2} y$

- if w_l^H is the optimal weights vector, y is "white", hence $E\{yy^H\} = I$

$$E\left\{R_l^{-1/2} X_0 X_0^H R_l^{-1/2}\right\} = I$$

$$R_l^{-1/2} E\{X_0 X_0^H\} R_l^{-1/2} = I \qquad R_l^{-1/2} R_0 R_l^{-1/2} = I$$

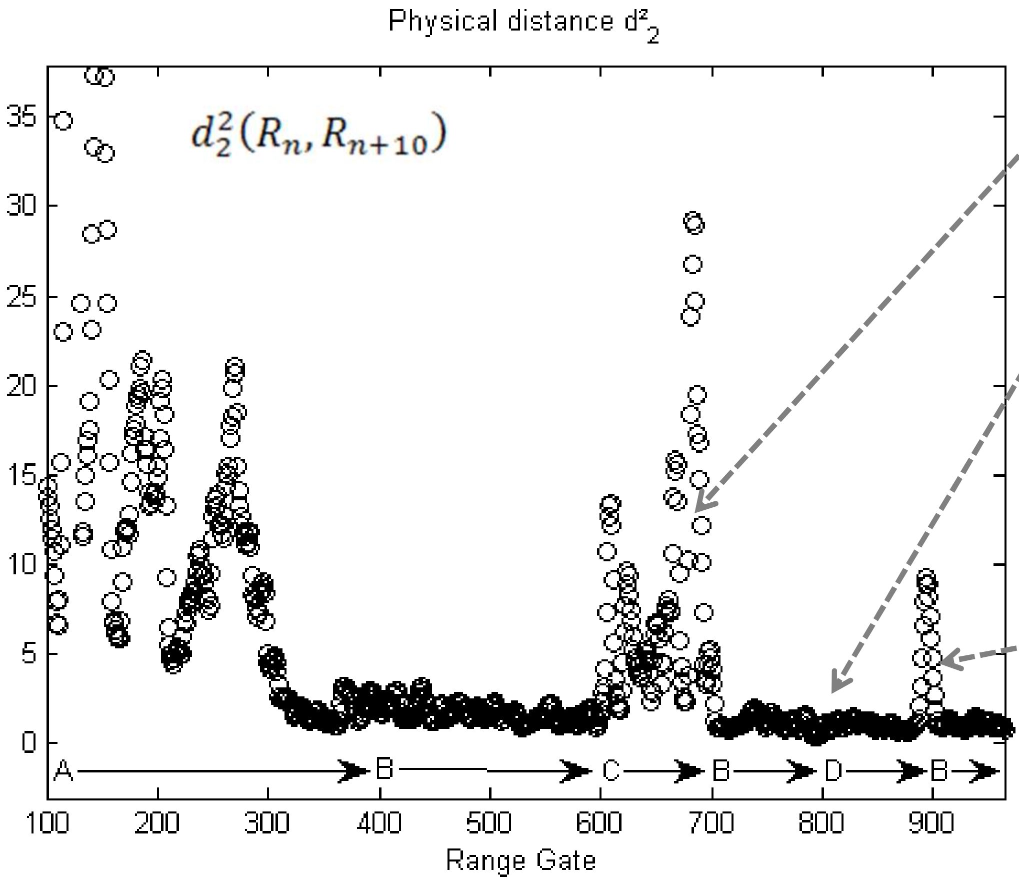
we want $\min_{R_l} \left\{R_l^{-1/2} R_0 R_l^{-1/2} - I\right\}$

- that leads to the optimal criterion or "distance" but this is **NOT** a distance:

$$d_2^2(R_l, R_0) = \left\| R_l^{-1/2} R_0 R_l^{-1/2} - I \right\|_F^2$$

Optimal criterion between covariance matrices:

$$d_2^2(R_l, R_0) = \left\| R_l^{-1/2} R_0 R_l^{-1/2} - I \right\|_F^2$$



Heterogeneous area (spiky clutter)

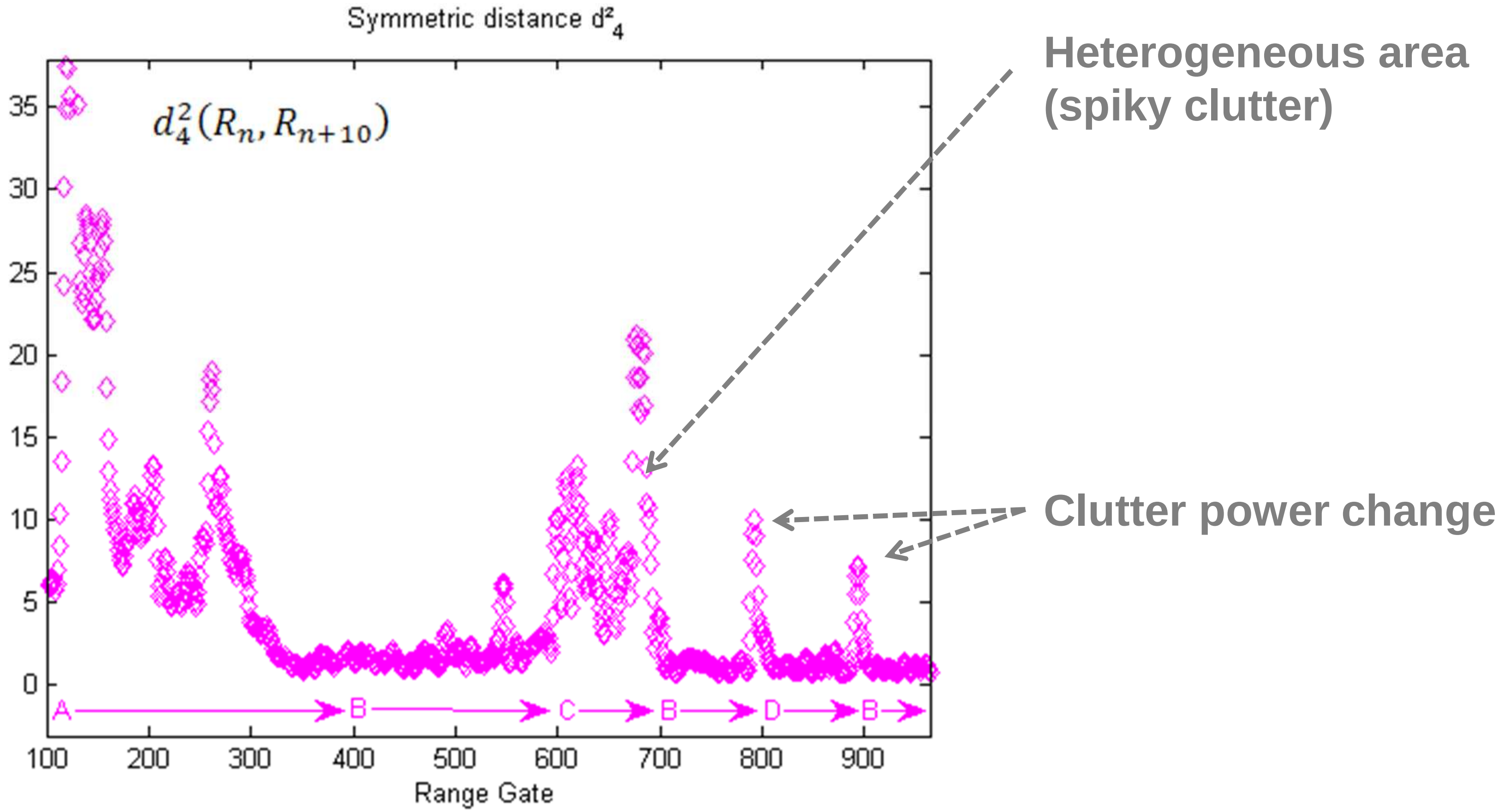
No detection of the high to low power clutter change

Clutter power change

Secondary data selection: physical POV

Symetrized physical criterion between covariance matrices:

$$d_4^2(R_l, R_0) = \left[\frac{1}{2}\left(\left\|R_l^{-1}R_0 - I\right\|_F + \left\|R_0^{-1}R_l - I\right\|_F\right)\right]^2$$



Link to the Riemannian geometry

- **For positive definite Hermitian covariance matrix, the right metric is defined by:**

$$ds^2 = Tr((R^{-1}(dR))^2)$$

and the distance: $d_3^2(R_l, R_0) = \sum_{k=1}^{n} \log^2(\lambda_k)$ *with* $\det(R_0 - \lambda R_l) = 0$

$$d_3^2(R_l, R_0) = \left\| \log\left(R_l^{-1/2} R_0 R_l^{-1/2}\right) \right\|_F^2$$

- **From the following serie**: $\log(A) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{(A-I)^n}{n}$

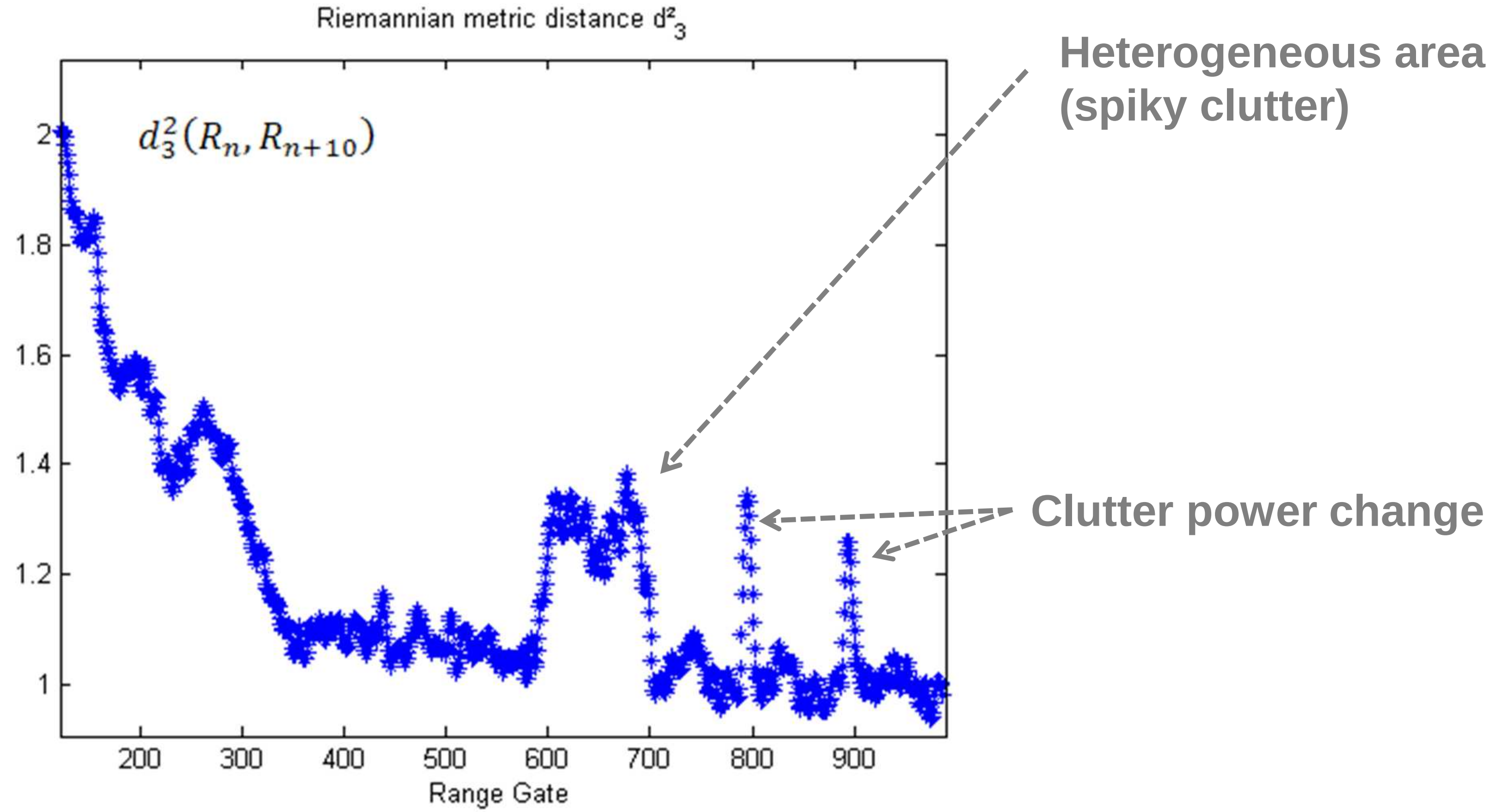
first order approximation: $d_3^2(R_l, R_0) \approx \left\| R_l^{-1/2} R_0 R_l^{-1/2} - I \right\|_F^2$

This approximation is the physical criterion $d_3^2(R_l, R_0) \approx d_2^2(R_l, R_0)$

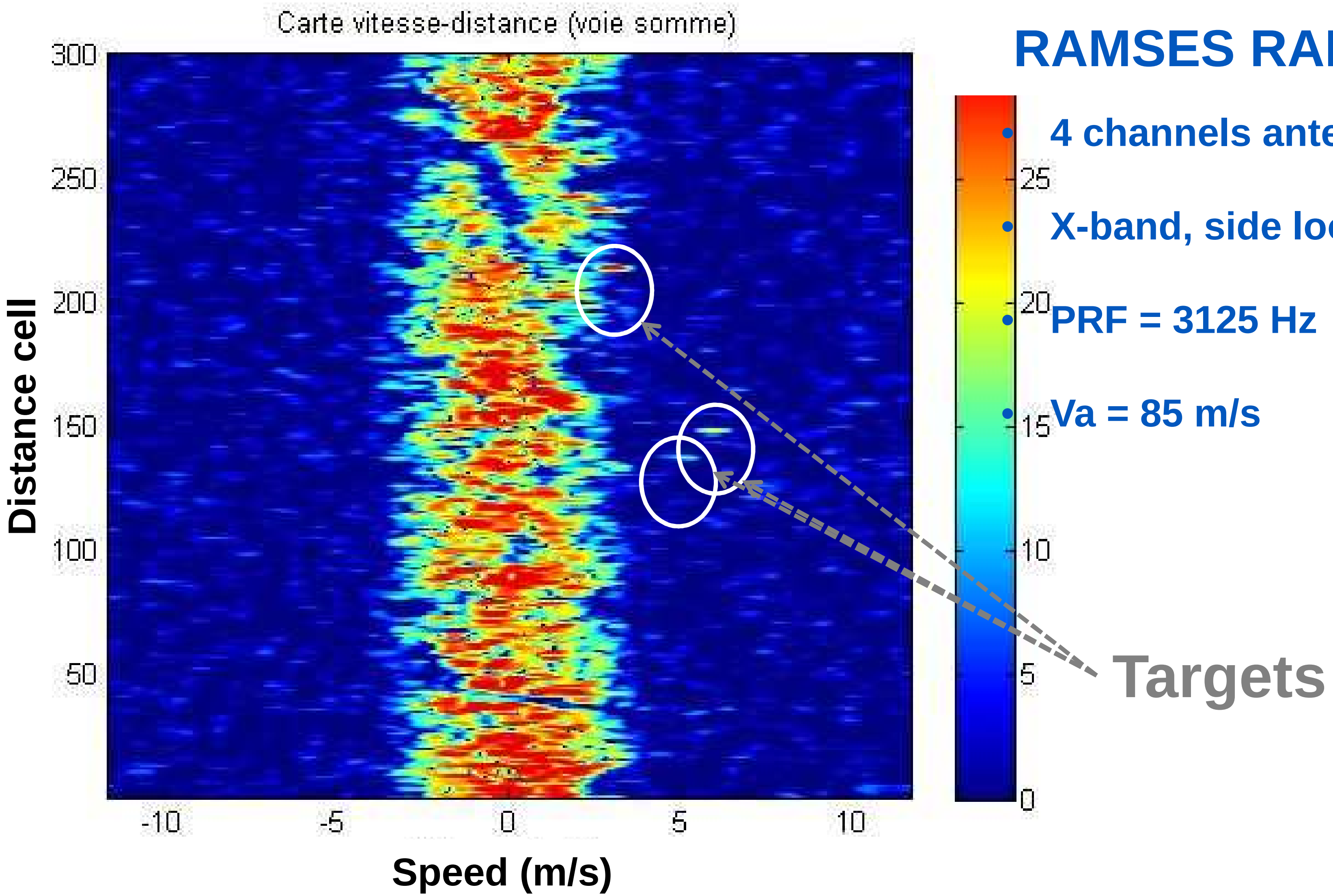
Secondary data selection using Riemannian metric distance

Riemannian distance between covariance matrices:

$$d_3^2(R_l, R_0) = \left\| \log\left(R_l^{-1/2} R_0 R_l^{-1/2}\right) \right\|_F^2$$



Real data performance: SAREX



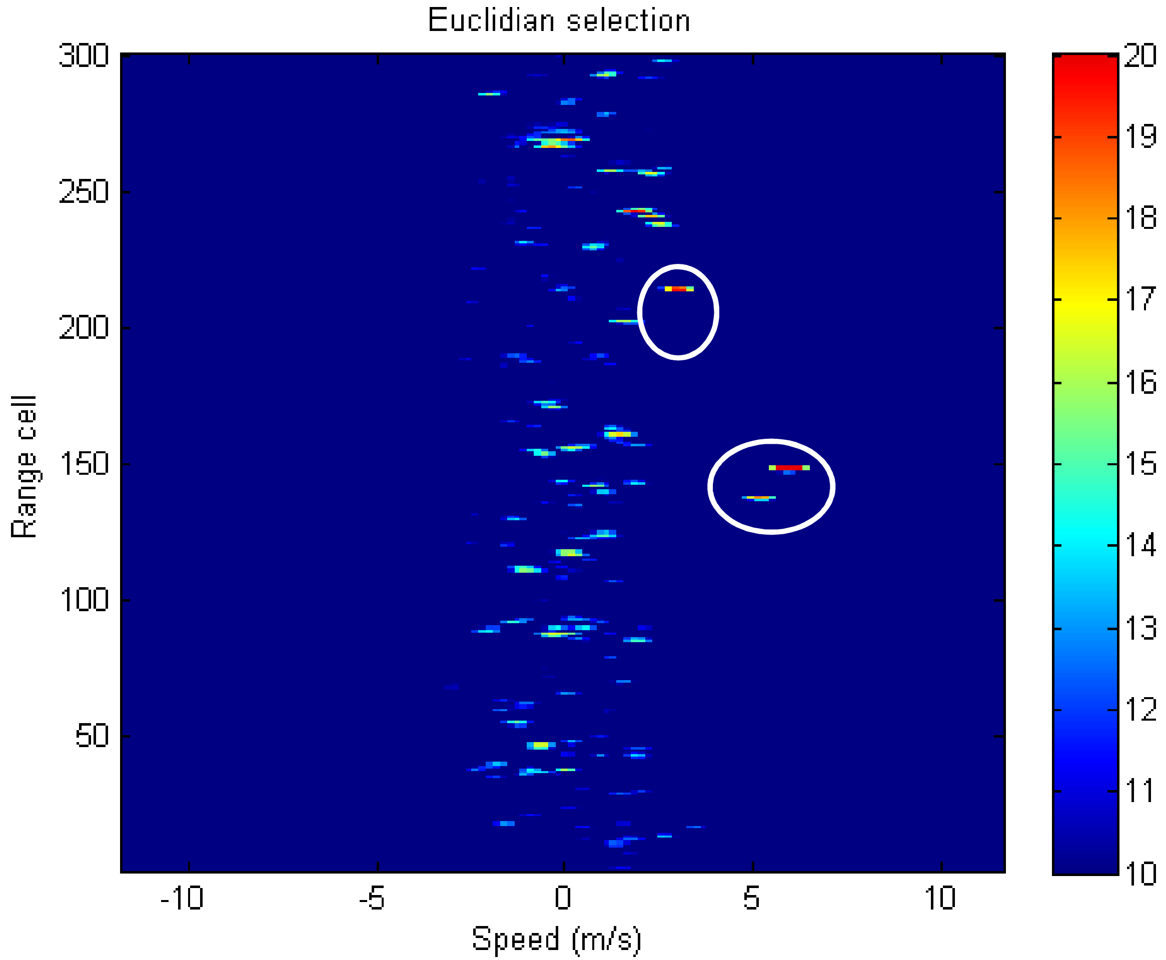
RAMSES RADAR

- 4 channels antenna
- X-band, side looking
- PRF = 3125 Hz
- Va = 85 m/s

Real data performance: SAREX (Euclidean Geometry)

<u>Selection strategy</u>: guardcells = **2**, we pick the **10** closest matrices out of **20**

a maximum distance value can be set so the number of chosen matrices could be < 10

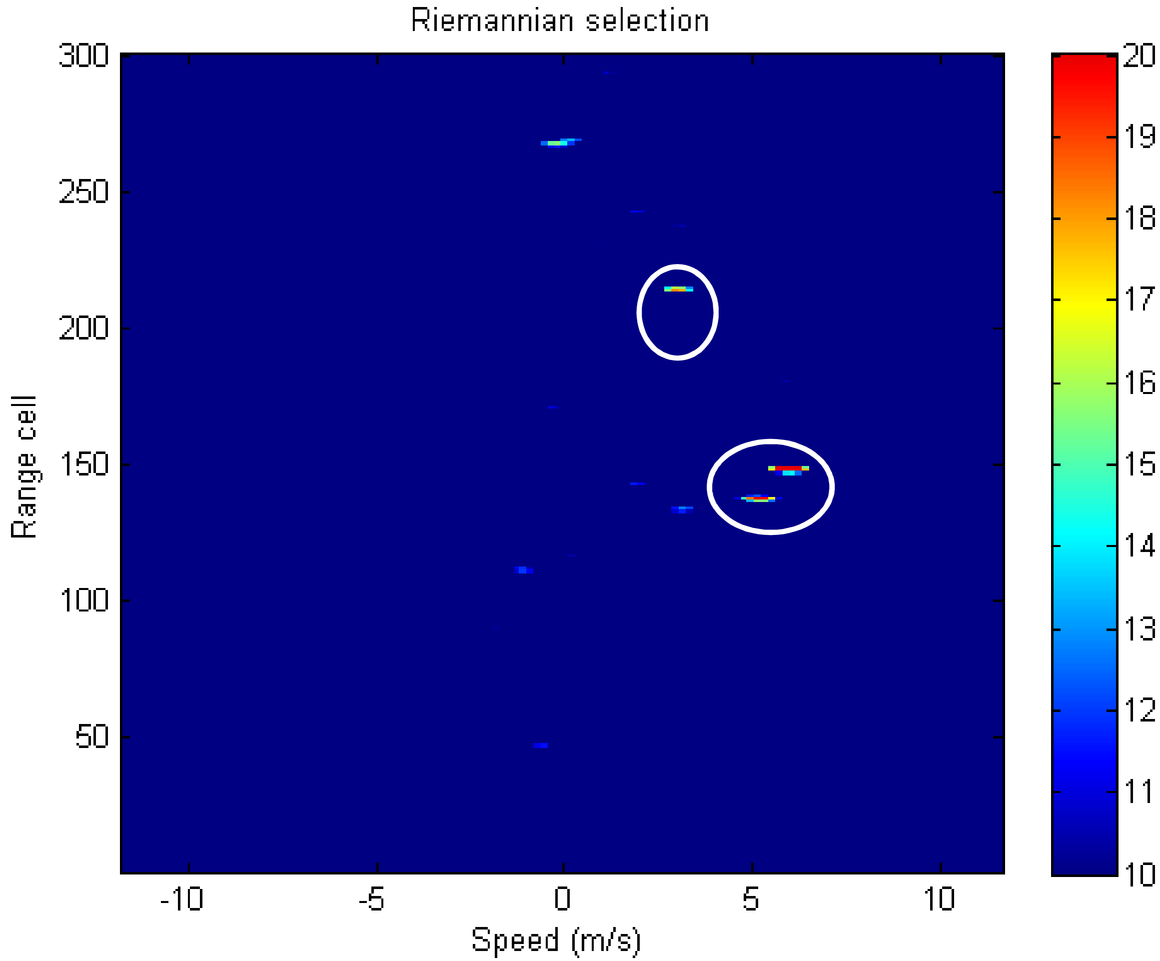


Threshold is set to 10dB

Real data performance: SAREX (Riemannian Geometry)

<u>Selection strategy</u>: guardcells = **2**, we pick the **10** closest matrices out of **20**

a maximum distance value can be set so the number of chosen matrices could be < 10



Threshold is set to 10dB

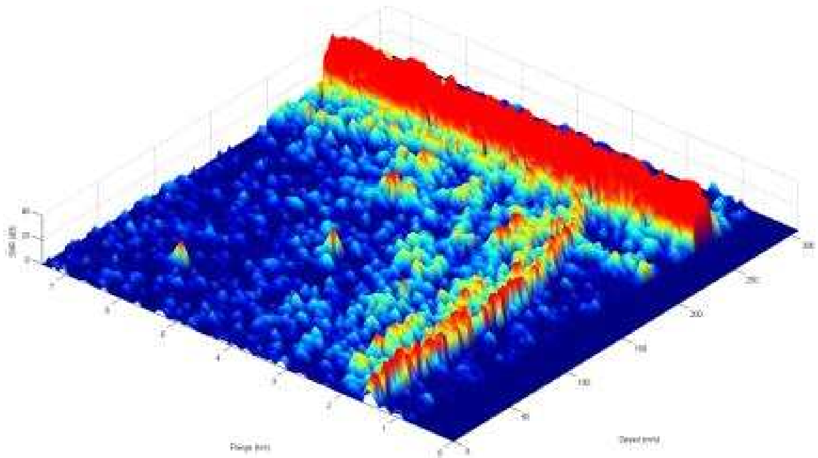
- Clutter rejection is much better
- Small signal loss (1 dB) on target 3
- Small signal gain (0.5dB) on target 1

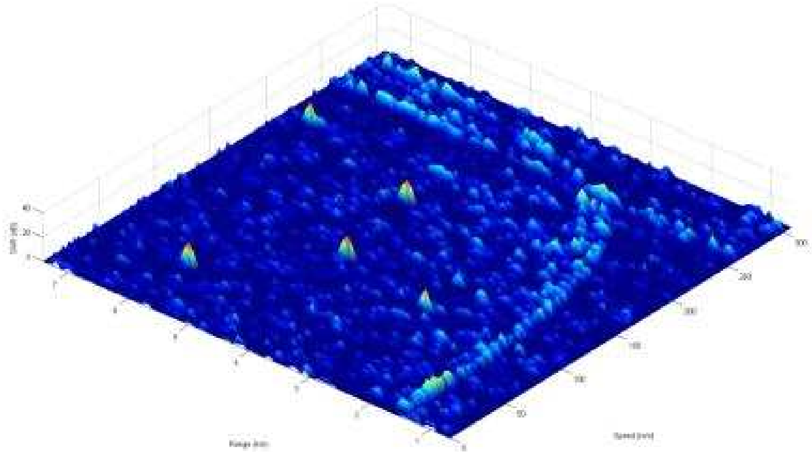
Way forward

- **Use geometric mean in Riemannian geometry instead of classical arithmetic mean**

$$R = \frac{R_1 + R_2}{2} \quad \Rightarrow \quad R_1 \circ R_2 = R_1^{1/2}\left(R_1^{-1/2} R_2 R_1^{-1/2}\right)^{\gamma} R_1^{1/2}$$

- **Application to air-to-air radar mode: high heterogeneous sidelobes clutter + continuous mainlobe clutter**





- **Algorithm implementation on GPU to increase calculation speed**
  - **massively parallel algorithms, should run very fast!**

www.thalesgroup.com

Extension to Non-Stationary Time-Doppler Signal

THALES

Non-Stationary Doppler Processing

- **New challenge in Radar is the processing of non-stationary signal**
  - fast time variation of Doppler Spectrum in one burst
- **Most of processing chains implemented in radars make the assumption of**
  - Doppler stationary signal during the burst waveform duration.
- **This assumption is not always true, especially in case of:**
  - high speed Doppler fluctuation
  - abrupt Doppler changes

  of target or clutter signal
- **We can observe non-stationarity for**
  - high speed or abrupt Doppler variations of clutter or target signal but also in case of target migration during the burst duration due to high range resolution.

Non-Stationary Doppler: Target Signal

Detection of Target Non-Stationary Doppler Signal

- **High speed Doppler fluctuation:**
  - Rotor and blades of helicopter
- **Abrupt Doppler changes:**
  - helicopter pop-up or low altitude target demasking behind a creast line
  - rocket with splitting events (multi-headed rocket, decoys, debris,...), Multiple Reentry vehicle payload for a ballistic missile that deploys multiple warheads
  - missile firing by an airborne platform
- **Target Range Migration in High Range Resolution Mode (UWB)**

Applications for other MFR radar functions

- **Kill Assessment for staring antenna (long burst/wafeform)**
- **Manoeuver Detection for Tracking (coupled with Radar Res. Mgt)**
- **Advanced NCTR (Non-Cooperative Target Recognition)**
  - Robust target recognition with time-scale invariance properties

Non-Stationary Doppler: Clutter Signal

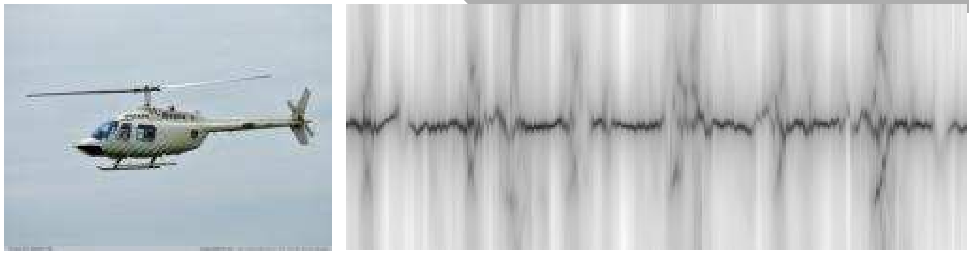
Non-Stationary Doppler Signal of Non-homogeneous Clutter (amplitude/Doppler/Polar fluctuations)

- **Sea Clutter**
  - Littoral Environment
  - Breaking Waves
  - Spikes
- **Atmospheric clutter**
  - Turbulent Atmosphere (high Eddy Dissipation Rate)
  - Wind-Shear, Micro-Burst, Downdraft

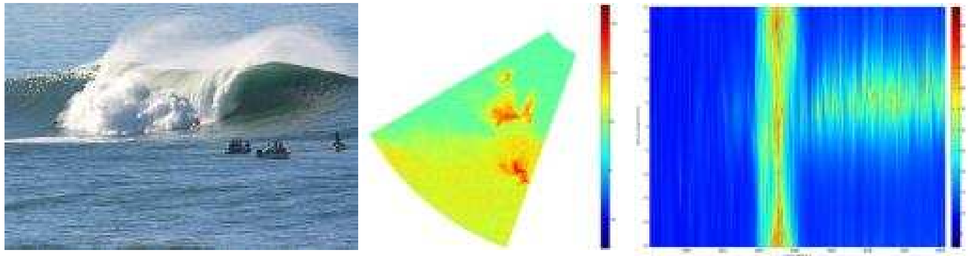
Wake-Vortex of all airborne platforms

- **Wake-Vortex of Civil Commercial Aircrafts on airport (European SESAR, US Nexten in US, Japanese CARATS , NTU ATM RI « Research Institute » in Singapore)**
- **Wake-Vortex of helicopters in Non Light of Sight**
- **Wake-Vortex of fighters / Delta Wing UCAV in Rain**

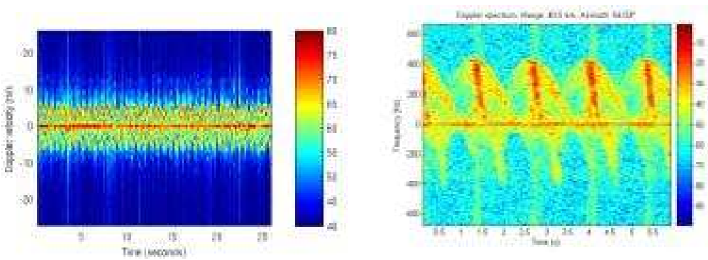
Non-Stationary Time-Doppler Signal in Radar



Time-Doppler Signature of BEL206 helico on one range cell staring antenna

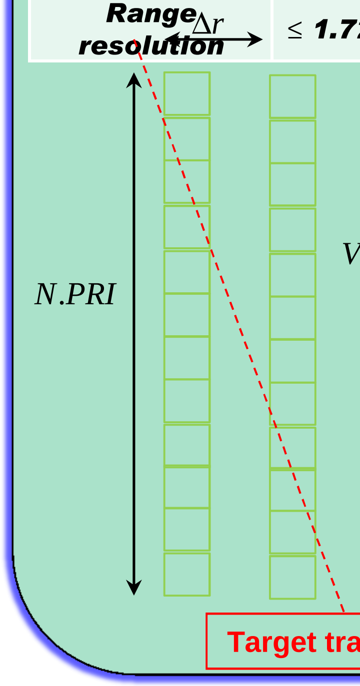


Doppler signature of Sea Clutter



Time-Doppler signature of wind farm turbine blades

Bande	UHF	S	X
F_{mean}	435 MHz	3.15 GHz	10.00 GHz
"UWB"	≥ 87 MHz	≥ 630 MHz	≥ 2 GHz
Range resolution	≤ 1.72 m	≤ 23.8 cm	≤ 7.5 cm



$$V_{rad}^{cible} > \frac{\Delta r}{N.PRI}$$

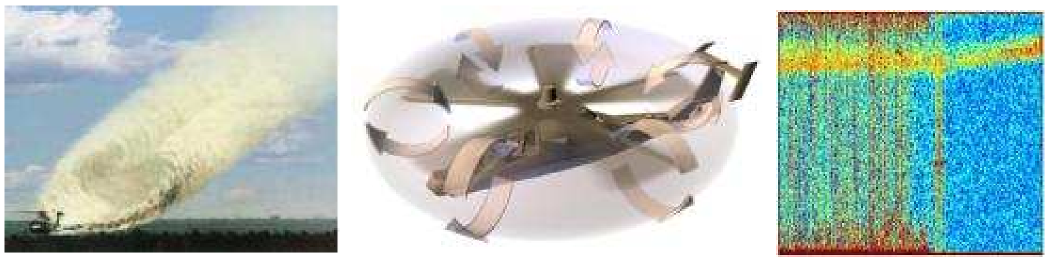
Exemple :

$N.PRI = 24\,ms\ (12 \times 2\,m/s)$

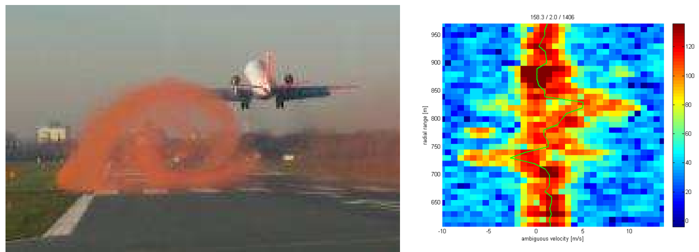
$Portée : 300\,km$

$\Delta r = 25\,cm$

$V_{rad}^{cible} > 10\,m/s\ (36\,km/h)$



Time-Doppler signature of Helco Blades Wake-Vortex



Doppler signature of aircraft wake-vortex

www.thalesgroup.com

NEW APPROACH AND MODEL FOR NON-STATIONNARY DOPPLER SIGNAL:

LOCAL STATIONARITY MODEL OF NON-STATIONARY TIME SERIES

THALES

State-of-the-Art: Active Research

Non-Stationnary Time Series Analysis

- Based on past work **: Y. Grenier, “Time dependent ARMA modelling of nonstationary signals”. IEEE, Trans. Acoust. Speech Signal Process. 31, 899–911, 1983**
- New active Research**: Workshop 2012 : SFdS (Société Française de statistique)**

PREDICTION OF TIME SERIES AND NON STATIONARY TIME SERIES

February 10th-11th, 2012

- **Time series do not benefit from generous properties of stationarity of ergodicity which allow to recover the law of a process by the only observation. Various alternatives are envisaged:**
  - the **Local Stationarity** allows to take into account regular variations of behavior
  - The **Isotonique Regression** gives one simple model in which a test of stationarity is even possible
  - The **Hidden Markov Chains** models the times of these breaks

State-of-The-Art: LOCAL STATIONARITY

- **Heidelberg University, 2012 : R. Dahlhaus, « Locally Stationary Processes », Handbook of Statistics, 30, 2012, http://arxiv.org/abs/1109.4174**
  - *If one restrict to linear processes or even more to Gaussian processes then a much more general theory is possible. We give a general definition for linear processes and discuss time varying spectral densities in detail. Study contains the Gaussian likelihood theory for locally stationary processes.*
- **Mannheim University, 2011: Vogt, M. "Nonparametric regression for locally stationary time series". PhD, 2011**
  - we study nonparametric models allowing for locally stationary regressors and a regression function that changes smoothly over time. These models are a natural extension of time series models with time-varying coefficients. we show that the main conditions of the theory are satisfied for a large class of nonlinear autoregressive processes with time-varying regression function.
- **Télécom ParisTech: E. Moulines, P. Priouret, et F. Roueff., "On recursive estimation for locally stationary time varying autoregressive processes", Ann. Statist. 33, 2610–2654, 2005**
  - the properties of recursive estimates of tvAR-processes have been investigated in the framework of locally stationary processes. The asymptotic properties of the estimator have been proved including a minimax result.

INFORMATION GEOMETRY MODEL OF STATIONARY TIME SERIES (Duality Concept):

SIEGEL METRIC FOR HERMITIAN POSITIVE DEFINITE (HPD) MATRIX

THALES

Siegel HPD Matrix Metric = IG Metric of Multiv. Gaussian Law

- **Information Geometry for Multivariate Gaussian Law of zero Mean and intrinsec Geometry of Hermitian Positive Definite Matrices (particular case of Siegel Upper-Half Plane) provide the same metric**

  - Information Geometry:

$$p(Z_n / R_n) = (\pi)^{-n} . \left|R_n\right|^{-1} . e^{-Tr\left[\hat{R}_n . R_n^{-1}\right]} \quad \text{with} \quad g_{ij}(\theta) = -E\left[\frac{\partial^2 \ln p(Z_n / \theta_n)}{\partial\theta_i . \partial\theta_j^*}\right]$$

$$\text{with} \quad \hat{R}_n = (Z_n - m_n).(Z_n - m_n)^+$$

$$\text{and} \quad E\left[\hat{R}_n\right] = R_n \qquad m_n = 0 \Longrightarrow ds^2 = Tr\left(\left(R_n^{-1}(dR_n)\right)^2\right)$$

  - Geometry of Siegel Upper-Half Plane:

$$SH_n = \left\{Z = X + iY \in Sym(n,C) / \operatorname{Im}(Z) = Y > 0\right\}$$

$$ds^2_{Siegel} = Tr\left(Y^{-1}(dZ)Y^{-1}(d\overline{Z})\right) \quad \text{with} \quad Z = X + iY$$

$$\begin{cases} X = 0 \\ Y = R_n \end{cases} \Longrightarrow ds^2 = Tr\left(\left(R_n^{-1}(dR_n)\right)^2\right)$$

Distance between HPD Matrices: particular case of Siegel

- **Siegel Distance:**
  - Particular Case (X=0) and General Case:
    - Particular Case (pure imaginary axis) : $Z = iR \quad \text{avec} \quad R > 0$

$$d^2(R_1,R_2) = \left\| \log\left(R_1^{-1/2}.R_2.R_1^{-1/2}\right) \right\|^2 = \sum_{k=1}^{n} \log^2(\lambda_k)$$

with $\det(R_2 - \lambda R_1) = 0$

    - General Case of Siegel Upper-Half Plane Distance:

$$Z = X + iY \in SH_n \quad \text{with} \quad X \neq 0$$

$$d^2_{Siegel}(Z_1,Z_2) = \left(\sum_{k=1}^{n} \log^2 \left(\frac{1+\sqrt{\lambda_k}}{1-\sqrt{\lambda_k}} \right) \right) \quad \text{with} \quad Z_1, Z_2 \in SH_n$$

with $\det\left(R(Z_1,Z_2) - \lambda.I\right) = 0$

$$R(Z_1,Z_2) = (Z_1 - Z_2)\left(Z_1 - \overline{Z}_2\right)^{-1}\left(\overline{Z}_1 - \overline{Z}_2\right)\left(\overline{Z}_1 - Z_2\right)^{-1}$$

THALES

Geometry of Hermitian Positive Definite Matrices given by:

- Geodesic : $d(R_X,\gamma(t)) = t.d(R_X,R_Y)$ with $t \in [0,1]$

$$\gamma(t) = R_X^{1/2} e^{t\log\left(R_X^{-1/2} R_Y R_X^{-1/2}\right)} R_X^{1/2} = R_X^{1/2}\left(R_X^{-1/2} R_Y R_X^{-1/2}\right)^t R_X^{1/2}$$

$$\gamma(0) = R_X \quad , \quad \gamma(1) = R_Y \text{ and } \gamma(1/2) = R_X \circ R_Y$$

Properties of this space

- Symetric Space as studied by Elie Cartan : Existence of bijective geodesic isometry

$$G_{(A,B)} X = (A \circ B) X^{-1} (A \circ B) \quad \text{avec} \quad A \circ B = A^{1/2}\left(A^{-1/2} B A^{-1/2}\right)^{1/2} A^{1/2}$$

- Bruhat-Tits Space : semi-parallelogram inequality

$$\forall x_1, x_2 \quad \exists z \ \text{ tel que } \forall x$$

$$d(x_1,x_2)^2 + 4d(x,z)^2 \leq 2d(x,x_1)^2 + 2d(x,x_2)^2 \quad \forall x \in X$$

- Cartan-Hadamard Space (Complete, simply connected with negative sectional curvature Manifold)

www.thalesgroup.com

MODEL NON-STATIONARY TIME SERIES AS GEODESIC PATH OF LOCALLY STATIONARY SIGNAL ON THPD MATRIX MANIFOLD

THALES

NS TIME SERIES = GEODESIC PATH OF LS TIME SERIES ON MANIFOLD

Locally Stationarity Assumption

- **We will assume that each non-stationary signal in one burst can be split into several short signals**
  - with less Doppler resolution
  - but locally stationary,

  represented by
  - time sequence of stationary covariance matrices (Toeplitz Hermitian PD matrix)
  - a geodesic path/polygon on covariance matrix manifold (THPD Matrix Manifold)
- **we will consider Time-Doppler signature:**
  - as a geodesic path on an Information Geometry manifold (of covariance matrices)
  - by analyzing the time series in the burst with a short sliding window

THALES

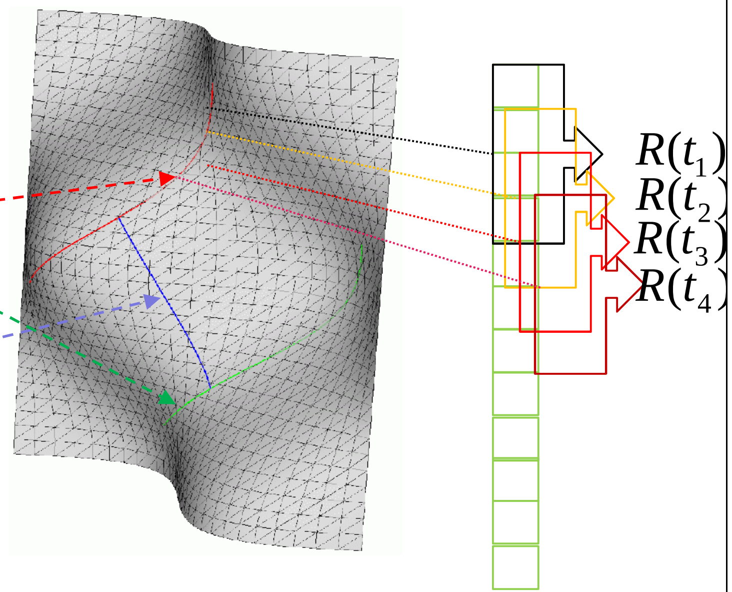
NS TIME SERIES = GEODESIC PATH OF LS TIME SERIES ON MANIFOLD

$$\gamma(t) = R_X^{1/2} e^{t\log\left(R_X^{-1/2} R_Y R_X^{-1/2}\right)} R_X^{1/2} = R_X^{1/2}\left(R_X^{-1/2} R_Y R_X^{-1/2}\right)^t R_X^{1/2}$$

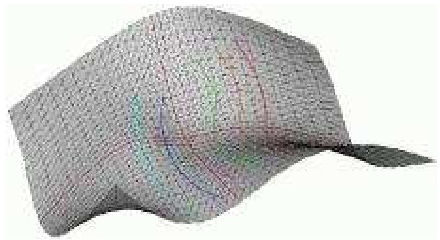
$$\gamma(0) = R_X \quad , \quad \gamma(1) = R_Y \quad \text{et} \quad \gamma(1/2) = R_X \circ R_Y$$

$$d^2(R_1, R_2) = \left\|\log\left(R_1^{-1/2}.R_2.R_1^{-1/2}\right)\right\|^2 = \sum_{k=1}^{n} \log^2(\lambda_k)$$

$$\det\left(R_2 - \lambda R_1\right) = 0$$



EACH TIME SERIE IS A PATH ON COVARIANCE MATRIX MANIFOLD



www.thalesgroup.com

FRECHET DISTANCE BETWEEN CURVES AND GEODESIC EXTENSION ON MANIFOLD

THALES

FRECHET Distance between Curves/Paths in Euclidean Space R^n

Classical Frechet/Hausdorff distances between Curves

- **Classically, Hausdorff distance, that is the maximum distance between a point on one curve and its nearest neighbor on the other curve, does not take into account the flow of the curves**
- **The Fréchet distance between two curves is defined as:**

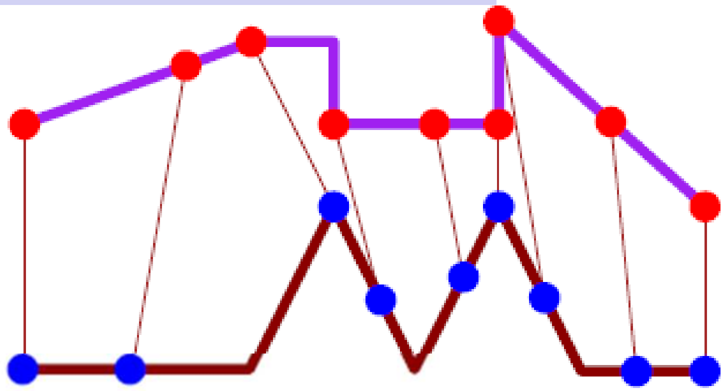
  The minimum length of a leash required to connect a dog and its owner as they walk without backtracking along their respective curves from one endpoint to the other.

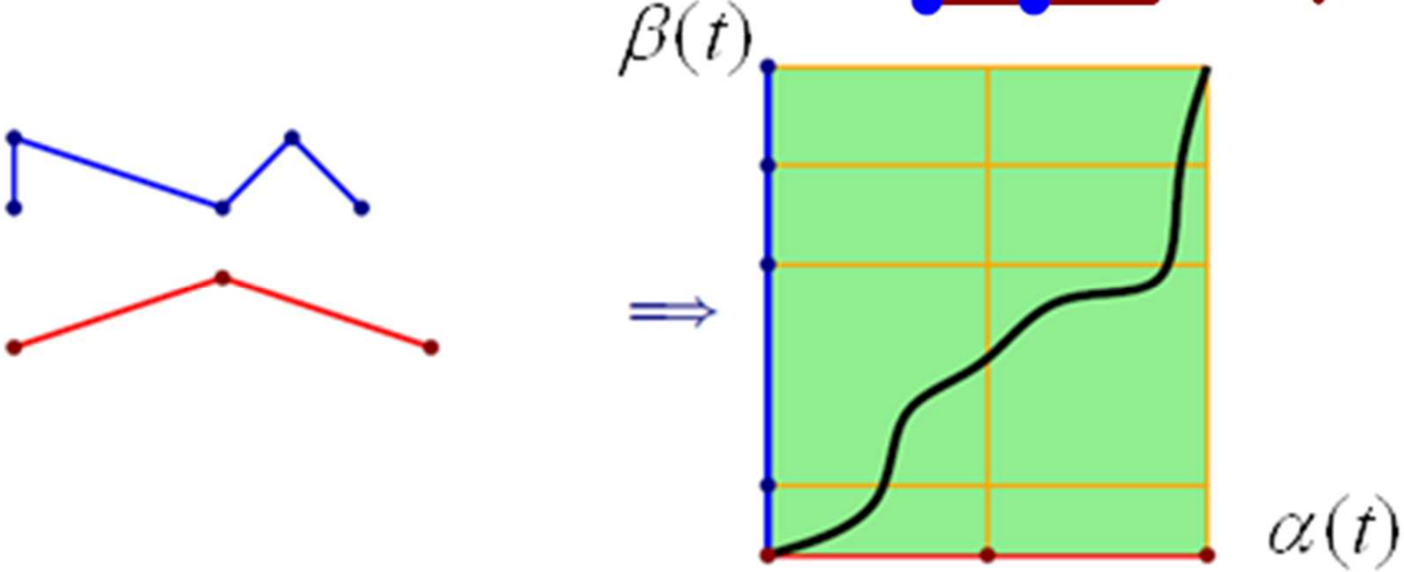
- **Let *P* and *Q* be two given curves, the Fréchet distance between *P* and *Q* is defined as the infimum over all reparameterizations α and β of $[0,1]$ of the maximum over all $t \in [0,1]$ of the distance in between $P(\alpha(t))$ and $Q(\beta(t))$. In mathematical notation:**

$$d_{Fréchet}(P,Q) = \underset{\alpha,\beta}{Inf}\ \underset{t\in[0,1]}{Max}\left\{d\left(P(\alpha(t)),Q(\beta(t))\right)\right\}$$

FRECHET Distance between Curves/Paths in Euclidean Space R^n

$$\begin{cases} d_{Fréchet}(P,Q) = \underset{\alpha,\beta}{Inf} \underset{t\in[0,1]}{Max} \left\{ d\left(P(\alpha(t)),Q(\beta(t))\right)\right\} \\ \alpha \text{ and } \beta : [0,1] \rightarrow [0,1] \text{ Nondecreasing and surjective} \end{cases}$$





THALES

Polynomial Time Algorithm: Free-Space Diagram

- Alt and Godau have introduced a polynomial-time algorithm to compute the Fréchet distance between two polygonal curves in Euclidean space.
- For two polygonal curves with m and n segments, the computation time is $O\left(mn\log(mn)\right)$
- Alt and Godau have defined the free-space diagram between two curves for a given distance threshold ε is a two-dimensional region in the parameter space that consist of all point pairs on the two curves at distance at most ε:

$$D_\varepsilon(P,Q) = \left\{\left(\alpha,\beta\right) \in [0,1]^2 \,/\, d_{Fréchet}\left(P(\alpha(t)), Q(\beta(t))\right) \le \varepsilon\right\}$$

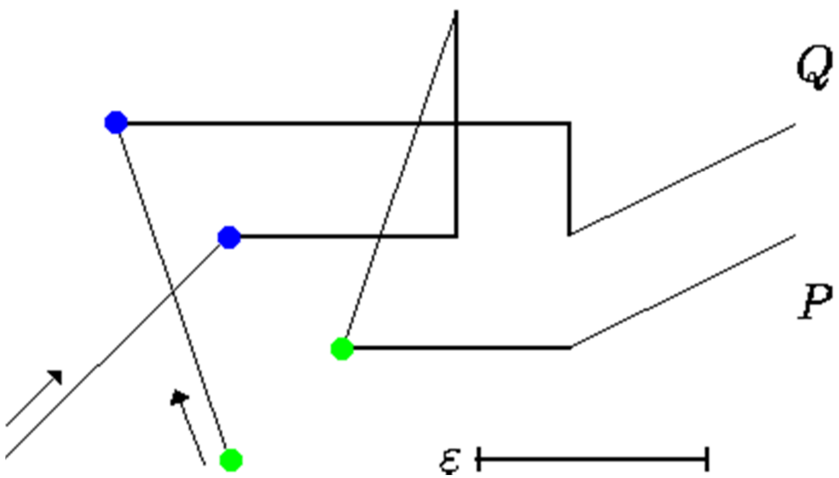
- The Fréchet distance $d_{Fréchet}\left(P,Q\right)$ is at most ε if and only if the free-space diagram $D_\varepsilon(P,Q)$ contains a path which from the lower left corner to the upper right corner which is monotone both in the horizontal and in the vertical direction.

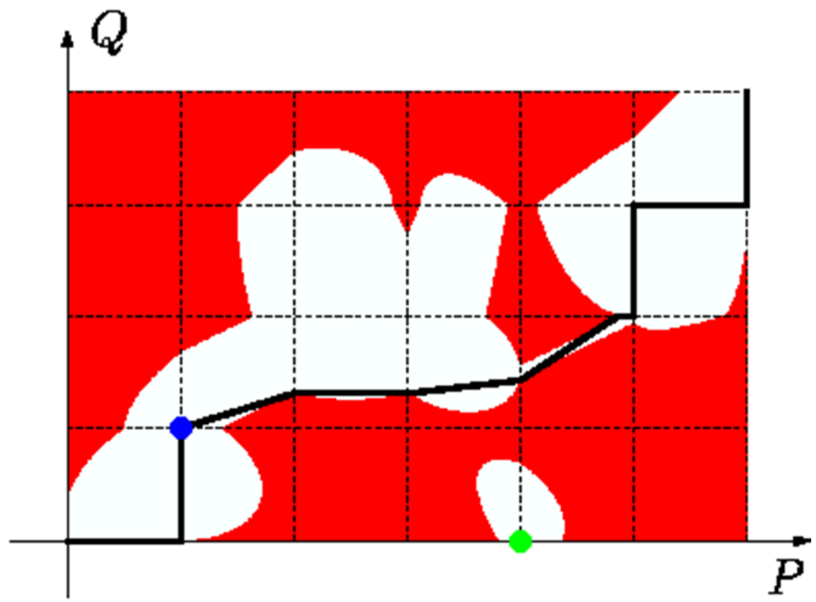
THALES

FRECHET Distance: Free-Space Diagram

Polynomial Time Algorithm: Free-Space Diagram

- In an n m free-space diagram, shown in following figure, the horizontal and vertical directions of the diagram correspond to the natural parametrizations of *P* and *Q* respectively.
- Therefore, if there is a monotone increasing curve from the lower left to the upper right corner of the diagram (corresponding to a monotone mapping), it generates a monotonic path that defines a matching between point-sets *P* and *Q*





www.thalesgroup.com

FRECHET DISTANCE OF GEODESIC CURVES REPRESENTATIVE OF TIME-DOPPLER SIGNATURE ON INFORMATION GEOMETRY MANIFOLD

THALES

FRECHET Distance of Geodesice Curves in Information Geometry

Fréchet Distance between 2 geodesic paths

- **We extend previous Fréchet distance between paths in R^n to Information Geometry Manifold.**
- **When the two curves are embedded in a more complex metric space, the distance between two points on the curves is most naturally defined as the geodesic length of the shortest path between them.**
- **If we consider *N* subsets of *M* Radar pulses in the burst, the Doppler burst can then be described by a poly-geodesic lines on Information Geometry Manifold.**
- **The set of N covariance matrices time serie $\{R(t_1), R(t_2), \ldots, R(t_N)\}$ describe a discrete "polygonal" geodesic path on Information Geometry Manifold, and we can extend previous Frechet Distance but with Geodesic distance:**

$$\begin{cases} d_{Fréchet}(R_1, R_2) = \underset{\alpha,\beta}{Inf}\ \underset{t\in[0,1]}{Max}\left\{ d_{geo}\left(R_1(\alpha(t)), R_2(\beta(t))\right)\right\} \\ \text{with } d^2_{geo}\left(R_1(\alpha(t)), R_2(\beta(t))\right) = \left\| \log\left(R_1^{-1/2}(\alpha(t)) R_2(\beta(t)) R_1^{-1/2}(\alpha(t))\right)\right\|^2 \end{cases}$$

Extended Fréchet Distance between 2 geodesic paths

- **As classical Fréchet distance doesn't take into account with *Inf[Max]* close dependence of elements between points of time series paths, we propose to define a new distance given by:**

$$d_{geo-path}\left(R_1,R_2\right)=\underset{\alpha,\beta}{Inf}\left\{\int_0^1 d_{geo}\left(R_1(\alpha(t)),R_2(\beta(t))\right)dt\right\}$$

- ***We have then to find the solution for computing the geodesic minimal path on the Fréchet free-space diagram. The length of the path is not given by euclidean metric*** $ds^2=dt^2$ ***where*** $L=\int_L ds$)

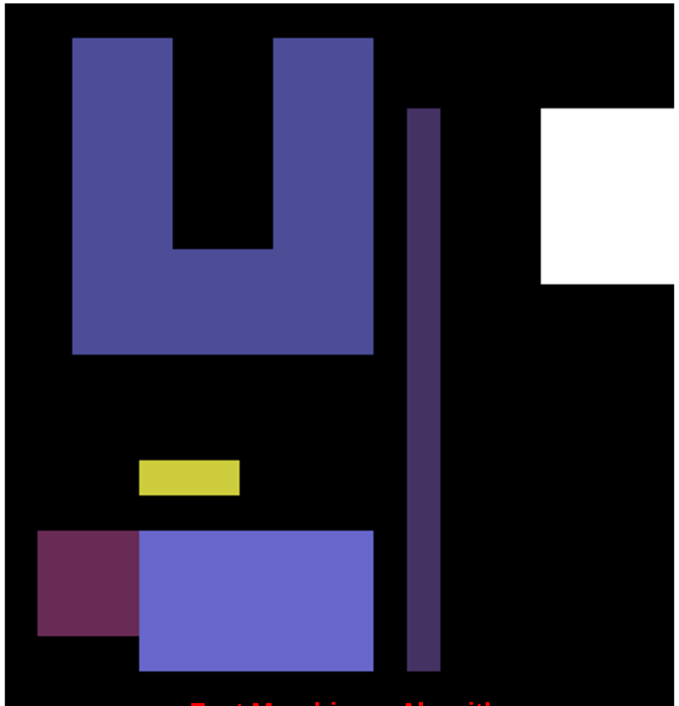
but geodesic metric weighted by d(.,.) of the free-space diagram :

$$L_g=\int_L g.ds=\int_L ds_g$$ **with** $$ds_g=d\left(R_1(\alpha(t),R_2(\beta(t))\right).dt$$

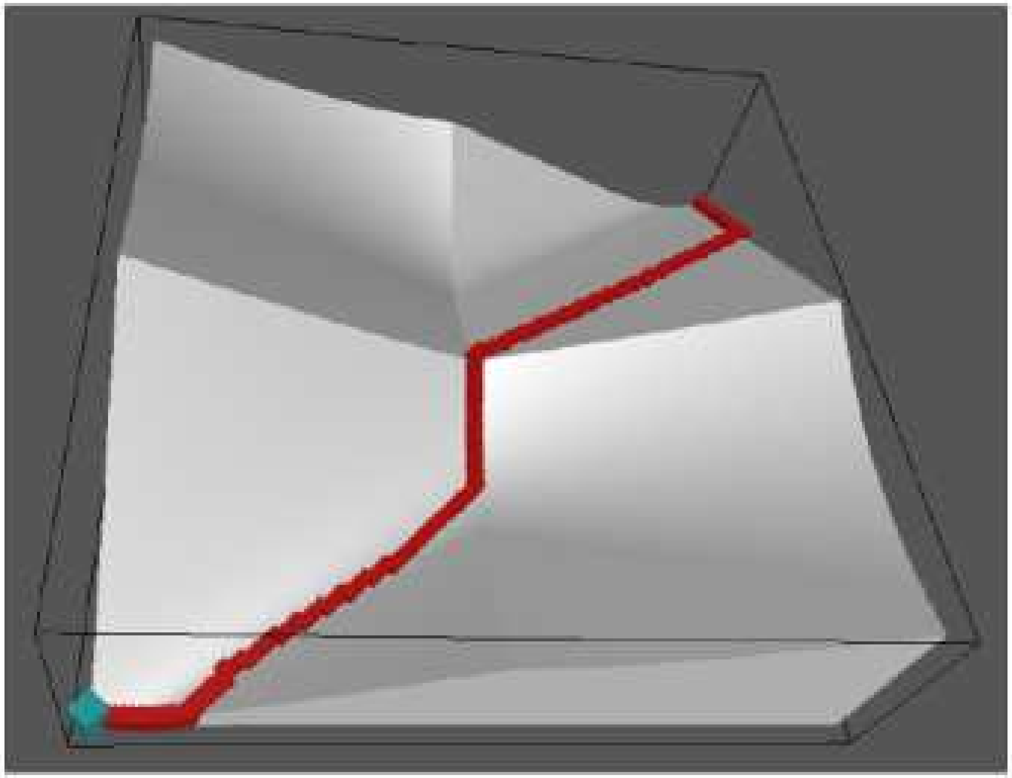
THALES

Extended Fréchet Distance between 2 geodesic paths computable by « fast marching / shortest path » algorithms

$$d_{geo-path}\left(R_1, R_2\right) = \underset{\alpha,\beta}{Inf}\left\{\int_0^1 d_{geo}\left(R_1(\alpha(t)), R_2(\beta(t))\right)dt\right\}$$



« Fast Marching » Algorithm



THALES

Computation of « Median Curves » in Covariance Matrix Manifold

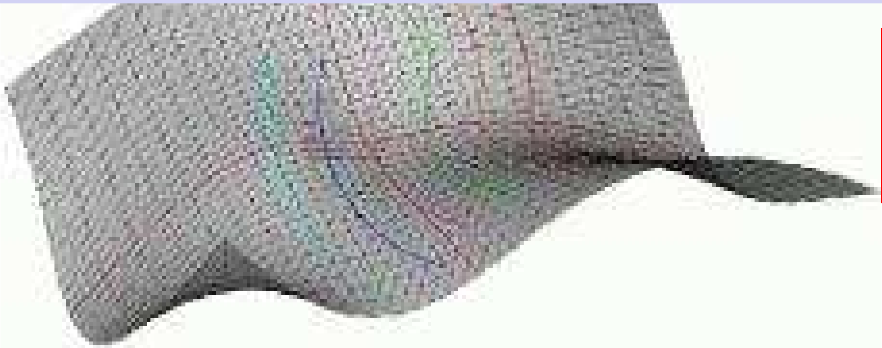
Ordered CFAR for Non-Stationary Signal:

$$d_{geo-path}\left(R_1^{Time-Serie}, R_2^{Time-Serie}\right) = \underset{\alpha,\beta}{Inf}\left\{\int_0^1 d_{geo}\left(R_1(\alpha(t)), R_2(\beta(t))\right)dt\right\}$$

- « Median Curve/Curve » on Matrix Manifold should Minimize :

$$R_{Median}^{Time-Serie} = Arg \underset{R^{Time-Serie}}{Min} \sum_{k=1}^{n} d_{geo-path}\left(R^{Time-Serie}, R_k^{Time-Serie}\right)$$

Approximated solution:
take Path of the set that
minimizes this functional



PhD Thesis 2013-2016

- Non-Stationary Ordered-Statistic CFAR:

$$d_{geo-path}\left(R_{Median}^{Time-Serie}, R_{cell\_under\_test}^{Time-Serie}\right) \geq Threshold$$

THALES

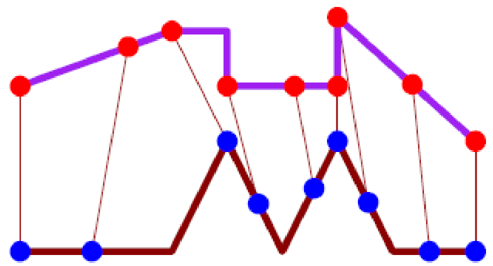
Modified Fréchet distance between paths on Manifold

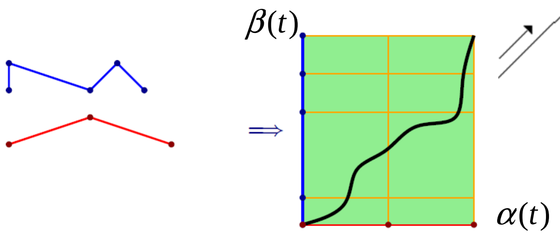
$$d_{Fréchet}(P,Q)=\underset{\alpha,\beta}{Inf}\ \underset{t\in[0,1]}{Max}\{d(P(\alpha(t)),Q(\beta(t)))\}$$

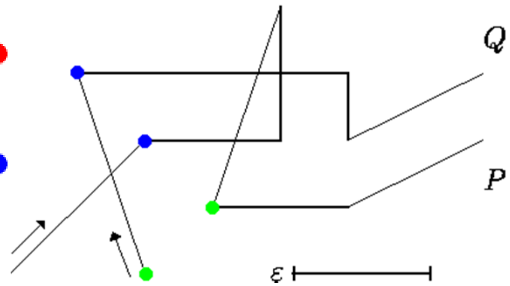
α et $\beta:[0,1]\rightarrow[0,1]$ surjection non décroissante

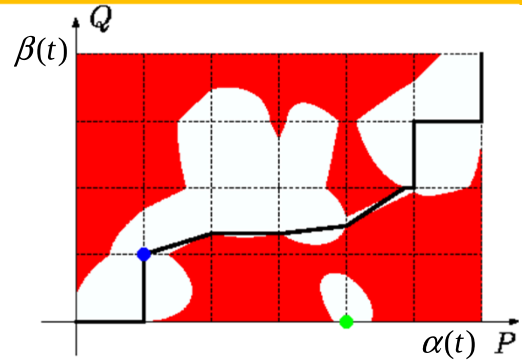
$$d_{THALES}(P,Q)=\underset{\alpha,\beta}{Inf}\int_0^1\{d(P(\alpha(t)),Q(\beta(t)))\}dt$$

α et $\beta:[0,1]\rightarrow[0,1]$ surjection non décroissante









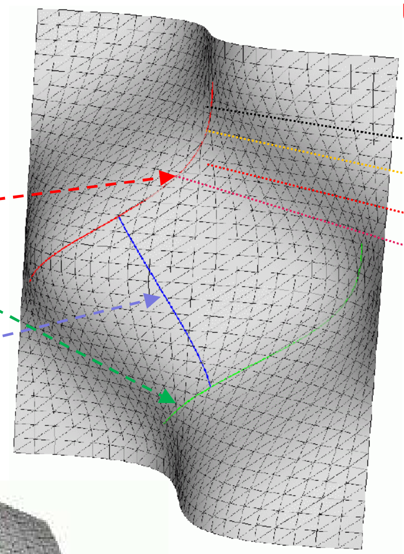
Utilisation de techniques de « Fast Marching »

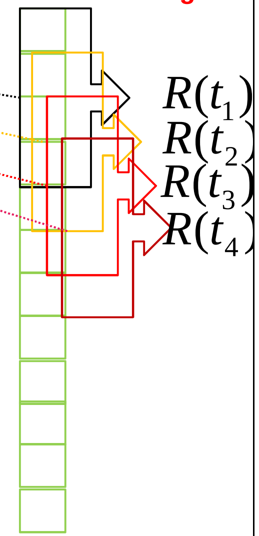
$$\gamma(t)=R_X^{1/2}e^{t\log\left(R_X^{-1/2}R_YR_X^{-1/2}\right)}R_X^{1/2}=R_X^{1/2}\left(R_X^{-1/2}R_YR_X^{-1/2}\right)^tR_X^{1/2}$$

$\gamma(0)=R_X$, $\gamma(1)=R_Y$ et $\gamma(1/2)=R_X\circ R_Y$

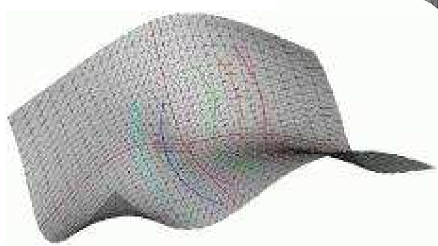
$$d^2(R_1,R_2)=\left\|\log\left(R_1^{-1/2}.R_2.R_1^{-1/2}\right)\right\|^2=\sum_{k=1}^{n}\log^2(\lambda_k)$$

$$\det(R_2-\lambda R_1)=0$$





Statistique de signature temps-Doppler: statistiques de chemins géodésiques sur une variété ou un espace métrique



References

[1] M. Fréchet, “Sur quelques points du calcul fonctionnel », Rendiconti del Circolo Mathematico di Palermo, n°22, p1-74, 1906

[2] M. Fréchet, « L’espace dont chaque élément est une courbe n’est qu’un semi-espace de Banach », Annales scientifiques de l’ENS, 3ème série, tome 78, n°3, p.241-272, 1961

[3] M. Bauer, P. Harms, P.W. Michor, « Almost local metrics on shape space of hypersurfaces in n-spaces », arXiv:1001.0717v4, 27 Jan 2011

[4] A.Chouakria-Douzal and P.N. Nagabhushan, « Improved Fréchet distance for time series », in Data Sciences and Classiffication, Springer 2006, pp 13-20

[5] P.W. Michor and D. Mumford, « An overview of the Riemannian metrics on spaces of curves using the Hamiltonnian approach, » Appl. Comput. Harm. Analysis 23 (1): 74-113, 2007

[6] M. Bauer, M. Bruveris, C. Cotter, S. Marsland, P. W. Michor, “Constructing reparametrization invariant metrics on spaces of plane curves”, preprint http://arxiv.org/abs/1207.5965

[7] R. Dahlhaus, « Locally Stationary Processes », Handbook of Statistics, 30, 2012, http://arxiv.org/abs/1109.4174

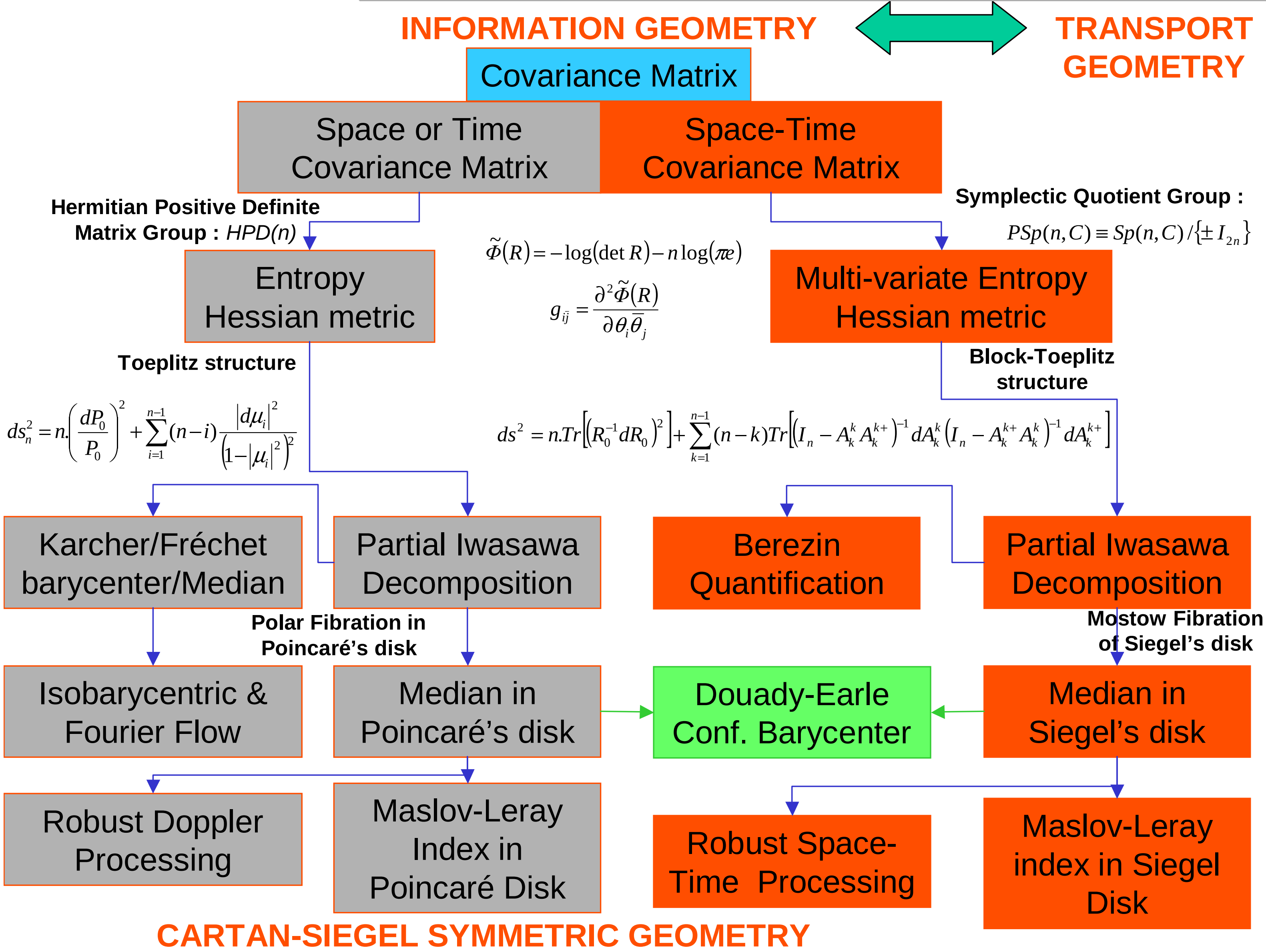
THALES

www.thalesgroup.com

Synthesis of theoretical framework of Structured Covariance Matrices Geometries study

THALES

Synthesis


INFORMATION GEOMETRY
TRANSPORT GEOMETRY
Covariance Matrix
Space or Time Covariance Matrix
Space-Time Covariance Matrix
Hermitian Positive Definite Matrix Group : HPD(n)
Symplectic Quotient Group :
$PSp(n,C) \equiv Sp(n,C)/\{\pm I_{2n}\}$
Entropy Hessian metric
$\tilde{\Phi}(R) = -\log(\det R) - n\log(\pi e)$
$g_{i\bar{j}} = \frac{\partial^2 \tilde{\Phi}(R)}{\partial \theta_i \partial \bar{\theta}_j}$
Multi-variate Entropy Hessian metric
Toeplitz structure
Block-Toeplitz structure
$ds_n^2 = n.\left(\frac{dP_0}{P_0}\right)^2 + \sum_{i=1}^{n-1}(n-i)\frac{|d\mu_i|^2}{\left(1-|\mu_i|^2\right)^2}$
$ds^2 = n.Tr\left[\left(R_0^{-1}dR_0\right)^2\right] + \sum_{k=1}^{n-1}(n-k)Tr\left[\left(I_n - A_k^k A_k^{k+}\right)^{-1} dA_k^k \left(I_n - A_k^{k+} A_k^k\right)^{-1} dA_k^{k+}\right]$
Karcher/Fréchet barycenter/Median
Partial Iwasawa Decomposition
Berezin Quantification
Partial Iwasawa Decomposition
Polar Fibration in Poincaré's disk
Mostow Fibration of Siegel's disk
Isobarycentric & Fourier Flow
Median in Poincaré's disk
Douady-Earle Conf. Barycenter
Median in Siegel's disk
Robust Doppler Processing
Maslov-Leray Index in Poincaré Disk
Robust Space-Time Processing
Maslov-Leray index in Siegel Disk
CARTAN-SIEGEL SYMMETRIC GEOMETRY

General Diffeomorphism for TBTHPD Matrices

$$ds^2 = n.Tr\left[\left(R_0^{-1}dR_0\right)^2\right] + \sum_{k=1}^{n-1}(n-k)Tr\left[\left(I_n - A_k^k A_k^{k+}\right)^{-1} dA_k^k \left(I_n - A_k^{k+} A_k^k\right)^{-1} dA_k^{k+}\right]$$

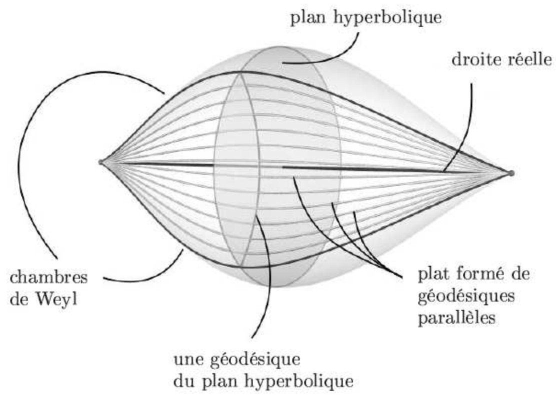
$$ds_m^2 = d\theta^{(m)+}\left[g_{ij}\right]d\theta^{(m)} = m.\left(d\log\left(P_0\right)\right)^2 + \sum_{i=1}^{m-1}(m-i)\frac{\left|d\mu_i\right|^2}{\left(1-\left|\mu_i\right|^2\right)^2}$$

$$\left(R_0, A_1^1, \ldots, A_{n-1}^{n-1}\right) \in THPD_m \times SD^{n-1}$$

$$SD = \left\{Z \,/\, ZZ^+ < I_m\right\}$$

$$R_0 \rightarrow \left(\log\left(P_0\right), \mu_1, \ldots, \mu_{m-1}\right) \in R \times D^{m-1}$$

$$D = \left\{z \,/\, zz^* < 1\right\}$$



Spatio - Doppler State coded by : $\in R \times D^{m-1} \times SD^{n-1}$

Main Theorem : Poincaré Everywhere

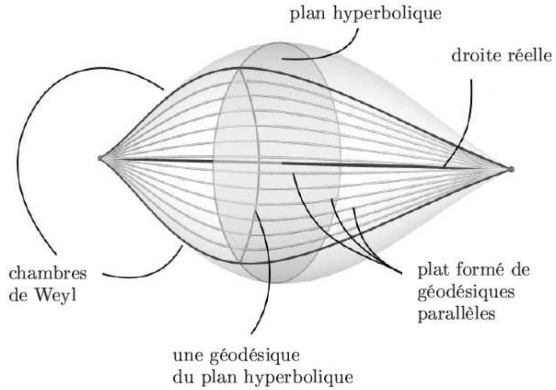
$$\left(R_0, A_1^1, ..., A_{n-1}^{n-1}\right) \in THPD_m \times SD^{n-1}$$

$$SD = \left\{ Z \,/\, ZZ^+ < I_m \right\}$$

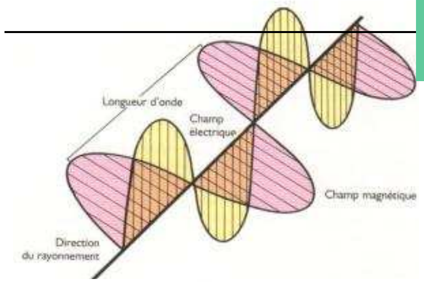
$$R_0 \rightarrow \left(\log(P_0), \mu_1, ..., \mu_{m-1}\right) \in R \times D^{m-1}$$

$$D = \left\{ z \,/\, zz^* < 1 \right\}$$

Compact Hadamard Space:
R x Poly-Poincaré Disk x Poly-Siegel Disk



$$\text{Spatio - Doppler State coded by} : \in R \times D^{m-1} \times SD^{n-1}$$

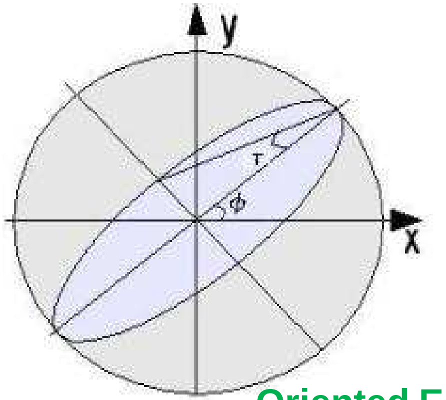


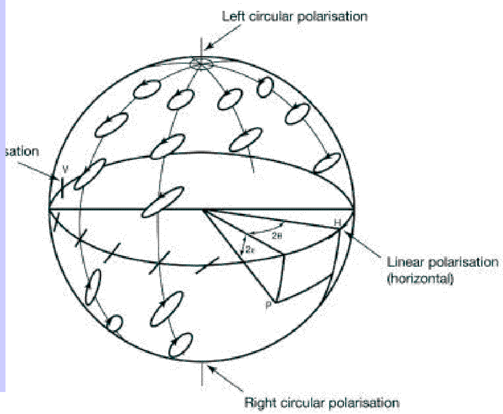
$$\text{EM State coded by} : \in R \times S^1 \times D^{m-1} \times SD^{n-1}$$

$$\vec{E} = \begin{bmatrix} z_H \\ z_V \end{bmatrix} \Rightarrow \vec{S} = \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} |z_H|^2 + |z_V|^2 \\ |z_H|^2 - |z_V|^2 \\ 2.\mathrm{Re}\left(z_V z_H^*\right) \\ 2.\mathrm{Im}\left(z_V z_H^*\right) \end{bmatrix} = s_0 . \begin{bmatrix} 1 \\ \cos(2\tau).\cos(2\phi) \\ \cos(2\tau).\sin(2\phi) \\ \sin(2\tau) \end{bmatrix}$$

$$\text{with} \quad \phi = \frac{\arctan(s_2 / s_1)}{2} \quad \text{and} \quad \tau = \frac{\arctan\left(s_3 / \sqrt{s_1^2 + s_2^2}\right)}{2}$$

$$\text{Polarimetry Poincaré Model} : \{s_0, \phi, \tau\}$$





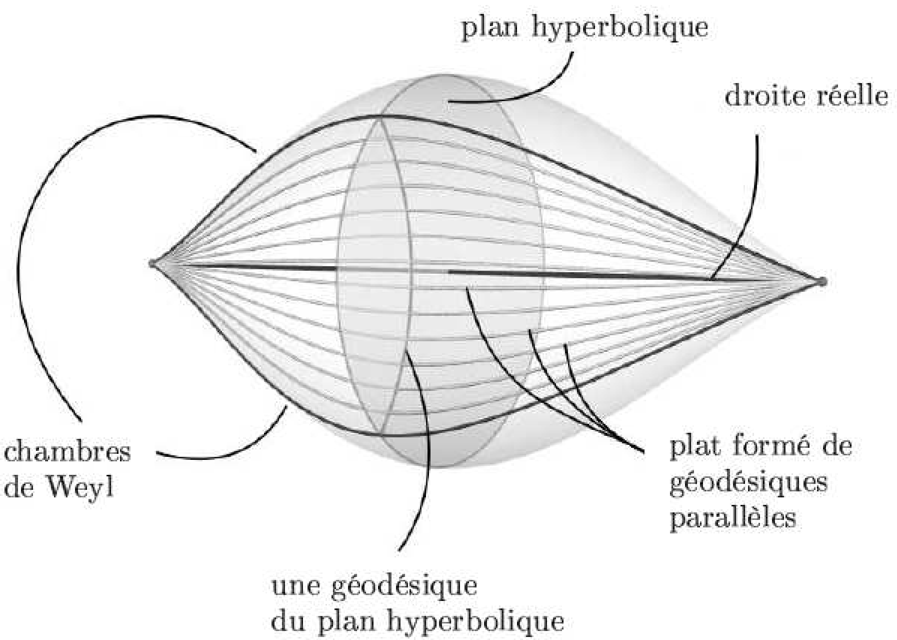
$$\text{Polar State coded by} : \in R \times S^1$$

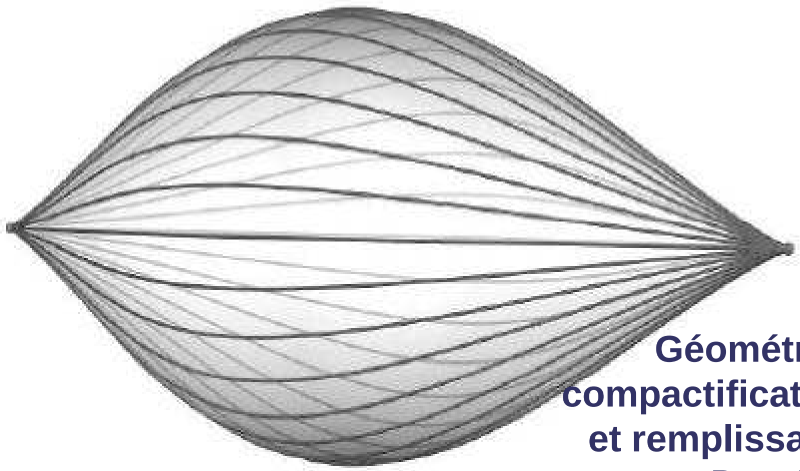
Oriented Ellipse

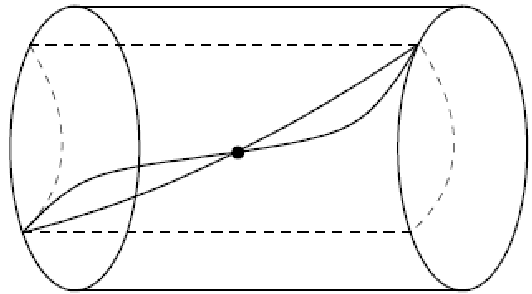
Polarimetry Poincaré Model: $R_+^* x S^1$

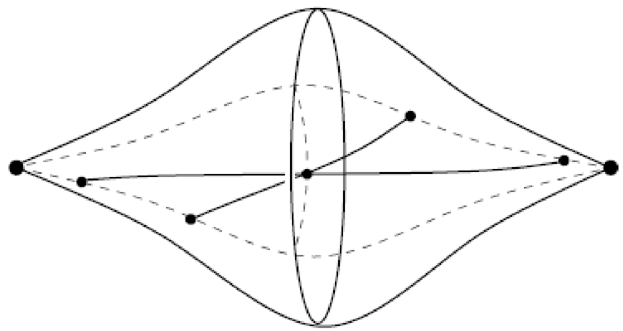
THALES

Hadamard Compactification

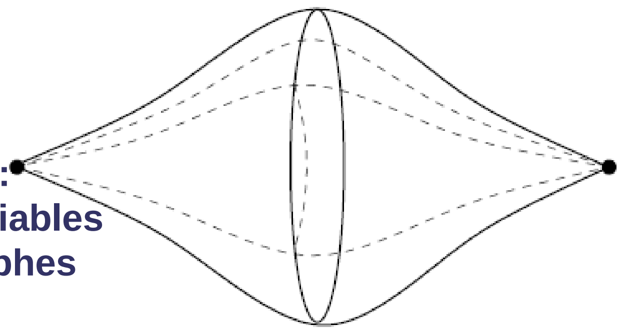








From the cylindrical representation of
we contract the caps and the side,
and blow-up the two circular corners



Toeplitz Hermitian PD Matrices: Simple case n=2

$$\Omega = \begin{bmatrix} h & a-ib \\ a+ib & h \end{bmatrix} > 0$$

$$\det \Omega > 0 \Leftrightarrow h^2 > a^2 + b^2$$

Toeplitz Hermitian Positive Definite

Scale Parameter

Shape Parameter

$$\{\log(h), \mu\} \quad \text{with} \quad \mu = \frac{a+ib}{h}$$

$$\log(h) \in R$$

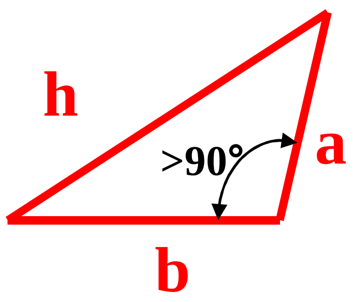
$$\mu \in D = \{z / |z| < 1\}$$

$$\Omega = h.\begin{bmatrix} 1 & \mu^* \\ \mu & 1 \end{bmatrix} > 0$$

$$h \in R_+^*$$

$$\mu = \frac{a+ib}{h} \in D = \{z / |z| < 1\}$$

D : Poincaré Unit Disk



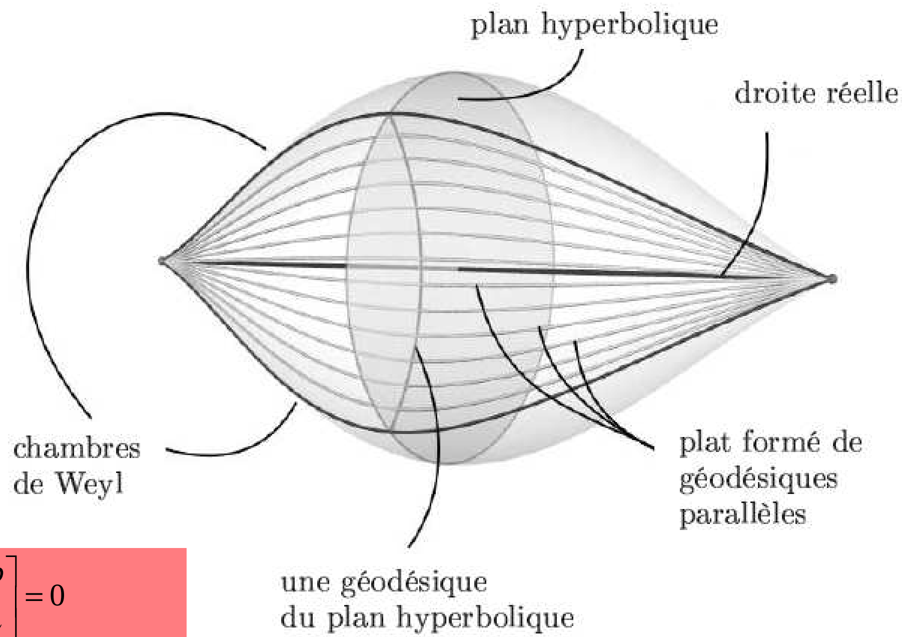
Ambligone Triangle
$h^2>a^2+b^2$

$$\det \begin{bmatrix} h-\lambda & a-ib \\ a+ib & h-\lambda \end{bmatrix} = 0$$

$$\Rightarrow \lambda^2 - 2h\lambda + \left(h^2 - \left(a^2 + b^2\right)\right) = 0$$

$$\Rightarrow \lambda = h \pm \sqrt{a^2 + b^2}$$

$$\lambda > 0 \Leftrightarrow h^2 > a^2 + b^2$$



Hadamard Compactification

Short Summary of the approach

$$\Omega = \begin{bmatrix} h & a-ib \\ a+ib & h \end{bmatrix}$$

Toeplitz HPD matrix

New parameterization $\{h,\mu\} \in R_+^* \times D$

$$\Omega = h.\begin{bmatrix} 1 & \mu^* \\ \mu & 1 \end{bmatrix} > 0$$

$$\Rightarrow \mu = \frac{a+ib}{h} \in D = \{z \in C \,/\, |z| < 1\}$$

Hadamard Compactification :

$\{\log(h),\mu\} \in R \times D$

$$\{\log(h)_{Mean}, \mu_{mean}\}$$

$$\Rightarrow \Omega_{Mean} = \begin{bmatrix} e^{\log(h_{mean})} & h_{mean}.\mu_{mean}^* \\ h_{mean}.\mu_{mean} & e^{\log(h_{mean})} \end{bmatrix}$$

$\{\log(h)_{Mean}, \mu_{mean}\}$ **FRECHET Barycenter in Metric Space**

$$= \arg \underset{\{\log(h),\mu\}}{Min} \sum_{i=1}^{N} d^{\,p}_{geodesique}\left(\{\log(h_i),\mu_i\},\{\log(h),\mu\}\right)$$

$$ds^2 = \left(\frac{dh}{h}\right)^2 + \frac{|d\mu|^2}{\left[1-|\mu|^2\right]^2} = \begin{bmatrix} d\log(h) \\ d\mu \end{bmatrix}^+ \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\left[1-|\mu|^2\right]^2} \end{bmatrix} \begin{bmatrix} d\log(h) \\ d\mu \end{bmatrix}$$

www.thalesgroup.com

Miscellaneous

THALES

www.thalesgroup.com

Quantization of Complex Symmetric Spaces by Berezin

THALES

Classical Symmetric Bounded Domains

Symmetric bounded domains in C^n are particular cases of symmetric spaces of noncompact type.

Elie Cartan has proved that there is :

- **4 types of classical symmetric bounded domains**
- **2 exceptional types (group of motion E6 and E7)**

Classical Symmetric Bounded Domains (extension of Poincaré Disk)

Z : rectangular Complex Matrix

$$ZZ^+ < I \quad (^+ : transposed - conjugate)$$

$\text{Type I} : \Omega^{I}_{p,q}$ Complex matrices with p rows and q columns

$\text{Type II} : \Omega^{II}_{p}$ Complex symmetric matrices of order p

$\text{Type III} : \Omega^{III}_{p}$ Complex skew-symmetric matrices of order p

$\text{Type IV} : \Omega^{IV}_{n}$ Complex matrices with n columns and 1 row such that :

$$\begin{cases} \left|ZZ^t\right| < 1 \\ 1 + \left|ZZ^t\right|^2 - 2ZZ^+ > 0 \end{cases}$$

THALES

Kernel function of Symmetric Bounded Domains

- **Luo-Geng Hua has computed the kernel functions for all classical domains :**

$$K(Z,W^*)=\frac{1}{\mu(\Omega)}\det(I-ZW^+)^{-\nu}$$

$$\text{for}\begin{cases}\text{Type I}:\Omega^I_{p,q}\ ,\nu=p+q\\ \text{Type II}:\Omega^{II}_{p}\ ,\nu=p+1\\ \text{Type III}:\Omega^{III}_{p}\ ,\nu=p-1\end{cases}$$

$$K(Z,W^*)=\frac{1}{\mu(\Omega)}\left(1+ZZ^tW^*W^+-2ZW^*\right)^{-\nu}\ \text{for Type IV}:\Omega^{IV}_n\ ,\nu=n$$

where $\mu(\Omega)$ is the euclidean volume of the domain

- **Particular case (p=q=n=1) : Poincaré Unit Disk**

$$\Omega^I_{1,1}=\Omega^{II}_1=\Omega^{III}_1=\Omega^{IV}_1=\{z\in C\ /\ zz^*<1\}$$

$$K(z,w^*)=\frac{1}{(1-zw^*)^2}$$

Analytic automorphisms of Bounded Domains

- **Groups of analytic automorphisms of these domains are locally isomorphic to the group of matrices which preserve following forms:**

$$\text{Type I}: \Omega^{I}_{p,q}\ ,\ AHA^{*} = H, H = \begin{pmatrix} I_p & 0 \\ 0 & -I_p \end{pmatrix}, \det A = 1$$

$$\text{Type II}: \Omega^{II}_{p}\ ,\ AHA^{*} = H, AKA^{t} = K, H = \begin{pmatrix} I_p & 0 \\ 0 & -I_p \end{pmatrix}, K = \begin{pmatrix} 0 & I_p \\ -I_p & 0 \end{pmatrix}$$

$$\text{Type III}: \Omega^{III}_{p}\ ,\ AHA^{*} = H, ALA^{t} = L, H = \begin{pmatrix} I_p & 0 \\ 0 & -I_p \end{pmatrix}, L = \begin{pmatrix} 0 & I_p \\ I_p & 0 \end{pmatrix}$$

$$\text{Type IV}: \Omega^{IV}_{n}\ ,\ AHA^{*} = H, AHA^{t} = H, H = \begin{pmatrix} -I_2 & 0 \\ 0 & I_n \end{pmatrix}$$

- **The group consits of block matrices A (generalization of the fractional linear transformation). For the first 3 types :**

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \Rightarrow gZ = (A_{11}Z + A_{12})(A_{21}Z + A_{22})^{-1}$$

THALES

Analytic automorphisms of Bounded Domains

- All classical domains are circular following from Cartan's general theory, and the point 0 is distinguished for the potential :

$$\Phi\left(Z,Z^{*}\right)=\ln\left[\frac{K\left(Z,Z^{*}\right)}{K\left(0,0\right)}\right]=\log\det\left(I-ZZ^{+}\right)^{-\nu}$$

- Berezin quantization is based on the construction of the Hilbert Space of functions analytic in Ω :

$$\left\langle f,g\right\rangle =c(h)\int f(Z)g(Z)\left[\frac{K(Z,Z^{*})}{K(0,0)}\right]^{-1/h}d\mu(Z,Z^{*})$$

$$c(h)^{-1}=\int\left[\frac{K(Z,Z^{*})}{K(0,0)}\right]^{-1/h}d\mu(Z,Z^{*})$$

$$K(gZ,gZ^{*})\,j(g,z)\,j(g,Z)^{*}=K(Z,Z^{*})\ \text{ with }\ j(g,Z)=\frac{\partial gZ}{\partial Z}$$

THALES

Berezin Quantification of Poincaré Unit Disk

- The most elementary example of Berezian quantification is, in the case of complex dimension 1, given by the Poincaré unit Disk with volume element : $1/2i.(1-|z|^2)^{-2}dz\wedge dz^*$

$$D=\{z\in C/|z|<1\}=SU(1,1)/S^1$$

$$g\in SU(1,1)\ \text{ with }\ g=\begin{bmatrix} a & b \\ b^* & a^* \end{bmatrix}\ \text{ where }\ |a|^2-|b|^2=1$$

$$\text{Kähler potential}: F(z)=-\ln\left(1-|z|^2\right)\Rightarrow F(gz)=2\operatorname{Re}\ln\left(b^*z+a^*\right)+F(z)$$

$$\Rightarrow\frac{\partial^2 F(gz)}{\partial z\partial z^*}=\frac{\partial^2 F(z)}{\partial z\partial z^*}$$

- Map from path on D to automorphy factor :

$$g(0)=b\left(a^*\right)^{-1}\Rightarrow F(g(0))=-\ln\left(1-\left|b\left(a^*\right)^{-1}\right|^2\right)\underset{|a|^2-|b|^2=1}{=}\ln\left(1+|b|^2\right)$$

$$g^{-1}=\begin{pmatrix} a^* & -b \\ -b^* & a \end{pmatrix}\Rightarrow F(g^{-1})=F(g)$$

Berezin Quantification of Siegel Unit Disk

- **Extension for Siegel Unit Disk :**

$$SD_n = \{Z \,/\, ZZ^+ < I\}$$

$$\text{with } \; g = \begin{bmatrix} A & B \\ B^* & B^* \end{bmatrix} \text{ and } g^t J g = J \;\; \text{with } \; J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

$$\text{where } \begin{cases} A^+ A - B^t B^* = I \\ B^+ A - A^t B^* = 0 \end{cases}$$

$$g(Z) = (AZ + B)(B^* Z + A^*)^{-1}$$

$$\text{Kähler potential}: F(z) = -\log\det\left(I - Z^+ Z\right) = -trace \ln\left(I - Z^+ Z\right)$$

$$F(g(Z)) = F(Z) + 2\,\mathrm{Re}\, trace \ln\left(A^* + B^* Z\right)$$

$$\partial\partial^* F\left(g(Z)\right) = \partial\partial^* F(Z)$$

- **The orbit of the matrix Z=0 is the space of matrices of the form :**

$$g(0) = B\left(A^*\right)^{-1} \Rightarrow F(g(0)) = \ln\det\left(I + B^+ B\right) = trace \ln\left(I + B^+ B\right)$$

Special Berezin Coordinates

- For every symmetric Riemannian space, there exist a dual space being compact. The isometry groups of all the compact symmetric spaces are described by block matrices (the action of the group in terms of special coordinates is described by the same formula as the action of the group of motions of the dual domain).

$$\Gamma = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \Rightarrow \Gamma(W) = (A_{11}W + A_{12})(A_{21}W + A_{22})^{-1}$$

$$\text{Isometry}: \Gamma = C\Gamma C^{-1} \text{ with } C = \frac{1}{\sqrt{2}} \begin{pmatrix} I & iI \\ iI & I \end{pmatrix}$$

- Berezin coordinates for Siegel domain :

$$\Gamma = \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix}, \Gamma^{-1} = \begin{pmatrix} A^+ & B^t \\ B^+ & A^t \end{pmatrix}$$

$$\text{or equivalently}: \Gamma\Gamma^+ = I \ , \ \Gamma L \Gamma^t = L \text{ with } L = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

$$g(0) = B(A^*)^{-1} \Rightarrow F(g(0)) = \ln\det(I + BB^+) = trace \ln(I + BB^+)$$

Invariant metric in Special Berezin Coordinates

- Let M be a classical complex compact symmetric space. The invariant volume and invariant metric in terms of special Berezin coordinates have the form :

$$d\mu(W,W^*) = F(W,W^*)\frac{d\mu_L(W,W^*)}{\pi^n}$$

$$ds^2 = \sum_{\alpha,\beta} g_{\alpha,\beta} dW^{\alpha} dW^{\beta*} \quad \text{with} \quad g_{\alpha\beta} = -\frac{\partial^2 \ln F(W,W^*)}{\partial W^{\alpha} \partial W^{\beta*}}$$

$$\text{where} \quad F(W,W^*) = \det(I + WW^+)^{-\nu}$$

- Link with : For arbitrary Kählerian homogeneous space, the logarithm of the density for the invariant measure is the potential of the metric

www.thalesgroup.com

Analysis on Symmetric Cone by Faraut

THALES

Link with Analysis on Symmetric Cones (J. Faraut)

- **Jacques Faraut has published 2 books on Analysis on Symmetric Cones & on Lie Group :**
  - [1] J. Faraut & A. Koranyi, « Analysis on Symmetric cones », Clarendon Press, Oxford, 1994
  - [2] J. Faraut, « Analyse sur les groupes de Lie », Calvage & Mounet, Paris, 2006
- **Harmonic Analysis in the special case of the cone of positive definite matrices in the vector space of all real symmetric matrices plays a fundamental role in :**
  - Number Theory (Minkowski, siegel, Maas,…)
  - Statistics (Wishart, Constantine, James, Muirhead)
  - Physic (study of Lorentz cone)
- **General Case has been studied by :**
  - Gindikin
  - Vinberg
- **Symmetric Cones & Tubes (over them) are example of Riemannian Symmetric Spaces**

Link with Analysis on Symmetric Cones (J. Faraut)

Convex cones :

- **Let *V* be a finite dimensional real Euclidean Space. A subset *C* of *V* is said to be a cone if :**

$$\left.\begin{array}{l} x \in C \\ \lambda > 0 \end{array}\right\} \Rightarrow \lambda x \in C$$

- **The closed dual cone of any cone C is defined by :**

$$C^{\#} = \{ y \in V \,/\, (x|y) \geq 0, \forall x \in C \}$$

$$\left(C^{\#}\right)^{\#} = C$$

- **The automorphism group $G(\Omega)$ of an open convex cone Ω is defined by :** $G(\Omega) = \{ g \in GL(V) \,/\, g\Omega = \Omega \}$
- $G(\Omega)$ **is a closed subgroup of $GL(\Omega)$, and hence is a Lie Group. The open cone Ω is said to be homogeneous if $G(\Omega)$ acts on it transitively. The open cone is said to be symmetric if it is homogeneous and self-dual :**

$$(gx|y) = (x|g^{*}y)$$

$$g^{*} : \text{the adjoint of an element of } g \in GL(\Omega)$$

Link with Analysis on Symmetric Cones (J. Faraut)

Convex cones :

- **For any proper open convex cone** Ω**:**

$$G(\Omega^*) = G(\Omega)^*$$

$$\text{if } \begin{cases} \Omega^* = \Omega \\ g \in G(\Omega) \end{cases} \Rightarrow g^* \in G(\Omega)$$

- $G(\Omega)^* = G(\Omega)$ **characterizes the symmetric cones**

Characteristic Function of a Cone :

- **Let** Ω **be a proper open convex cone, its characteristic function is :**

$$\varphi(x) = \int_{\Omega^*} e^{-(x|y)} dy$$

dy **Euclidean Measure on *V***

$$\forall g \in G(\Omega) \ , \ \varphi(gx) = |\det g|^{-1} \varphi(x)$$

$$gx = \lambda x \ , \ \lambda > 0 \ , \ \varphi(\lambda x) = \lambda^{-n} \varphi(x)$$

The second derivative $D^2 \log \varphi(x)$ **is positive definite at each point**

$$\begin{cases} G_x(u,v) = D_u D_v \log \varphi(x) \\ D_u \phi(x) = \left. \dfrac{d}{dt} \right|_{t=0} \phi(x+tu) \end{cases} \Rightarrow G_x(u,u) > 0$$

THALES

Link with Analysis on Symmetric Cones (J. Faraut)

- **Characteristic Function of a Cone :**
  - Riemannian structure g is given by

$$g = d^2 \log \varphi(x) \quad \text{with}$$

$$d^2 \log \varphi(x) = d^2 \left[\log \int \varphi_u du \right]$$

$$= \frac{\int \varphi_u d^2 \log \varphi_u du}{\int \varphi_u du} + \frac{1}{2} \frac{\iint \varphi_u \varphi_v \left(d \log \varphi_u - d \log \varphi_v \right)^2 dudv}{\iint \varphi_u \varphi_v dudv}$$

Link with Analysis on Symmetric Cones (J. Faraut)

Characteristic Function of a Cone :

- **The adjoint :** $x^* = -\nabla \log \varphi(x)$ with $\left(\nabla f(x) \middle| u\right) = D_u f(x)$
- **The map** $x \in \Omega \mapsto x^* \in \Omega^*$ **is a bijection :** $\left(x \middle| x^*\right) = n$ **and has unique fixed point**

Symmetric Cone as Riemannian Symmetric Space :

- **The bilinear form** $G_x(u,v) = D_u D_v \log \varphi(x)$ **is positive definite, therefore it defines a Riemannian metric on** Ω
- **The cone** Ω **equiped with this metric is a Riemannian Manifold**
- **Since the cone** Ω **is symmetric, the map** $x \mapsto x^* = -\nabla \log \varphi(x)$ **is a bijection and an isometry (the manifold is a Riemannian Symmetric Space given by this isometry)**

$$\left(x^*\right)^* = x$$

$$\varphi(x)\varphi(x^*) = const$$

THALES

Link with Analysis on Symmetric Cones (J. Faraut)

The cone of Positive Definite Symmetric matrices :

- **Inner Product & quadratic form:** $x, y \in Sym(n, R) \quad , \quad (x|y) = Tr(xy)$

  $$\forall \xi \in R^n \;,\; \xi \neq 0 \;,\; Q(\xi) = \left(x \middle| \xi\xi^T\right) > 0$$

- **Let** $\Omega = \Pi_n(R)$ **be the set of positive definite symmetric matrices.** **The set is an open conex cone and is self-dual :** $\Omega^* = \Omega$

- $\Omega = \Pi_n(R)$ **is homegeneous :** $x \in \Omega \Rightarrow \begin{cases} x = gg^T = \rho(g) I_n \\ \rho(g)x = gxg^T \end{cases}$

- **Characteristic function :**

$$x = \rho(g) I_n \Rightarrow \varphi(x) = \left|\det(\rho(g))\right|^{-1} \varphi(I_n)$$

$$\begin{cases} \det(x) = (\det(g))^2 \\ \det(\rho(g)) = \left|\det(g)\right|^{n+1} \end{cases} \Rightarrow \varphi(x) = [\det(x)]^{-\frac{n+1}{2}} \varphi(I_n)$$

$$\log \varphi(x) = -\frac{1}{2}(n+1)\log(\det x) + \log \varphi(I_n)$$

$$\nabla \log\left|\det(x)\right| = x^{-1} \Rightarrow x^* = \frac{1}{2}(n+1)x^{-1}$$

THALES

Link with Analysis on Symmetric Cones (J. Faraut)

Symmetric Cone & Exponential Familly of probability measure

- **Let μ be a positive Borel Measure on euclidean space *V*. Assume that the following integral is finite for all *x* in an open set $\Omega \subset V$:**

$$\varphi(x) = \int e^{-(x|y)} d\mu(y)$$

- **For $x \in \Omega$, consider the probability measure (exponential familly) :**

$$p(x,dy) = \frac{1}{\varphi(x)} e^{-(x|y)} d\mu(y)$$

Then

$$m(x) = \int y p(x,dy) = -\nabla \log \varphi(x)$$

$$(V(x)u|v) = \int (y - m(x)|u)(y - m(x)|v) p(x,dy) = D_u D_v \log \varphi(x)$$

THALES

www.thalesgroup.com

Homogeneous Hyperbolic Affine Hyperspheres by Sasaki

F. Barbaresco

THALES

Affine hyperbolic hypersphere

- **A locally strongly convex hypersurface in the affine space R^{n+1} is called an affine hyperbolic hypersphere if the affine normals through each point of the hypersurface either all intersect at one point, called its center, that is on the convex side.**
- **This class of hypersurfaces was first studied systematically by W. Blaschke in the frame of affine geometry.**
- **E. Calabi raised a conjecture that:**
  - these hypersurfaces are asymptotic to the boundary of a convex cone
  - every non-degenerate cone V determines a hyperbolic affine hypersphere, asymptotic to the boundary of V, uniquely by the value of its mean curvature.

  He proved this conjecture for homogeneous convex cones under some conditions on the action of the automorphism group of the cone.

Homogeneous Affine hyperbolic hypersphere

- The theory of homogeneous convex cones plays a central role.
- Let V be a nondegenerate open convex cone in $R^{n+1}(x)$ and V' be its dual. A(V) means the group of all linear transformations which leave V invariant. The characteristic function of V, is given by the equation:

$$\phi_V(x) = \int_{V'} e^{-\langle x,\xi \rangle} d\xi \quad , \quad x \in V$$

KOSZUL-VINBERG CHARACTERISTIC FUNCTION

with $\langle x, \xi \rangle$ the value of the linear functional ξ at x

- We denote by S_c the level surface of $\phi_V : \{\phi_V(x) = c\}$ which is a noncompact submanifold in V, and by ω the induced metric on S_c
- The Hessian $d^2 \log \phi_V$ defines the metric on V.
- Assuming the cone V is homogeneous under A(V), Sasaki proved that S_c is a homogeneous hyperbolic affine hypersphere and every such hyperspheres can be obtained in this way
- Sasaki remarks that *ω is identified with the affine metric and S_c* is a global Riemannian symmetric space when V is a self-dual cone.

THALES

Proper Affine hyperbolic hypersphere

- **Let S be a hypersurface in *R*$^{n+1}$ and $f : S \to R^{n+1}$ be the imbedding of S. The imbedding f defines a volume bundle valued quadratic form G on S by the equation:**

$$G = \sum_{i,j} \det\left(\frac{\partial^2 f}{\partial y^i \partial y^j}, \frac{\partial f}{\partial y^1}, \dots, \frac{\partial f}{\partial y^n} \right) dy^i dy^j \otimes dy^1 \wedge \dots \wedge dy^n$$

- **in terms of local coordinates $\left(y^1, \dots, y^n\right)$ of S. This is invariant under unimodular affine transformations in *R*$^{n+1}$. If this quadratic form is supposed to be non-degenerate, it defines a pseudoriemannian structure tensor *g* with corresponding volume element *dv(g)*, uniquely defined by the equation: $G = g \otimes dv(g)$**

Proper Affine hyperbolic hypersphere

- **we assume that the set S is locally strongly convex. Then the tensor *g* can be chosen to be positive definite choosing the orientation of S so that G is positive valued.**
- **With this Riemannian metric, called the affine metric, the affine normal is defined to be the vector $\vec{n} = (1/n)\Delta f$ where *Δ* is the Laplace-Beltrami operator with respect to *g*.**
- **For an affine hyperbolic hypersphere with the center at the origin, n satisfies the equation:**

$$\vec{n} = -Hf$$

where H, called the **affine mean curvature**, is a nonzero constant.

- **Calabi proved that the hypersurface *S* is a proper affine hypersphere with its center at the origin and the affine mean curvature *H* if *u* satisfies the equation:**

$$\det\left(\frac{\partial^2 u}{\partial \xi_i \partial \xi_j}\right) = \left(Hu(\xi)\right)^{-n-2}$$

THALES

Non-parametric characterization of an affine hypersphere

- Let $\left(x^1,...,x^{n+1}\right)$ be a linear coordinate system of R^{n+1} and $\left\{x_{n+1}=f\left(x_1,...,x_n\right)\right\}$ be the representation of S as the graph of a locally strongly convex function f for $x=\left(x^1,...,x^n\right)$ ranging in a domain $D\subset R^n$
- Let $\Omega\subset R^n\left(\xi_1,\xi_2,...,\xi_n\right)$ be the image of *D* under the locally invertible mapping $\xi=gradf=\left(f_1,...,f_n\right)$ where $f_i=\dfrac{\partial f}{\partial x^i}$
- We define the function $u\left(\xi_1,...,\xi_n\right)\in\Omega$ by the equation

$$u\left(gradf\,(x)\right)=-f\,(x)+\left\langle x,gradf\,(x)\right\rangle$$

  where $\left\langle .,.\right\rangle$ is the pairing giving the canonical duality.
- *u* is the Legendre transform of f and also the domain *Ω* the Legendre transform of *S* with respect to the coordinates $\left(x^1,...,x^{n+1}\right)$

THALES

Construction of hyperbolic affine hypersphere

- **Theorem [Wu and Sacksteder]**
  - Let *S* be a closed hyperbolic affine hypersphere with center at the origin and the affine mean curvature *H* . the hypersurface *S* is complete (with respect to the affine metric and with respect to the induced metric of the Riemannian metric of R^{n+1}), noncompact, orientable, smooth and locally strongly convex. In this situation we have:
  - Such a surface is the full boundary of some closed convex body and is the graph of a non-negative smooth strictly convex function defined in some hyperplane.
- **By this theorem the hypersurface *S* can be written globally as the set $\{x_{n+1} = f(x_1,...,x_n)\}$, where f is a positive smooth strictly convex function on $\{x_{n+1} = 0\}$**
- **The tangent plane at any point of *S* cannot contain the origin. In other words the affine normal is not tangent to *S*:**
  - the normal vector in Euclidean sense at one point in S is proportional to: $\vec{n}_E = (\xi_1,...,\xi_n,-1)$ with ξ_i the coordinate of the Legendre transformation
  - The affine normal at that point is $\vec{n} = \frac{1}{\rho^2}\left(\frac{\partial \rho}{\partial \xi_1},...,\frac{\partial \rho}{\partial \xi_{n1}}, \rho + \sum_{i=1}^{n} \frac{\partial \rho}{\partial \xi_i}\xi_i\right)$ with $\rho = \det(f_{ij})^{-1/n+2}$
  - Hence: $\langle \vec{n}, \vec{n}_E \rangle = -1/\rho \neq 0$

THALES

Asymptotic to boundary of convex cone

- ***S* is asymptotic to the boundary of a convex cone when the boundary is equal to the set of all asymptotic lines of *S*:**
  - Let $S_k = \{kp \in R^{n+1} / p \in S\}$ then $S_k \cap S_{k'} = 0 \text{ for } k \neq k'$ and $V = \bigcup_{k>0} S_k$ is an open non-degenerate convex cone.
  - *S* is asymptotic to the boundary of V
- **Theorem:**
  - Every closed hyperbolic affíne hypersphere is asymptotic to the boundary of a convex cone. Conversely, every non-degenerate cone V determines a hyperbolic affine hypersphere asymptotic to the boundary of V, and uniquely determined by the value of its mean curvature.
- **Legendre transformation is an isometry with respect to the LN-metric τ (introduced by Loewner & Nirenberg) and the affine metric *g*:**
  - For a negative convex solution u of $\det\left(\frac{\partial^2 u}{\partial \xi_i \partial \xi_j}\right) = \left(Hu(\xi)\right)^{-n-2}$

    Loewner & Nirenberg defined the metric on a bounded convex domain Ω in $R^n(\xi)$

$$\tau = \frac{1}{Hu} d^2 u$$

THALES

Loewner-Nirenberg Metric

- **LN-metric τ (introduced by Loewner & Nirenberg)**

$$\tau = \frac{1}{Hu} d^2 u \qquad \textbf{with} \qquad \det\left(\frac{\partial^2 u}{\partial \xi_i \partial \xi_j}\right) = \left(Hu(\xi)\right)^{-n-2}$$

- **This metric has the projective invariance in the following sense:**
  - Let $A \in SL(n+1, R)$ be a projective transformation which sends Ω onto $A\Omega$. Then A is an isometry with respect to LN-metrics of Ω and $A\Omega$.

- **Legendre transformation is an isometry with respect to the LN-metric τ and the affine metric *g***
  - *Let* $x^{n+1} = f\left(x^1, ..., x^n\right)$ *be an equation of a hyperbolic affine hypersphere.*

Since $\det\left(f_{i,j}, f_1, ..., f_n\right) = f_{ij}$ *the affine metric is written as* $g = \left(\det\left(f_{ij}\right)\right)^{-1/n+2} d^2 f$ *but* $d^2 f(x) = d^2 u(\xi)$ *at the corresponding points x and ξ by the Legendre transformation. Hence:*

$$\tau = \frac{1}{Hu} d^2 u = \rho d^2 u = g \quad \textbf{with} \quad \rho = \det\left(f_{ij}\right)^{-1/n+2}$$

THALES

Homogeneous hyperbolic affine hyperspheres

- **Let *V* be a non-degenerate convex cone in R^{n+1}. First we recall some properties of the characteristic function:**

$$\phi_V(x) = \int_{V'} e^{-\langle x,\xi\rangle} d\xi \quad , \quad x \in V$$

KOSZUL-VINBERG CHARACTERISTIC FUNCTION

  - $\phi_V(x)$ tends to infinity when *x* approaches to any point of the boundary of *V*
  - The measure $\phi_V(x)dx$ is invariant under $A(V)$:

$$\phi_V(Ax) = \phi_V(x)/\det(A) \quad \text{for} \quad A \in A(V)$$

  - $\log \phi_V$ is convex on *V*. Hence $d^2 \log \phi_V$ defines a metric on *V*
  - The level surface of $\phi_V : S_c = \{\phi_V(x) = c\}$ is a noncompact submanifold in *V* called the characteristic surface of *V*.
  - We denote by ω_c the induced metric of $d^2 \log \phi_V$ on S_c

Homogeneous hyperbolic affine hyperspheres

- **Assuming *V* is affinely homogeneous. The characteristic surface S_c is obviously homogeneous with respect to unimodular elements of *A(V)*:**
- **<u>THEOREM (SASAKI):</u> Every characteristic surface S_c is a complete hyperbolic affine hypersphere with mean curvature $ac^2/n+2$ where a is *a* negative constant depending only on *V*.**
- **Proof: Let $\psi(x)=\log\phi_V(x)$. $S_c=\{\psi=\log c\}$ can be written locally as $x^{n+1}=f(x^1,\ldots,x^n)$ by a smooth function f, the coordinate $\left(x^1,\ldots,x^{n+1}\right)$ being chosen such that $\psi_{n+1}\neq 0$.**
  Let u the Legendre transform of f. Since ψ is constant on S_c we have $f_i=\psi_i/\psi_{n+1}$ on S_c.
  By the definition: $u=-\left(f\psi_{n+1}+\sum_{i=1}^{n}x^i\psi_i\right)/\psi_{n+1}$
  But
  $$\psi(kx)=\psi(x)-(n+1)\log k\ ,\ \forall x\in V,\forall k>0$$
  Hence $$\sum_{\alpha=1}^{n+1}x^\alpha\psi_\alpha(x)=-(n+1)$$
  Therefore $u=(n+1)/\psi_{n+1}$

Homogeneous hyperbolic affine hyperspheres

- **Consider**

$$\det\left(-\psi_{n+1} f_{ij}\right) = \det\left(\psi_{ij} - \frac{\psi_{in+1}\psi_j + \psi_i\psi_{jn+1}}{\psi_{n+1}} + \frac{\psi_i\psi_j\psi_{n+1n+1}}{\psi_{n+1}^2}\right), \ 1 \le i, j \le n$$

$$\det\left(-\psi_{n+1} f_{ij}\right) = \frac{1}{\psi_{n+1}^2} \begin{vmatrix} \psi_{ij} & \psi_{in+1} & \psi_i \\ \psi_{n+1j} & \psi_{n+1n+1} & \psi_{n+1} \\ \psi_j & \psi_{n+1} & 0 \end{vmatrix}$$

- **Set** $\Phi(x) = \begin{vmatrix} \psi_{\alpha\beta} & \psi_\alpha \\ \psi_\beta & 0 \end{vmatrix}, \ 1 \le \alpha, \beta \le n+1$

- **We have** $\Phi(x) = (\det A)^2 \Phi(Ax) \ , \ A \in A(V)$

- **Hence** $\Phi(x) = b\phi^2(x)$ **by the homogeneity for some constant *b* which depends only on *V* itself. This means**

$$\psi_{n+1}^2 \det\left(-\psi_{n+1} f_{ij}\right) = \Phi(x) = bc^2 \ \text{ on } \ S_c$$

Remark: $u = \dfrac{n+1}{\psi_{n+1}}$

- **Since** $\det\left(f_{ij}\right) = \det\left(\dfrac{\partial^2 u}{\partial\xi_i \partial\xi_j}\right)^{-1}$**, then** $\det\left(\dfrac{\partial^2 u}{\partial\xi_i \partial\xi_j}\right) = \left(-\dfrac{n+1}{u}\right)^{n+2} \dfrac{1}{bc^2}$

- **THEOREM: Let *S* a complete hyperbolic affine hypersphere with its center at the origin which is homogeneous under the subgroup *G* of the unimodular group. Let V, the convex cone to whose boundary the hypersurface S is asymptotic: $V=\bigcup_{k>0} S_k$.**
  Let $\tilde{G}=G\times R^{+}$, The element $g=\left(\bar{g},t\right)\in\tilde{G}$ acts on *V* by $g=t.\bar{g}(x)$
  ***V* is homogeneous under $\tilde{G}$.**
  Then *S* is a characteristic surface of *V*.

Proof:

Let function γ on V by the equation $\gamma(x)=k^{-n-1}\text{ for }x\in S_k$

Since $S_k\cap S_{k'}=\phi\text{ for }k\neq k'$ ***, γ is well defined.***

Then $\gamma(Ax)=\gamma(x)/\det(A)$ ***for A e G.***

Therefore, by the homogeneity, $\gamma=b\phi_V$ ***for some nonzero constant b.***

THALES

Homogeneous hyperbolic affine hyperspheres

- ***2 previous Sasaki 's Theorems prove that the classification of homogeneous hyperbolic affine hyperspheres is reduced to the classification of homogeneous convex cones:***
  - *Rothaus, O. S., The construction of homogeneous convex cones, Ann. of Math., 83 , pp. 358-376., 1966*
  - *Vinberg, E. B., The theory of convex homogeneous cones, Trans. Moscow Math. Soc, 12 (1963), 340-403*
- ***E.B. Vinberg has defined an inductive method producing all homogeneous convex cones :***
  - *From a given convex cone V_i in $R^{n+1}(x)$, one can construct another homogeneous cone V in $R(t) \times R^m(y) \times R^{n+1}(x)$ by the equation :*

$$V = \{(t, y, x) / t > y^t h^{-1}(x) y\}$$

where h is a linear mapping on R^{n+1} whose values are real symmetric positive-definite matrices of order m and, corresponding to each element B of some transitive subgroup of A{V_1), there exists a matrix $A \in GL(m, R)$ such that $A^t h(x) A = h(Bx)$. This method can be transposed to obtain all projectively homogeneous bounded convex domains or all homogeneous hyperbolic aífine hyperspheres

Homogeneous hyperbolic affine hyperspheres

- **THEOREM: Suppose V is homogeneous. Then the metric ω_c is identified with the affine metric g up to a constant factor.**

Proof :

Let $\psi(x) = \log \phi_V(x)$. $S_c = \{\psi = \log c\}$ *can be written locally as* $x^{n+1} = f(x^1,\ldots,x^n)$ *by a smooth function* f *, the coordinate* $(x^1,\ldots,x^{n+1})$ *being chosen such that* $\psi_{n+1} \neq 0$

Since $dx^{n+1} = -\sum_{i=1}^{n} \frac{\psi_i}{\psi_{n+1}} dx^i$

we have $\omega_c = d^2\psi\big|_{S_c} = \sum_{1\leq i,j\leq n} -\psi_{n+1} f_{ij} dx^i dx^j$

But $\psi_{n+1} = (n+1)/u$. *Therefore* $\omega_c = -(n+1)Hg$

The assumption needed is not the homogeneity of V but the condition that the level surface is an affine hypersphere.

Homogeneous hyperbolic affine hyperspheres

- **Remark: The solution u of the equation**

$$\det\left(\frac{\partial^2 u}{\partial \xi_i \partial \xi_j}\right) = \left(Hu(\xi)\right)^{-n-2}$$

is given as a first order logarithmic derivative of the characteristic function, i.e. $u = (n+1)/\psi_{n+1}$

Since the characteristic function ϕ **of the product of convex cones V and W is equal to** $\phi_V \phi_W$ **, the derivative of** ϕ **is written using the derivatives of** ϕ_V **and** ϕ_W **. This gives the composition formula for reducible cones.**

THALES

Homogeneous hyperbolic affine hyperspheres

- **Considering the self-dual cone, to formulate another description of the Legendre transformation we introduce the *-mapping due to Koecher , It is a mapping from V to its dual V defined by the equation:** $\xi = x^* = -(\psi_1(x),\dots,\psi_{n+1}(x))$ for $x \in V$
- **This mapping * has the following properties:**
  - * sets up a one to one correspondence between *V* and *V* and $(Ax)^* = (A^t)^{-1} x^*$ holds for every $A \in A(V)$
  - If *V* is homogeneous, then *V'* is also homogeneous and $\phi_V(x)\phi_{V'}(x^*)$ is constant for all *x*. We denote this constant by κ^2
- **In homogeneous case, the *-image of S_c is also a characteristic surface of *V* which we denote by S'_c. Taking a hyperplane *H* such that $U = H \cap V'$ is a non-empty bounded convex domain, we correspond, for every point $x^* \in S_c^{'}$ the intersection point of the line through the origin and *x** with *U*. Thus we have a mapping from S'_c to *U*.**
- **LEMMA: The mapping * followed by this mapping is defined on S_c and projectively equivalent to every Legendre transformation of S_c.**

Homogeneous hyperbolic affine hyperspheres

- **PROPOSITION: Suppose *V* is homogeneous. Then * is an isometry with respect to the metrics** $d^2 \log \phi_V$ **and** $d^2 \log \phi_{V'}$
- **Proof:** $\psi_V = \log \phi_V , \psi_{V'} = \log \phi_{V'}, x^* = -grad_x \psi_V$

$$d\xi_i(x^*) = -\sum_j \frac{\partial^2 \psi_V}{\partial x^i \partial x^j}(x) dx^j , dx^j(\xi^*) = -\sum_k \frac{\partial^2 \psi_{V'}}{\partial \xi_j \partial \xi_k}(\xi) d\xi_k$$

$$^{**} = I_d \Rightarrow \sum_j \frac{\partial^2 \psi_V}{\partial x^i \partial x^j}(x) . \frac{\partial^2 \psi_{V'}}{\partial \xi_j \partial \xi_k}(x^*) = \delta_i^k$$

$$\left(^*\right)^* \left(d^2 \psi_{V'}\right) = \left(^*\right)^* \left(\sum_{i,j} \frac{\partial^2 \psi_{V'}}{\partial \xi_i \partial \xi_j} d\xi_i d\xi_j \right)$$

$$= \sum_{i,j,k,l} \frac{\partial^2 \psi_{V'}}{\partial \xi_i \partial \xi_j}(x^*) \frac{\partial^2 \psi_V}{\partial x^i \partial x^k}(x) \frac{\partial^2 \psi_V}{\partial x^j \partial x^l}(x) dx^k dx^l$$

$$= d^2 \psi_V$$

THALES

Homogeneous hyperbolic affine hyperspheres

- **Now assume V is a self-dual cone. Then $K(x) = -\partial x^* / \partial x$ is an automorphism of *V* and, moreover, $x^* = K(x)x$**
- **Let $S = S_{K^*}$, * / *S* is an involutive automorphism of *S*.**
  Since $\phi_V(x^*) = \phi_V\big(K(x)x\big) = \phi_V(x)\det K$, we have
  $\det(K(x)) = 1$ for $x \in S$ **.Therefore *K(x)* is an automorphism of S**
- **PROPOSITION: If the homogeneous cone *V* is self-dual, then the characteristic surface S_c is a globally symmetric space.**
- **Proof: For one point $x_0 \in S$ define an automorphism *s* of *S* by**
  $s(x) = K(x_0)^{-1} x^*$ **. Since *K(x)* is symmetric, *s* is an involution by**
  By $s(x_0) = x_0$ and $\dfrac{\partial s}{\partial x}(x_0) = K(x_0)^{-1} \dfrac{\partial x^*}{\partial x}(x_0) = -I_d$
  For a general S_c, it is enough to translate the symmetry of *S* to S_c
  by the mapping $x \in S_x \rightarrow \dfrac{\kappa}{c} x \in S$

Homogeneous hyperbolic affine hyperspheres

- Let $H^{+}(n, K)$ be the cone of positive-definite hermitian symmetric matrices over K = *fields R, C, H* (quaternions) or the Cayley algebra *Ca*. Then the followings are the list of all irreducible self-dual cones *V* and the corresponding globally symmetric spaces *S*

$$V = H^{+}(n,R) \Rightarrow S = SL(n,R)/SO(n)$$

$$V = H^{+}(n,R) \Rightarrow S = SL(n,C)/SU(n)$$

$$V = H^{+}(n,H) \Rightarrow S = SU^{*}(2n)/S_{p}(n)$$

$$V = C(n) \Rightarrow S = SO(1,n-1)/SO(n-1)$$

$$V = H^{*}(2,Ca) \Rightarrow S = \text{the space of type EIV}$$

THALES

Homogeneous hyperbolic affine hyperspheres

- **Corollary: Let be a regular convex cone and let $g = Dd\log\psi$ be the canonical Hessian metric. Then each level surface of the characteristic function ψ is a minimal surface of the Riemannian manifold (Ω, g)**
- **Example: Let be a regular convex cone consisting of all positive definite symmetric matrices of degree n. Then $(D, g = -Dd\log\det x)$ is a Hessian structure on Ω , and each level surface of $\det x$ is a minimal surface of the Riemannian manifold: $(\Omega, g = -Dd\log\det x)$**

www.thalesgroup.com

Kähler-Ricci Flow

THALES

Complex Autoregressive Model & Kähler Geometry

- **Erich Kähler Geometry is given by :**
  - complex Manifold of n dimensions, compact or not, with Kählerian metric, that could be locally given by positive definite Riemannian Form :

$$ds^2 = 2\sum_{i,j=1}^{n} g_{i\bar{j}}.dz^i dz^{\bar{j}}$$

  - Kähler condition : Local Existence of Kähler potential function Φ , (and Pluri-harmonic equivalent) such that :

$$g_{i\bar{j}} = \frac{\partial^2 \Phi}{\partial z^i \partial z^{\bar{j}}}$$

  - Ricci Tensor is given by remarkable Expression [Erich Kähler] :

$$R_{i\bar{j}} = -\frac{\partial^2 \log\left(\det g_{k\bar{l}}\right)}{\partial z_i \partial \bar{z}_j}$$

  - And scalar curvature :

$$R = \sum_{k,l=1}^{n} g^{k\bar{l}}.R_{k\bar{l}}$$

THALES

Kähler metric for Complex Autoregressive Model

- **Inthe framework of Affine Information Geometry, metric is given by Hessian of Entropy :**
  - Entropy for Multivariate Gaussian model of zero mean :

$$\tilde{\Phi}(R) = -\log(\det R) - n\log(\pi e)$$

$$g_{ij} \equiv \frac{\partial^2 \tilde{\Phi}}{\partial H_i \partial H_j} \quad \text{and} \quad H = -R$$

  - In cas of Compex Autoregressive Model of order n, Entropy could be expressed by reflection coefficients :

$$\tilde{\Phi}(R_n) = \sum_{k=1}^{n-1}(n-k).\ln\left[1-|\mu_k|^2\right] + n.\ln\left[\pi.e.\alpha_0^{-1}\right]$$

$$\alpha_n^{-1} = \left[1-|\mu_n|^2\right]\alpha_{n-1}^{-1}$$

$$\text{with} \quad \alpha_0^{-1} = P_0 = \frac{1}{n}\sum_{k=1}^{n}|x_k|^2$$

$$\det(R_n) = \prod_{k=0}^{n-1}\alpha_k^{-1} = \alpha_0^{-n}\prod_{k=1}^{n-1}\left[1-|\mu_k|^2\right]^{n-k}$$

Kähler metric for Complex Autoregressive Model

- **We define « Doppler » metric in case of Complex Autoregressive Model by Hessian of Kähler Potential, where Potential is given as in Information Affine Geometry by Entropy :**
  - Kähler Potential is given by Entropy parametrized by reflection coefficients :

$$\tilde{\Phi}(R_n) = \sum_{k=1}^{n-1} (n-k).\ln\left[1 - |\mu_k|^2\right] + n.\ln\left[\pi.e.\alpha_0^{-1}\right]$$

  - Metric can be explicitly computed :

$$\theta^{(n)} = \begin{bmatrix} P_0 & \mu_1 & \cdots & \mu_{n-1} \end{bmatrix}^T = \begin{bmatrix} \theta_1^{(n)} & \cdots & \theta_n^{(n)} \end{bmatrix}^T$$

$$g_{11} = n\alpha_0^2 = nP_0^{-2} \qquad g_{ij} = \frac{(n-i).\delta_{ij}}{\left(1 - |\mu_i|^2\right)^2}$$

$$ds_n^2 = n.\left(\frac{dP_0}{P_0}\right)^2 + \sum_{i=1}^{n-1} (n-i) \frac{|d\mu_i|^2}{\left(1 - |\mu_i|^2\right)^2}$$

Scalar Curvature of Complex Autoregressive Model

- **We use Ricci Tensor expression given by Erich Kähler in framework of Kähler Geometry**
  - In Kähler Geometry, Ricci Tensor is given by :

$$R_{i\bar{j}} = -\frac{\partial^2 \log\left(\det g_{k\bar{l}}\right)}{\partial z_i \partial \bar{z}_j}$$

  - We can compute Ricci tensor for Complex Autoregressive Model :

$$\begin{cases} R_{11} = -2\dfrac{1}{P_0^2} \\ R_{k\bar{l}} = -2\dfrac{\delta_{kl}}{\left(1-\left|\mu_k\right|^2\right)^2} \quad \text{for } k = 2,\ldots,n-1 \end{cases}$$

  - Its negative scalar curvature is given by :

$$R = \sum_{k,l} g^{k\bar{l}}.R_{k\bar{l}}$$

$$R = -2.\left[\sum_{j=0}^{n-1} \frac{1}{(n-j)}\right] \underset{n\to\infty}{\longrightarrow} -\infty$$

THALES

Autoregressive Model # Kähler-Einstein Metric

- **Previous metric is not a Kähler-Einstein metric, but a close matrix structure**
  - A metric is called Kähler-Einstein metric if its Ricci tensor is proportional to the metric :

$$R_{i\bar{j}} = k_0 . g_{k\bar{l}} \quad \text{with} \quad k_0 : \text{constant} \Rightarrow -\frac{\partial^2 \log\left(\det g_{k\bar{l}}\right)}{\partial z_i \partial \bar{z}_j} = k_0 . \frac{\partial^2 \Phi}{\partial z^i \partial z^{\bar{j}}}$$

  - In case of Kähler-Einstein Metric, Kähler Potential is solution Monge-Ampère Equation :

$$\det(g_{k\bar{l}}) = |\psi|^2 e^{-k_0 \Phi} \quad \text{with} \quad \begin{cases} \Phi : \text{Kähler Potential} \\ \psi : \text{holomorphe function} \end{cases}$$

  - For Complex Autoregressive Model, we have :

$$\left[R_{ij}\right] = B^{(n)}\left[g_{ij}\right] \text{ with } R = Tr\left[B^{(n)}\right] = -2.\left[\sum_{j=0}^{n-1} \frac{1}{(n-j)}\right]$$

$$\text{and where} \quad B^{(n)} = -2 diag\left\{.., (n-i)^{-1}, ..\right\}$$

Study of Calabi & Kähler-Ricci Flows for CAR Model

- **Intrinsic Geometric Flow are of fundamental interest in Mathematic & Physics based on variational Approach :**
  - Kähler-Ricci Flow :

Hilbert Action

$$\int_M R(g)dV(g)$$

Kähler-Ricci Flow

$$\frac{\partial g_{i\bar{j}}}{\partial t} = -R_{i\bar{j}} + \frac{\overline{R}}{n} g_{i\bar{j}}$$

$$\overline{R} = \frac{\int_M R(g)dV(g)}{\int_M dV(g)}$$

  - Calabi Flow :

Calabi Action

$$S(g) = \int_M R^2(g)dV(g)$$

Calabi Flow

$$\frac{\partial g_{i\bar{j}}}{\partial t} = \frac{\partial^2 R}{\partial z_i \partial z_j^*}$$

$$g_{i\bar{j}} = \frac{\partial^2 \phi(z,z^*)}{\partial z_i \partial z_j^*}$$

Calabi Flow

$$\frac{\partial \phi}{\partial t} = R - \overline{R}$$

  - Inegality between both functionals

$$S(g) \geq \left(\int_M R(g)dV(g) \right)^2 / \int_M dV(g)$$

THALES

I. Bakas : Relation between Calabi & Kähler-Ricci Flows

In 2 dimension of the complex variable, Calabi & Kähler-Ricci flows are closely related :

- Calabi & Kähler-Ricci Flows can be related by :

$$\frac{\partial g_{i\bar{j}}}{\partial t} = -R_{i\bar{j}} \Rightarrow \begin{cases} \dfrac{\partial^2 g_{i\bar{j}}}{\partial t^2} = -\dfrac{\partial R_{i\bar{j}}}{\partial t} \\ R_{i\bar{j}} = -\dfrac{\partial^2 \log(\det(g))}{\partial z_i \partial z_j^*} \end{cases} \Rightarrow \begin{cases} \dfrac{\partial^2 g_{i\bar{j}}}{\partial t^2} = \dfrac{\partial^2}{\partial z_i \partial z_j^*}\left(\dfrac{\partial \log(\det(g))}{\partial t}\right) \\ \dfrac{\partial \log(\det(g))}{\partial t} = -\sum_{i,j} g^{i\bar{j}} R_{i\bar{j}} = -R \end{cases}$$

$$\Rightarrow \frac{\partial^2 g_{i\bar{j}}}{\partial t^2} = -\frac{\partial^2 R}{\partial z_i \partial z_j^*}$$

- If second derivative according to Kähler-Ricci flow time is identificated with opposite of second derivative according to Calabi flow time , both flows are equivalent: $\dfrac{\partial^2}{\partial t_R^2} = -\dfrac{\partial t}{\partial t_C}$

THALES

Ionnas Bakas : Flows action on Kähler Potential

Given Kähler metric : $ds^2 = 2e^{\Phi(z,z^*,t)} dz.dz^*$

$$R_{zz^*} = -\frac{\partial^2 \Phi}{\partial z \partial z^*} \qquad R_{zz^*} = -\frac{\partial^2 \log(\det(g))}{\partial z \partial z^*} \quad \text{and} \quad g_{zz^*} = e^{\Phi}$$

Kähler-Ricci Flow

$$\frac{\partial g_{i\bar{j}}}{\partial t} = -R_{i\bar{j}} \Rightarrow \frac{\partial e^{\Phi}}{\partial t} = \frac{\partial^2 \Phi}{\partial z \partial z^*} \Rightarrow e^{\Phi} \frac{\partial \Phi}{\partial t} = \frac{\partial^2 \Phi}{\partial z \partial z^*}$$

$$\Rightarrow \frac{\partial \Phi}{\partial t} = \Delta \Phi \quad \text{with} \quad \Delta . = e^{-\Phi} \frac{\partial^2 .}{\partial z \partial z^*}$$

Calabi Flow

$$\frac{\partial g_{i\bar{j}}}{\partial t} = \frac{\partial^2 R}{\partial z \partial z^*} \quad \text{and} \quad R = \sum_{i,j} g^{i\bar{j}} R_{i\bar{j}} = -e^{-\Phi} \frac{\partial^2 \Phi}{\partial z \partial z^*} = -\Delta \Phi$$

$$\Rightarrow \frac{\partial e^{\Phi}}{\partial t} = -\frac{\partial^2 \Delta \Phi}{\partial z \partial z^*}$$

$$\Rightarrow \frac{\partial \Phi}{\partial t} = -\Delta \Delta \Phi$$

Kähler-Ricci and Calabi Flows for CAR Model

Based on previous Rao metric computation for Autoregressive Model :

- Kähler-Ricci Flow on Reflection Coefficients :

$$1-\left|\mu_i(t)\right|^2 = \left(1-\left|\mu_i(0)\right|^2\right).e^{-\frac{t}{(n-i)}} \quad \text{and} \quad P_0(t) = P_0(0)e^{-\frac{t}{n}}$$

$$ds_n^2(t) = n\left(\frac{dP_0}{P_0(0)}\right)^2 e^{\frac{2t}{n}} + \sum_{i=1}^{n-1}(n-i)\frac{\left|d\mu_i\right|^2}{\left(1-\left|\mu_i(0)\right|^2\right)^2} e^{\frac{2t}{(n-i)}}$$

- Calabi Flow on Reflection Coefficients :

$$1-\left|\mu_i(t)\right|^2 = \left(1-\left|\mu_i(0)\right|^2\right).e^{-\frac{2t}{(n-i)^2}} \quad \text{and} \quad P_0(t) = P_0(0)e^{-\frac{2t}{n^2}}$$

$$ds_n^2(t) = n\left(\frac{dP_0}{P_0(0)}\right)^2 e^{\frac{4t}{n^2}} + \sum_{i=1}^{n-1}(n-i)\frac{\left|d\mu_i\right|^2}{\left(1-\left|\mu_i(0)\right|^2\right)^2} e^{\frac{4t}{(n-i)^2}}$$

THALES