

THALES



14ème école d'été de Peyresq
en traitement du signal et des images

Peyresq, du 30 juin au 6 juillet 2019

Les structures élémentaires de la géométrie de l'information et la métrique de Fisher-Koszul-Souriau : exemples d'applications pour le signal radar

Frédéric Barbaresco, Expert Radar
Segment Leader « SENSING »
Représentant du segment technologique
« traitements, commandes et cognition »
des systèmes terrestres et aériens
Groupe THALES

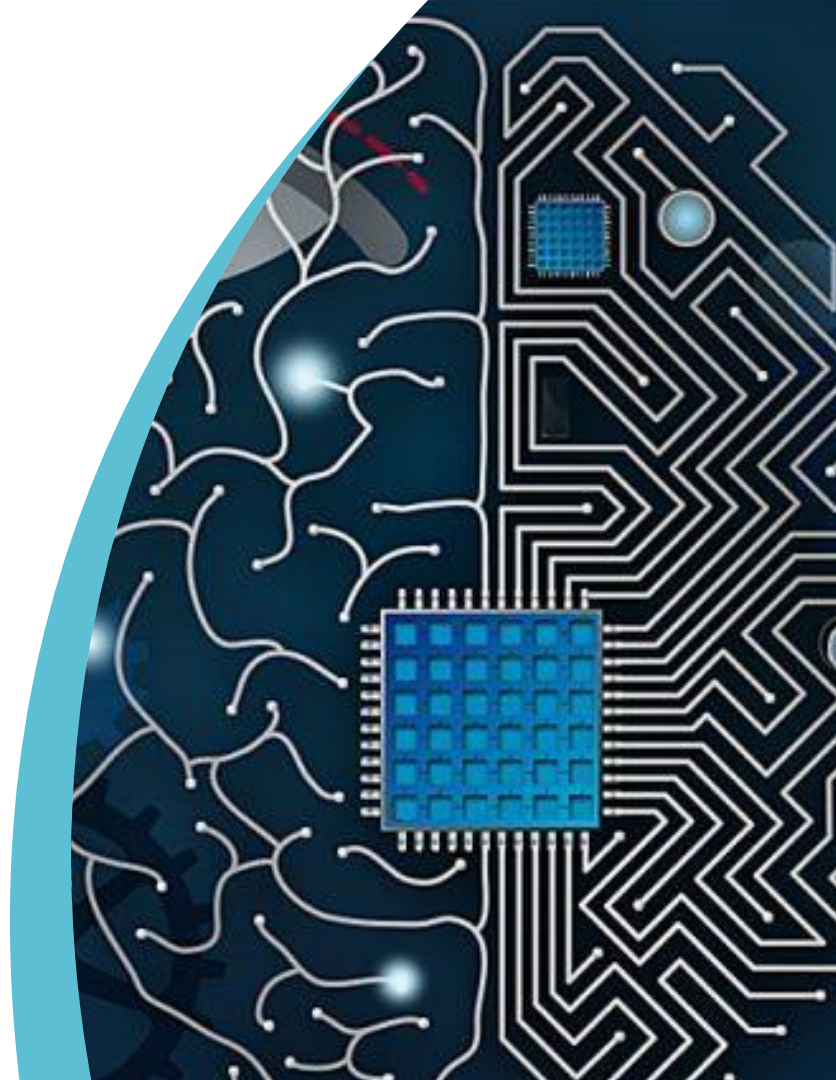
www.thalesgroup.com

ALEAE GEOMETRIA
Géométrie du hasard



Blaise Pascal

Borne de Fréchet-Darmonis-Cramer-Rao et métrique de Fisher



La borne de Fréchet-Darmonis-Cramer-Rao

Borne de la matrice de covariance tout estimateur non biaisé:

➤ La matrice de Fisher et la borne de Fréchet-Darmonis-Cramer-Rao



1945

Information and the Accuracy Attainable in the Estimation of Statistical Parameters

C. Radhakrishna Rao

Borne FDCR

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

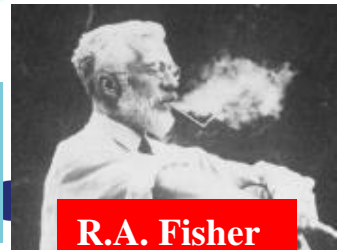
SUR L'EXTENSION DE CERTAINES EVALUATIONS STATISTIQUES AU CAS DE PETITS ECHANTILLONS

par Maurice Fréchet.

Matrice de Fisher

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

1943
(1939 cours IHP)



Axiomatisation de la géométrisation par Nikolai Nikolaevich Chentsov

Travaux de C.R. Rao à partir de 1945 (recherche d'une distance dans l'espace des paramètres des densités invariante par changement de paramétrisation, axiomatisé par N.N. Chentsov



1945

Calyampudi Radhakrishna Rao



1964

(axiomatisatio

Nikolai Nikolaevich Chentsov

➤ Divergence de Kulback-Leibler :

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

➤ Approche de Rao (métrique invariante par changement de paramétrisation)

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

$$w = W(\theta) \Rightarrow ds^2(w) = ds^2(\theta)$$

Page Facebook de Maurice Fréchet (création personnelle)



<https://www.facebook.com/FrechetMaurice/>

Frechet Maurice

Messages Notifications Statistiques Outils de publication Paramètres Aide

Mathématicien français père des probabilités et des statistiques dans des espaces abstraits.

Envoyer un e-mail

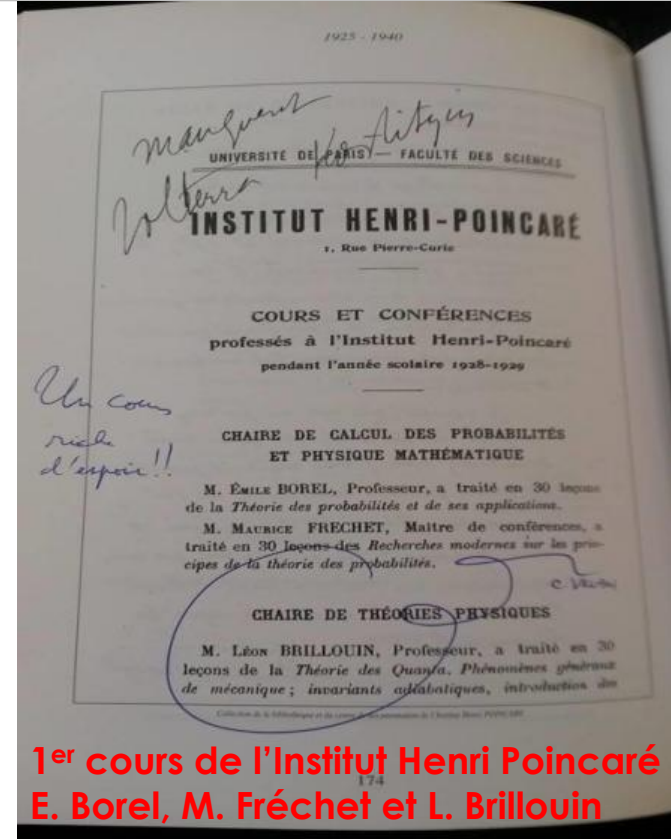
J'aime déjà Déjà abonné(e) Partager

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales. - © Thales - 2015 All rights reserved.

Maurice Fréchet et l'Institut Henri Poincaré

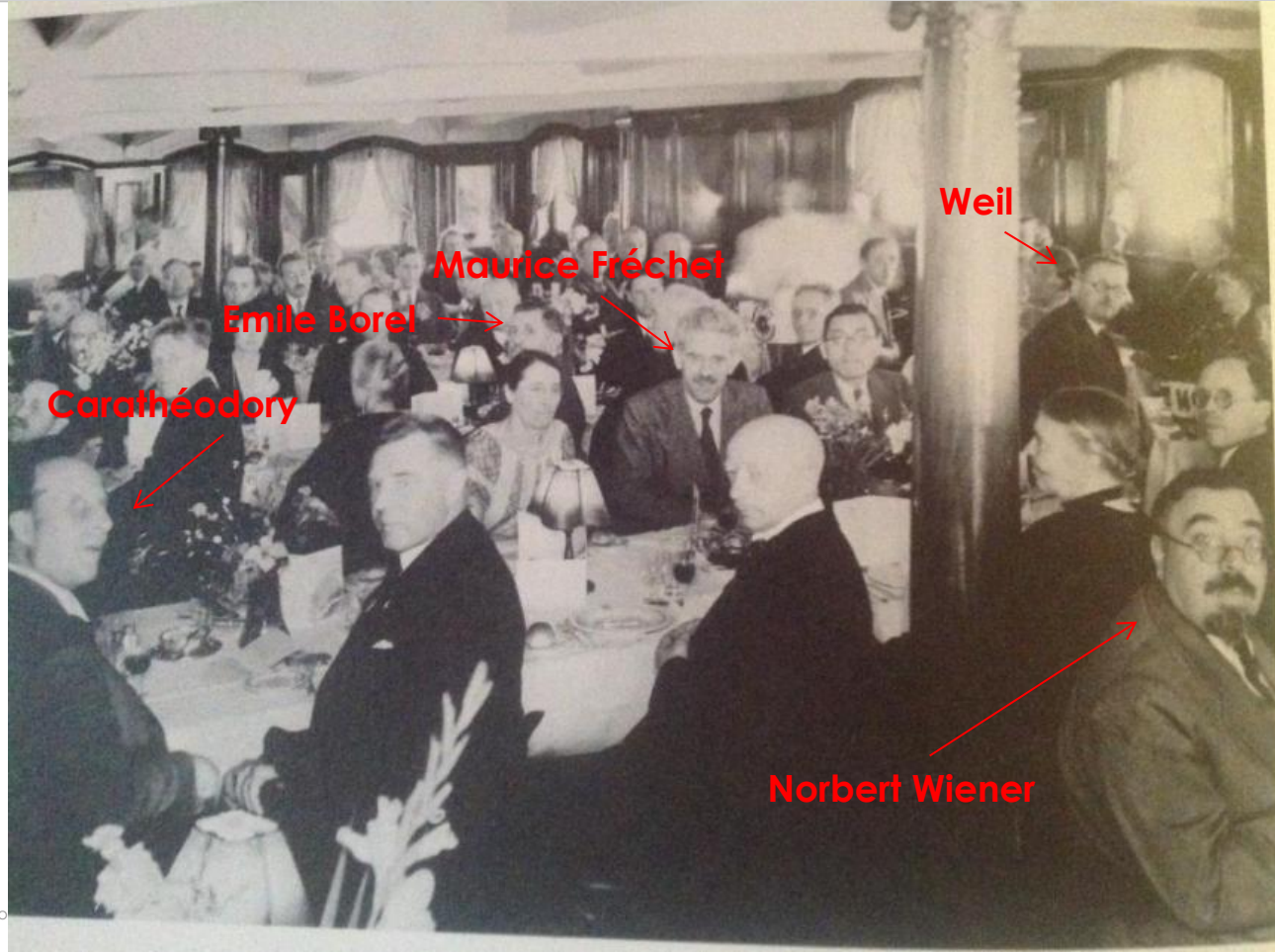
Fréchet à l'IHP

- En 1928 est créé à l'université de Paris l'institut Henri-Poincaré, notamment pour développer l'enseignement des probabilités, sous la direction d'Émile Borel. Maurice Fréchet est nommé maître de conférences de probabilité le 1er novembre 1928 et professeur sans chaire le 19 décembre.
- Années 1930: l'Institut Henri Poincaré devient ainsi un pôle attracteur pour ceux qui veulent étudier les probabilités autour des personnalités de Maurice Fréchet et de Georges Darmois.
- Le 1er octobre 1941 il succède à Émile Borel à la chaire de calcul des probabilités et physique mathématiques jusqu'à sa retraite en 1949.
- Il fut directeur des laboratoires de calculs et de statistiques de l'institut Henri Poincaré.



1^{er} cours de l'Institut Henri Poincaré
E. Borel, M. Fréchet et L. Brillouin

Maurice Fréchet, Borel, Wiener, Weil et Carathéodory Congrès d'Oslo 1936



This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales. - © Thales 2015 All rights reserved.

Maurice Fréchet, Darmois, Cramer et Fisher Congrès Calcul des probabilités – Genève 1938



A Congress on Probability in Geneva (Congrès, Calcul des Probabilités, 1938).
Along the front you see Lévy, R. A. Fisher, and Georges Darmois. On the stairs
are Cramér, Fréchet, and Jean Piaget.

Maurice Fréchet et André Blanc-Lapierre

Colloque international de Calcul des Probabilités - Lyon 1948



Thèse de 1945 d'
André Blanc-Lapierre

- Président du Jury:
Maurice Fréchet
- Rapporteur:
Georges Darmois

Ouverture d'un colloque GRETSI:
A. Blanc-Lapierre, H. Mermoz, P. Aigrain, B. Picinbono



Cinq normaliens (Dugué, Malécot, Ville, Fortet et **Blanc-Lapierre**) suivent des cours auprès de Maurice Fréchet à l'IHP et à l'École normale supérieure.

Fig. 9 Colloque International sur le Calcul des Probabilités, Lyon 1948. First row: Paul Lévy and Maurice Fréchet. On the picture one can find among others J. Doob, R. Fortet, D. Van Dantzig, E. Mourier, J. Kampé de Fériet, A. Blanc-Lapierre... (Photo: © Private collection F. Lederer)

Maurice Fréchet – 12 Square Desnouettes Paris 17ème

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - © Thales 2015 All rights reserved.



Actualités de la borne de Fréchet-Darmois: Ondes gravitationnelles et Boson de Higgs

Borne de Fréchet, matrice de Fisher et détection des ondes gravitationnelles par le réseau LIGO-VIRGO-KAGRA

ACCURACY OF ESTIMATING HIGHLY ECCENTRIC BINARY BLACK HOLE PARAMETERS WITH GRAVITATIONAL-WAVE DETECTIONS

LÁSZLÓ GONDÁN^{1,2}, BENCE KOCSIS¹, PÉTER RAFFAI^{1,2}, AND ZSOLT FREI^{1,2}

¹Eötvös University, Institute of Physics, Pázmány P. s. 1/A, Budapest, Hungary 1117 and

²MTA-ELTE EIRSA "Lendület" Astrophysics Research Group, Budapest, Hungary 1117

Draft version May 31, 2017

ABSTRACT

Mergers of stellar-mass black holes on highly eccentric orbits are among the targets for ground-based gravitational-wave detectors, including LIGO, VIRGO, and KAGRA. These sources may commonly form through gravitational-wave emission in high velocity dispersion systems or through the secular Kozai-Lidov mechanism in triple systems. Gravitational waves carry information about the binaries' orbital parameters and source location. Using the Fisher matrix technique, we determine the measurement accuracy with which the LIGO-VIRGO-KAGRA network could measure the source parameters of eccentric binaries using a matched filtering search of the repeated burst and eccentric inspiral phases of the waveform. We account for general relativistic precession and the evolution of the orbital eccentricity and frequency during the inspiral. We find that the signal-to-noise ratio and the parameter measurement accuracy may be significantly higher for eccentric sources than for circular sources. This increase is sensitive to the initial pericenter distance and component masses. For instance, compared to a $30M_{\odot} - 30M_{\odot}$ non-spinning circular binary, the chirp mass and sky localization accuracy can improve for an initially highly eccentric binary by a factor of ~ 30 (3.6) and ~ 1.5 (6) assuming an initial pericenter distance of $20M_{\text{tot}}$ ($10M_{\text{tot}}$).

Keywords: black hole physics – gravitational waves

Optimal solution to the inverse problem for the gravitational wave signal of a coalescing compact binary

Piotr Jaranowski and Andrzej Krolak
Phys. Rev. D **49**, 1723 – Published 15 February 1994

Borne de Fréchet, matrice de Fisher et détection du Boson de Higgs au LHC

New Ideas for Effective Higgs Measurements

Brehmer, Johann

Second, we use information geometry to understand and optimise Higgs measurements at the LHC. Our novel approach is based on the Fisher information, which encodes the maximum precision with which theory parameters can be measured in an experiment. We develop an algorithm to calculate the Fisher information in LHC processes and compute the information on dimension-six operators in different Higgs signatures. We demonstrate how information geometry lets us improve event selections, determine the most powerful observables, and compare the power of modern multivariate techniques to that of traditional histogram-based analyses.

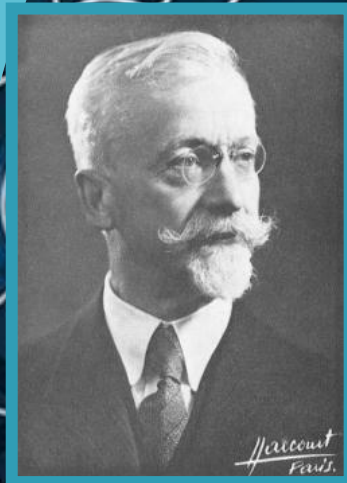
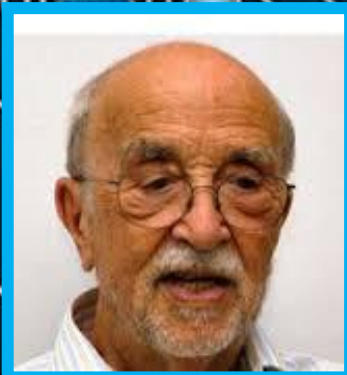
- [4] J. Brehmer, K. Cranmer, F. Kling, and T. Plehn:
Better Higgs boson measurements through information geometry.
Phys. Rev. D95 (7), p. 073002, 2017. arXiv:1612.05261.

« La Physique mathématique, en incorporant à sa base la notion de groupe, marque la suprématie rationnelle...Chaque géométrie – et sans doute plus généralement chaque organisation mathématique de l'expérience – est caractérisée par un groupe spécial de transformations.... Le groupe apporte la preuve d'une mathématique fermée sur elle-même. Sa découverte clôt l'ère des conventions, plus ou moins indépendantes, plus ou moins cohérentes » -
Gaston Bachelard, Le nouvel esprit scientifique, 1934

Thermodynamique des groupes de Lie et métrique de Fisher-Koszul-Souriau: de la géométrie de l'Information de Jean-Louis Koszul à la physique statistique des systèmes dynamiques de Jean-Marie Souriau

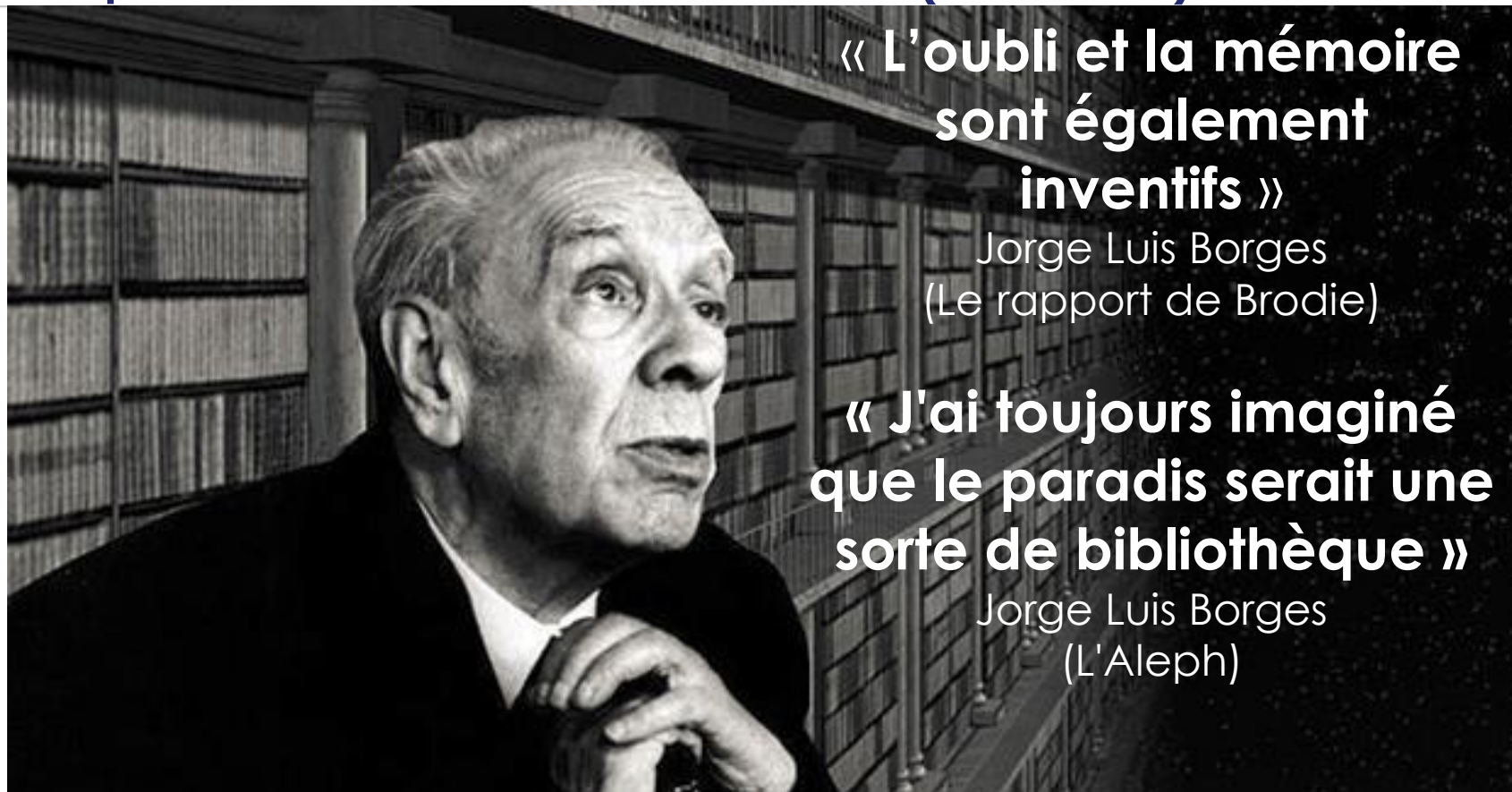
**Tout mathématicien sait qu'il est
impossible de comprendre un cours
élémentaire en thermodynamique.**

Vladimir Arnold



L'oubli et la mémoire: cet exposé est dédié à Jean-Louis Koszul (1921-2018)

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales. - © Thales 2015 All rights reserved.



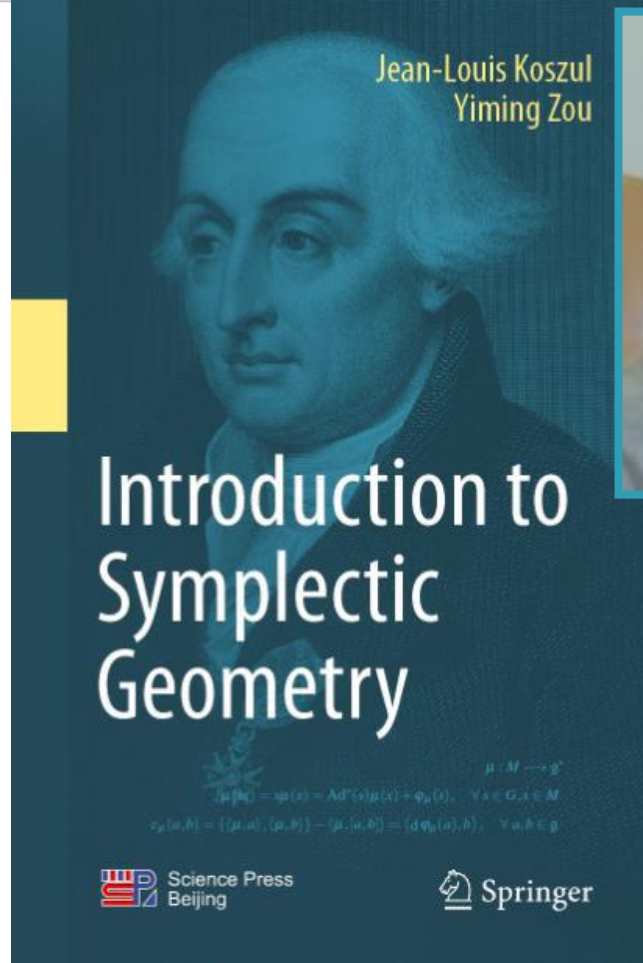
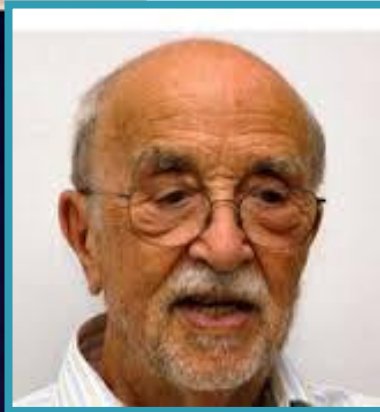
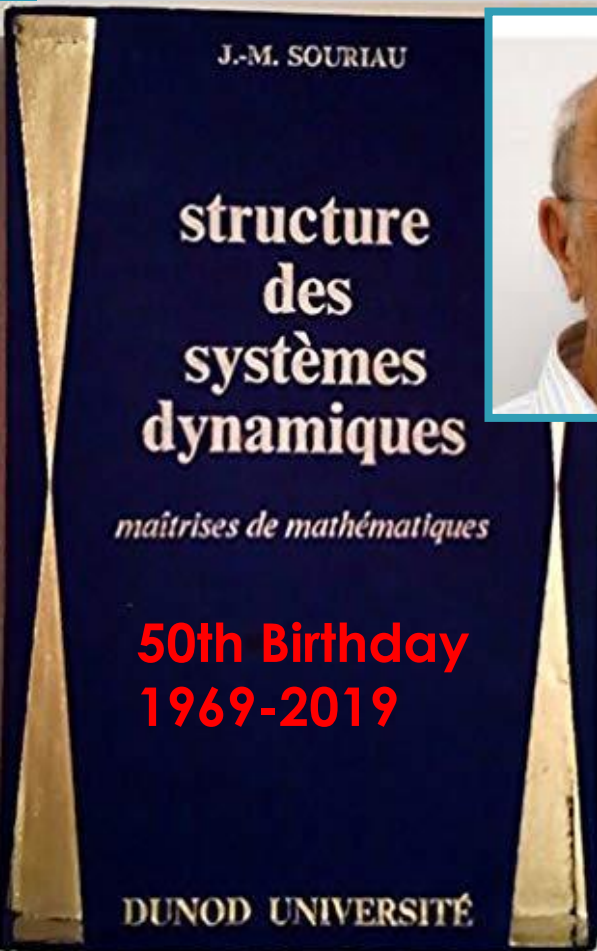
« L'oubli et la mémoire
sont également
inventifs »

Jorge Luis Borges
(Le rapport de Brodie)

« J'ai toujours imaginé
que le paradis serait une
sorte de bibliothèque »

Jorge Luis Borges
(L'Aleph)

Jean-Marie Souriau and Jean-Louis Koszul



1921-2018

**2019 Book,
1 year after
J.L. Koszul
passed away**

THALES

Koszul Book on Souriau Work

Jean-Louis Koszul · Yiming Zou

Introduction to Symplectic Geometry

Forewords by Michel Nguiffo Boyom, Frédéric Barbaresco and Charles-Michel Marle

This introductory book offers a unique and unified overview of symplectic geometry, highlighting the differential properties of symplectic manifolds. It consists of six chapters: Some Algebra Basics, Symplectic Manifolds, Cotangent Bundles, Symplectic G-spaces, Poisson Manifolds, and A Graded Case, concluding with a discussion of the differential properties of graded symplectic manifolds of dimensions (o, n) . It is a useful reference resource for students and researchers interested in geometry, group theory, analysis and differential equations.

$$\mu : M \longrightarrow \mathfrak{g}^*$$

$$\mu(sx) = s\mu(x) = \text{Ad}^*(s)\mu(x) + \phi_\mu(s), \quad \forall s \in G, x \in M$$

$$c_\mu(a, b) = \{ \langle \mu, a \rangle, \langle \mu, b \rangle \} - \langle \mu, [a, b] \rangle = \langle d\phi_\mu(a), b \rangle, \quad \forall a, b \in \mathfrak{g}$$

Jean-Louis Koszul
Yiming Zou

Introduction to Symplectic Geometry

$$\mu : M \longrightarrow \mathfrak{g}^*$$

$$\mu(sx) = s\mu(x) = \text{Ad}^*(s)\mu(x) + \phi_\mu(s), \quad \forall s \in G, x \in M$$

$$c_\mu(a, b) = \{ \langle \mu, a \rangle, \langle \mu, b \rangle \} - \langle \mu, [a, b] \rangle = \langle d\phi_\mu(a), b \rangle, \quad \forall a, b \in \mathfrak{g}$$

Koszul & Souriau: Bedrock of Information Geometry



- For the 50th birthday of Jean-Marie Souriau Book “**Structure des systèmes dynamiques**” published in 1969, and Jean-Louis Koszul Book Translation by Springer “**Introduction to Symplectic Geometry**”, FGS1’19 has celebrated the influence of the Triumvirate Elie Cartan (ENS, 1888), Jean-Louis Koszul (ENS, 1940) and Jean-Marie Souriau (ENS, 1942) on Foundations of Geometric Structure of Information.
- Both Koszul and Souriau were influenced by Elie Cartan works on symmetric homogeneous spaces. Jean-Louis Koszul has developed theory of hessian geometry introducing **Koszul forms** that are fundamental structures in Information Geometry. In parallel Souriau has developed in the framework of Geometrical Mechanics applied for Statistical Mechanics, a **Lie Group Thermodynamics** in Homogeneous Symplectic Manifold. Based on Souriau cocycle, this thermodynamics defines a **generalized Fisher metric** where the Gibbs Maximum Entropy density is covariant with respect to dynamic groups of Physics.

En mémoire à Jean-Louis Koszul 1921-2018

[A] Barbaresco, F. Jean-Louis Koszul and the elementary structures of Information Geometry. In Geometric Structures of Information Geometry; Nielsen, F.; Ed.; Springer: Berlin, Germany, 2018

[B] Barbaresco, F. Koszul Contemporaneous Lectures: Elementary Structures of Information Geometry and Geometric Heat Theory. In Introduction to Symplectic Geometry; Koszul, J.L., Ed.; Springer: Berlin, Germany, 2018.

[C] Barbaresco, F. Jean-Louis Koszul et les Structures Élémentaires de la Géométrie de l'Information; Revue SMAI Matapli; SMAI Editor; Volume 116, pp.71-84, Novembre 2018

[D] Barbaresco, F. Les densités de probabilité « distinguées » et l'équation d'Alexis Clairaut: regards croisés de Maurice Fréchet et de Jean-Louis Koszul, Conférence Histoire de la discipline, GRETSI'17, Juan-Les-Pins, Septembre 2017



Jean-Louis Koszul
Avec Yann Ollivier
GSI'13



Jean-Louis Koszul au colloque « Topologie différentielle »,
Strasbourg 1953 avec Chern, de Rham, Eckmann,
Ehresmann, Godeaux, Hopf, Lichnerowicz, Malgrange,
Milnor, Reeb, Schwartz, Süß, Thom, Libermann et Weil.



Jean-Louis Koszul
Créateur du CIRM
(photo anniversaire CIRM)

En mémoire à Jean-Louis Koszul 1921-2018



KOSZUL



Thèse avec Henri Cartan
(Jean Leray
membre du Jury)



Groupe Bourbaki



Professeur à
Strasbourg et à
Grenoble



Jean-Louis Koszul à GSI'13 aux Mines de Paris



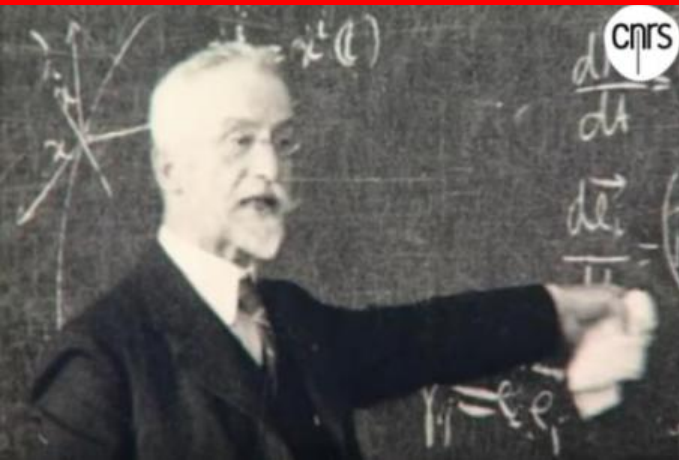
JEAN-LOUIS KOSZUL - DIRECTEUR DU LABORATOIRE, 1978-1981

1869-2019 Elie Cartan 150^{ème} anniversaire



Elie Cartan avec sa femme et ses 4 enfants

Demain, il y a 150 ans le 9 Avril 1869 naissait jour pour jour un esprit, élevé à la chaleur de la FORGE et au son de l'enclume et du marteau de son père Joseph Maréchal Ferrand de Dolomieu.



"des paysans sans prétention qui, au cours de leur longue vie, ont montré à leurs enfants un exemple de travail accompli avec joie et d'acceptation courageuse des fardeaux"
- Elie Cartan



Joseph Cartan (Dolomieu, 1890)

Cartan & Koszul families

The family settled in Dolomieu. **Joseph Cartan was the village blacksmith.** Elie Cartan recalled that his childhood had passed under "**blows of the anvil, which started every morning from dawn**".



Origin of Koszul Family: Ancestor from Poland was engaged in **Napoléon "Grande Armée"**. He came back to France with **Napoléon Army** and Married with Woman from **Strasbourg**

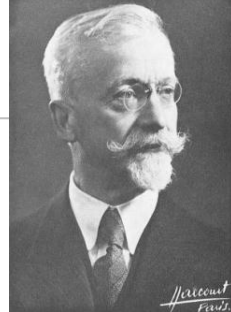
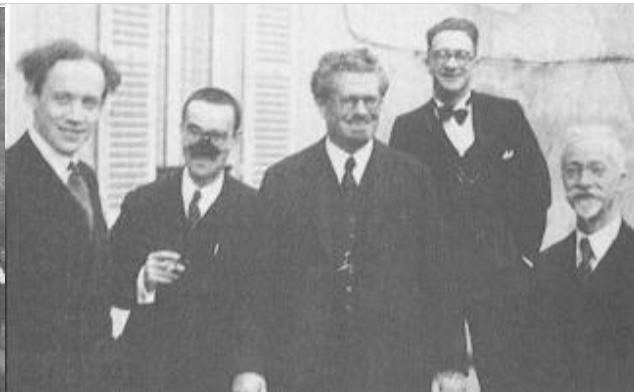
ALLOCUTION DE MONSIEUR HENRI CARTAN

Je n'ai nullement l'intention de faire un « discours », contrairement à ce qu'annonçait le programme de ces journées. Je voudrais simplement évoquer ici brièvement quelques souvenirs qui, avec les années qui passent inexorablement, tendent malheureusement à s'estomper.

Ces souvenirs commencent, il est vrai, avant la naissance de Koszul. En effet, ma mère, dans sa jeunesse, avait été une amie intime de celle qui devait devenir la mère de Jean-Louis Koszul. Il arriva que ces deux amies se marièrent ; l'une épousa un mathématicien connu, l'autre un angliciste non moins connu. Malgré l'éloignement consécutif à leurs mariages, des liens d'amitié subsistèrent, qui expliquent pourquoi, lorsque beaucoup plus tard, au printemps de 1929, j'arrivai à Strasbourg comme jeune chargé de cours à la Faculté des Sciences, je fus reçu dans la famille du professeur Koszul de la Faculté des Lettres. J'ai oublié le menu du repas familial, mais je vois toujours un jeune garçonnet de 8 ans, nommé Jean-Louis, qui évoluait dans l'appartement au milieu de ses grandes sœurs. L'aînée d'entre elles était mariée à un agrégatif de mathématiques que j'avais comme élève à la Faculté. Je ne restai à Strasbourg que quelques mois et perdis donc de vue le jeune Jean-Louis.

Puisque j'ai évoqué le souvenir de ses parents, permettez-moi de nommer aussi le grand-père paternel de Jean-Louis. Je ne l'ai pas connu, certes ; mais comme directeur du Conservatoire de musique de Roubaix-Tourcoing, il joua un rôle historique, car c'est lui qui donna au jeune Albert Roussel, qui venait d'abandonner la carrière navale, les conseils décisifs qui lui permirent de devenir l'un des plus grands compositeurs de musique du début du siècle. On aura l'occasion d'en parler cette année, puisqu'on va célébrer le cinquantenaire de la mort d'Albert Roussel.

Elie Cartan : 1869-1951



Jean CARTAN, December 1, 1906 – March 26, 1932

Marie-Louise BIANCONI, the spouse of Élie CARTAN

February 18, 1880 – May 21, 1950

Élie CARTAN, April 9, 1869 – May 6, 1951

The right half of the same horizontal tombstone reads
Hélène CARTAN, October 12, 1917 – June 7, 1952



Translations of
**MATHEMATICAL
MONOGRAPHS**

Volume 123

Élie Cartan
(1869–1951)

M. A. Akivis
B. A. Rosenfeld



Henri Poincaré's Sorbonne report on Elie Cartan Works



RAPPORT SUR LES TRAVAUX DE M. CARTAN

fait à la Faculté des Sciences de l'Université de Paris.

PAR

H. POINCARÉ.

Si alors on dépouille la théorie mathématique de ce qui n'y apparaît que comme un accident, c'est-à-dire de sa matière, il ne restera que l'essentiel, c'est-à-dire la forme; et cette forme, qui constitue pour ainsi dire le squelette solide de la théorie, ce sera la structure du groupe.

M. CARTAN a fait faire des progrès importants à nos connaissances sur trois de ces catégories, la 1^{ère} la 3^e et la 4^e. Il s'est principalement placé au point de vue le plus abstrait de la structure, de la forme pure, indépendamment de la matière, c'est-à-dire, dans l'espèce, du nombre et du choix des variables indépendantes.

Conclusions.

On voit que les problèmes traités par M. CARTAN sont parmi les plus importants, les plus abstraits et les plus généraux dont s'occupent les Mathématiques; ainsi que nous l'avons dit, la théorie des groupes est, pour ainsi dire, la Mathématique entière, dépouillée de sa matière et réduite à une forme pure. Cet extrême degré d'abstraction a sans doute rendu mon exposé un peu aride; pour faire apprécier chacun des résultats, il m'aurait fallu pour ainsi dire lui restituer la matière dont il avait été dépouillé; mais cette restitution peut se faire de mille façons différentes; et c'est cette forme unique que l'on retrouve ainsi sous une foule de vêtements divers, qui constitue le lien commun entre des théories mathématiques qu'on s'étonne souvent de trouver si voisines.

« A l'Exception d'Henri Poincaré qui écrit peu avant sa mort un rapport sur les travaux d'Elie Cartan à l'occasion de la candidature de celui-ci à la Sorbonne, les mathématiciens français ne voyaient pas l'importance de l'œuvre. »

Paulette LIBERMANN

La géométrie différentielle d'Elie Cartan à Charles Ehresmann et André Lichnerowicz, Géométrie au XXI^{ème} siècle, HERMANN, 2005



**Paulette
Libermann**

1869-2019 Elie Cartan 150th Birthday

- Jean-Louis Koszul organisa l'évènement du centenaire d'Elie Cartan à Grenoble, et c'est Jean-Louis Koszul qui fut un des orateurs du centenaire Elie Cartan organisé en 1969 à Bucarest, comme le rappelle Henri Cartan dans son allocution au colloque Koszul:
 - « Pour terminer, permettez-moi de dire à Koszul que je n'oublie pas l'hommage qu'il a rendu à l'oeuvre d'Élie Cartan en Géométrie différentielle lors de la **célébration, à Bucarest, en 1969**, du centenaire de sa naissance. Ce n'est sans doute pas un hasard si ce **centenaire a aussi été célébré à Grenoble la même année**. Comme toujours, Koszul s'exprima avec la discrétion et le tact que nous lui connaissons, et que nous aimons tant chez lui. C'est de tout coeur que je lui souhaite de poursuivre encore longtemps son activité, toujours pareil à lui-même. » - Henri Cartan
- Koszul, J.L.: **L'oeuvre d'Élie Cartan en géométrie différentielle**, in Élie Cartan, 1869–1951. Hommage de l'Académie de la République Socialiste de Roumanie à l'occasion du centenaire de sa naissance. Comprenant les communications faites aux séances du 4e Congrès du Groupement des Mathématiciens d'Expression Latine, tenu à Bucarest en 1969 (Editura Academiei Republicii Socialiste Romania, Bucharest, 1975) pp. 39–45

Elie Cartan Colloquium 1984

INTRODUCTION

Le séminaire conjoint NSF-CNRS "Elie Cartan et les mathématiques d'aujourd'hui" s'est tenu du 25 au 29 juin à l'Université de Lyon I. Il était centré sur la présentation de thèmes importants de la recherche actuelle en mathématiques et en physique mathématique dans des domaines où Elie CARTAN a joué un rôle de pionnier. Des discussions très animées ont eu lieu à propos de ces thèmes. La partie centrale du programme comprenait vingt-deux conférences qui ont été suivies par un public nombreux et enthousiaste de plus de deux cents mathématiciens et physiciens mathématiciens venus d'au moins dix-sept pays. Les conférences se sont tenues dans le grand amphithéâtre de mathématiques de l'Université de Lyon I, la salle Camille Jordan.

Ce volume regroupe les contributions écrites des conférenciers à l'exception de trois d'entre eux. Nous publions en annexe les résumés que ces auteurs ont bien voulu nous communiquer.

Le programme du séminaire avait été préparé par un Comité Scientifique sous la co-présidence de Shing-shen CHERN et d'Henri CARTAN. Les détails pratiques pour l'organisation du séminaire ont été réglés par un Comité mis sur pied par le Département de mathématiques de l'Université de Lyon I, sous la responsabilité d'Edmond COMBET. L'organisation a été très efficace et a créé une atmosphère dans laquelle la communication mathématique était stimulée. Le professeur GELFAND a reçu un diplôme de docteur honoris causa de l'Université de Lyon I lors de la séance de clôture du séminaire. Sa participation au séminaire, la participation simultanée de trois mathématiciens soviétiques émigrés parmi les plus éminents (Victor KAC du Massachusetts Institute of Technology de Boston, U.S.A., Ilya PIATETSKII-SHAPIRO de Tel-Aviv et Mikhail GROMOV de l'Institut des Hautes Études Scientifiques de Bures-sur-Yvette) ainsi que celle du physicien mathématicien polonais Andrzej TRAUTMAN de Varsovie ont élevé le séminaire au-delà du niveau d'une rencontre entre les écoles américaine et française à celui d'un événement mathématique réellement international.

Special Issue in *Astérisque* en 1984 :

Elie Cartan et les mathématiques d'aujourd'hui

proceedings of ENS Lyon Colloquium from 25th to 29th June 1984 (200 attendees from Mathematics & Physics) .

Among Speakers: Jean-Louis Koszul et Jean-Marie Souriau

Introduction	3
Une lettre d'André WEIL à Henri CARTAN	5
BERGER (Marcel). — La géométrie métrique des variétés riemanniennes (variations sur la formule $a^2 = b^2 + c^2 - 2bc \cos \alpha$)	9
CHERN (Shiing-shen). — Moving frames	67
CHOQUET-BRUHAT (Yvonne). — Causalité des théories de supergravité	79
FEFFERMAN (Charles) and GRAHAM (C. Robin). — Conformal invariants	95
GELFAND (I.M.) and ZELEVINSKY (A.V.). — Representation models for classical groups and their higher symmetries	117
GROMOV (M.). — Isometric immersions of Riemannian manifolds ..	129
GUILLEMIN (Victor). — The integral geometry of line complexes and a theorem of Gelfand-Graev	135
HELGASON (Sigurdur). — Fourier transform on symmetric spaces ..	151
KAC (V.G.) and PETERSON (D.H.). — Defining relations of certain infinite dimensional groups	165
KOSTANT (Bertram). — The McKay-correspondence, the Coxeter element and representation theory	209
KOSZUL (Jean-Louis). — Crochet de Schouten-Nijenhuis et cohomologie	257
KURANISHI (Masatake). — Cartan connections and CR structures with non-degenerate Levi-form	273
MOSTOW (G.D.). — Discrete subgroups of Lie groups	289
SCHMID (Wilfried). — Boundary value problems for group invariant differential equations	311

SINGER (I.M.). — Families of Dirac operators with applications to physics	323
SOURIAU (Jean-Marie). — Un algorithme générateur de structures quantiques	341
TRAUTMAN (Andrzej). — Optical structures in relativistic theories ..	401
WEINSTEIN (Alan). — Poisson structures and Lie algebras	421
Annexe : Résumés des autres conférences (R. BRYANT, M. DUFLO et J. TITS)	435

Le séminaire n'a été rendu possible que par le soutien de la National Science Foundation, du Centre National de la Recherche Scientifique, de l'American Mathematical Society et de la Société mathématique de France, ainsi que celui de l'Université Claude-Bernard (Lyon I), des villes de Lyon et Villeurbanne et du conseil général du Rhône. Nous espérons que ces institutions trouveront dans ce volume une preuve concrète du bien-fondé de leur effort!

1937 Elie Cartan Lectures at ENS Ulm in Julia Seminar : La Structure des groupes infinis

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Inales. - © Inales 2015 All rights reserved.

UNIVERSITÉ DE PARIS
FACULTÉ DES SCIENCES
CABINET DU DÉPARTEMENT
DES SCIENCES MATHÉMATIQUES

IV.- G.

SEMINAIRE DE MATHÉMATIQUES

Quatrième année 1936-1937

Les TRAVAUX de M. ELIE CARTAN

Le STRUCTURE des GROUPES INFINIS

Exposé fait par M. Elie CARTAN, le lundi 1er Mars 1937

Exemplaire n° 3

UNIVERSITÉ DE PARIS
FACULTÉ DES SCIENCES
CABINET DU DÉPARTEMENT
DES SCIENCES MATHÉMATIQUES

IV.- H.

SEMINAIRE DE MATHÉMATIQUES

Quatrième année 1936-1937

Les TRAVAUX de M. ELIE CARTAN

Le STRUCTURE des GROUPES INFINIS
(suite)

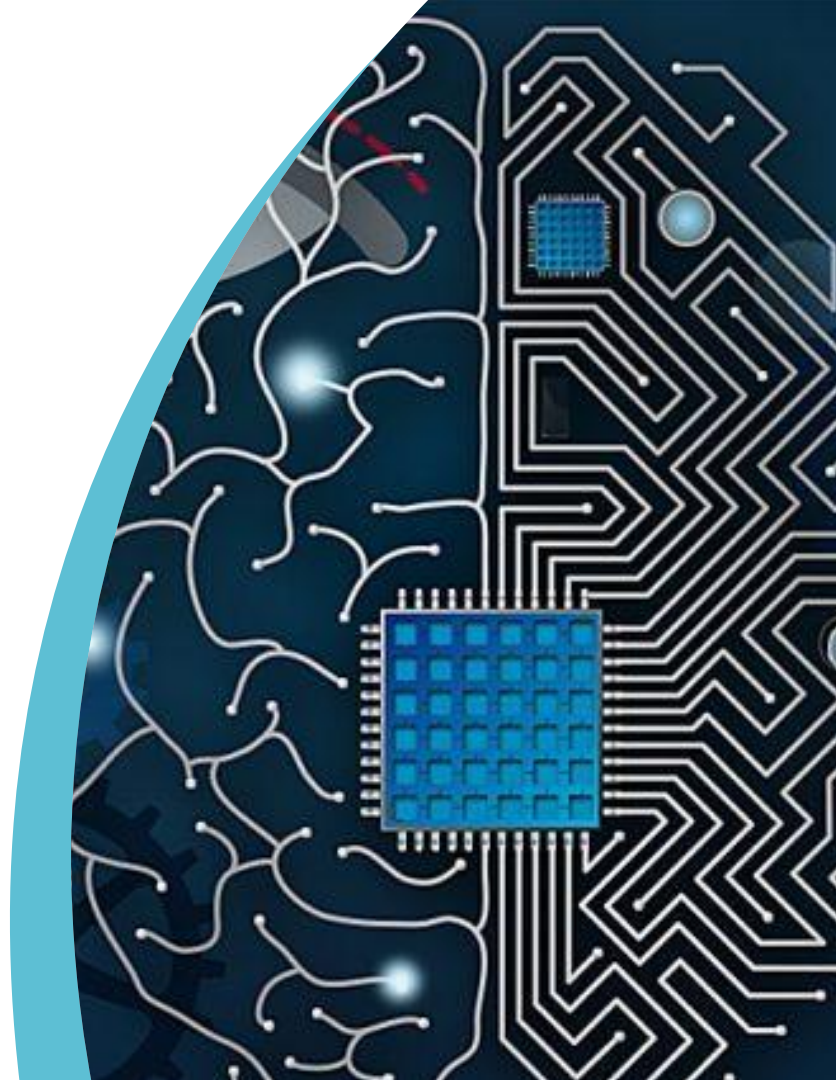
Exposé fait par M. Elie CARTAN, le lundi 15 Mars 1937

Exemplaire n° 3

THALES



Preamble



Elementary Structures of Information Geometry

This document may not be reproduced, modified, distributed, published, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - © Thales 2015 All rights reserved.

$$\omega = ig_{\alpha\beta} dz^\alpha \wedge dz^\beta$$

$$d\omega = 0$$

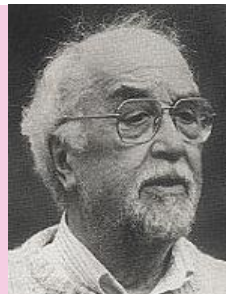
$$g_{\alpha\beta} = \partial\bar{\partial}^* U$$



Seminal work of Elie Cartan

Geometry of Jean-Marie Souriau

Study of homogeneous symplectic manifolds geometry with the action of dynamical groups. Introduction of the Lagrange-Souriau 2-form and Lie Groups Thermodynamics.



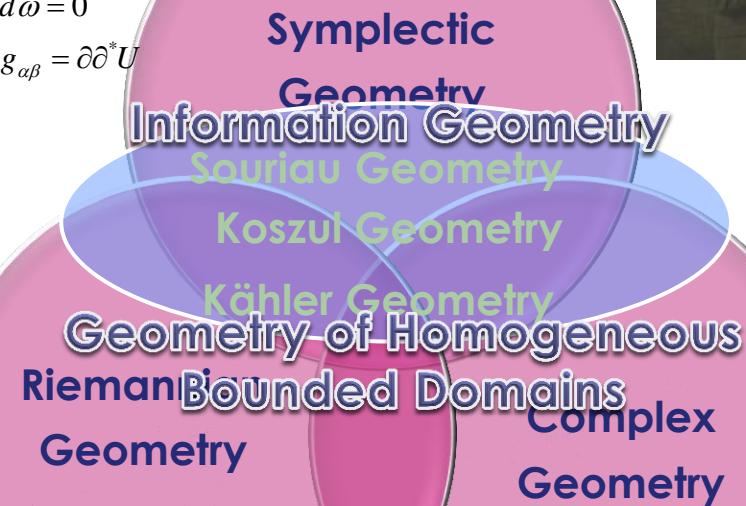
Geometry of Jean-Louis Koszul

Study of homogeneous bounded domains geometry, symmetric homogeneous spaces and sharp convex cones. Introduction of an invariant 2-form.



Geometry of Erich Kähler

Study of differential manifolds geometry equipped with a unitary structure satisfying a condition of integrability. The homogeneous Kähler case studied by André Lichnerowicz.



Information Geometry and Interdisciplinarity of Science

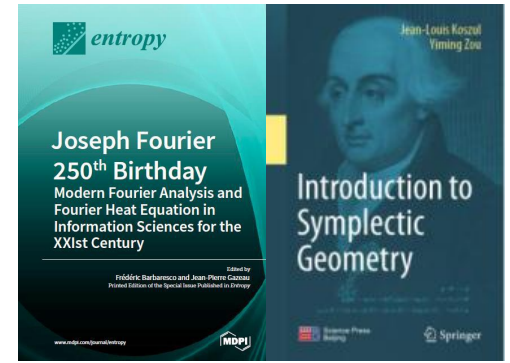
This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - © Thales 2015 All rights reserved.



J.M. Souriau
Souriau-Fisher Metric & Lie Groups Thermodynamics



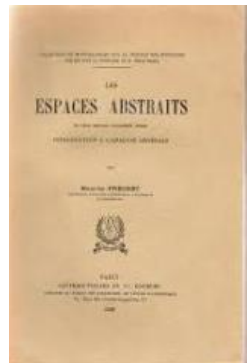
Roger Balian
Quantum Balian-Fisher Metric (qbit)



THALES



M. Fréchet
Statistics in Fisher Metric spaces



Souriau-Koszul-Fisher Metric & Electro-Magnetic Wave Statistics

THALES

Esprit de finesse et esprit de géométrie

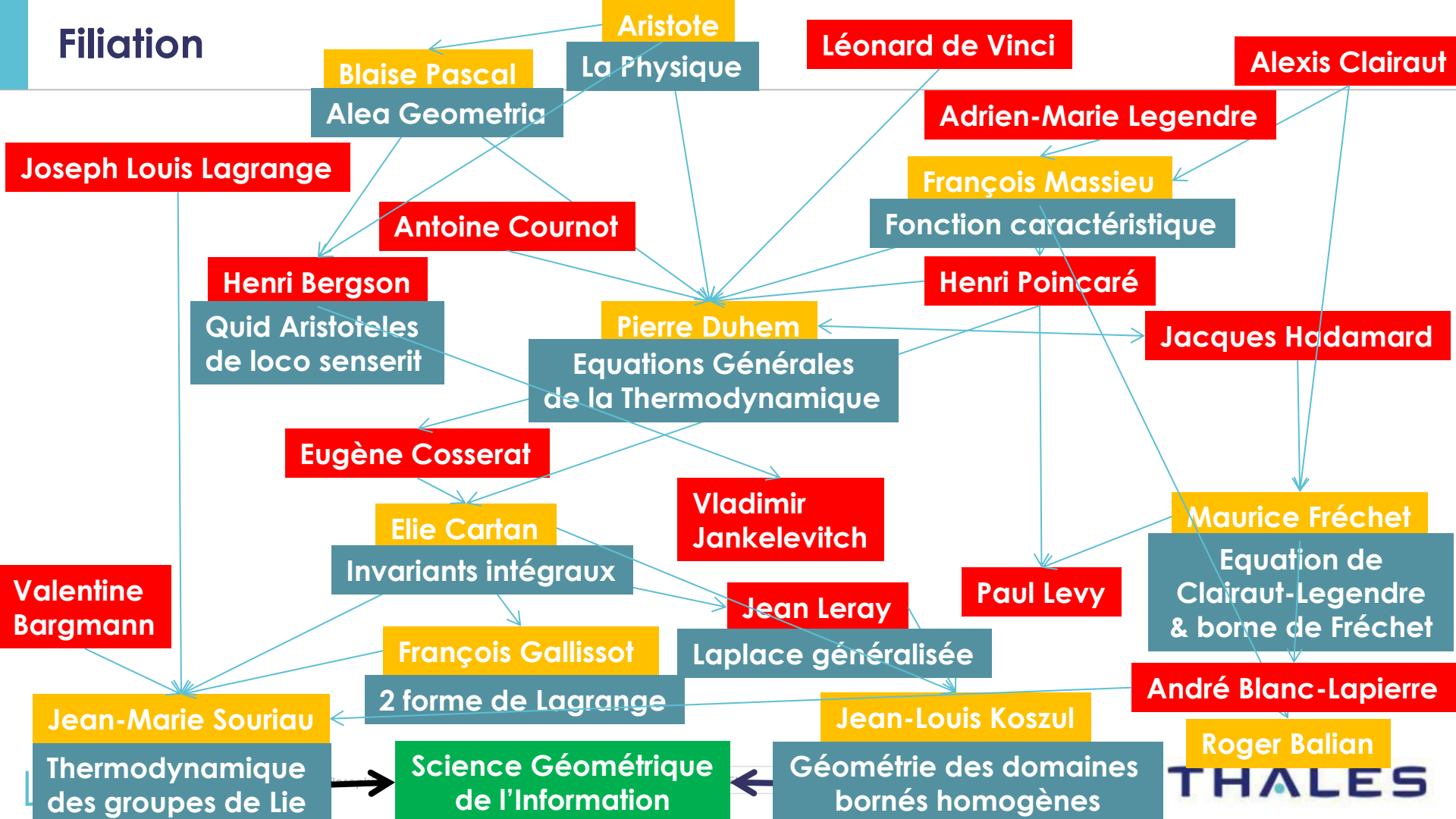


Pour la théorie de la connaissance mais aussi pour les sciences est fondamentale la notion de perspective.

Or, les expériences faites dans la géométrie algébriques, dans la théorie des nombres, et dans l'algèbre abstraite m'induisent à tenter une formulation mathématique de cette notion **pour surmonter ainsi au moyen de raisonnements d'origine géométrique la géométrie**. Il me semble en effet, que la tendance vers l'abstraction observée dans les mathématiques d'aujourd'hui, loin d'être l'ennemi de l'intuition ait le sens profond de quitter l'intuition pour la faire renaître dans une alliance entre **« esprit de géométrie » et « esprit de finesse »**, alliance rendue possible par les réserves énormes des mathématiques pures dont Pascal et Goethe ne pouvaient pas encore se douter.

Erich Kähler – Sur la théorie des corps purement algébriques, 1952

Filiation



From Souriau Lie Group Thermodynamics to Souriau-Fisher Metric



ENS Ulm 1942

Souriau Jean Marie

Introduction (1/2)

- L'étude des densités exponentielles invariantes par un groupe remonte aux travaux de Muriel Casalis dans sa thèse de 1990 [1-2].
- Ce problème avait été étudié précédemment dans un cadre géométrique par Jean-Louis Koszul dans les années 60 [3-7], en parallèle des travaux d'Ernest Vinberg [8-10] en Russie, pour définir des densités, sur des cônes convexes saillants, invariantes par les automorphismes de ces cônes.
- Le problème général a été résolu pour les groupes de Lie par Jean-Marie Souriau en Mécanique Géométrique en 1969 [11-12], en définissant une «Thermodynamique des groupes de Lie » en Mécanique Statistique.
- Ce modèle de Souriau considère la statistique des systèmes dynamiques dans leur « espace d'évolution » associée à une variété symplectique, et définit grâce à des outils cohomologiques (non équivariance de l'opérateur coadjoint pour l'application moment avec apparition d'un cocycle) une densité (de Gibbs) qui est covariante sous l'action des groupes dynamiques de la physique (par exemple, le groupe de Galilée en physique classique).

Introduction (2/2)

- Le cas des familles exponentielles invariantes par un groupe est un cas particulier associé au groupe affine.
- L'approche de Koszul et Souriau utilise la représentation affine des groupes de Lie et algèbre de Lie.
- Jean-Louis Koszul revient sur ce modèle de Jean-Marie Souriau en 1987 (ouvrage qui vient d'être traduit en anglais [13]).
- Ces structures géométriques associées aux familles exponentielles permettent de définir des généralisations de la métrique de Fisher [14-16] en géométrie de l'Information (métrique de Fisher-Koszul liée à la 2-forme de Koszul sur les cônes convexes saillants et métrique de Fisher-Souriau liée aux orbites coadjointes et l'application moment).
- L'équation de Clairaut-Legendre à la base de ces structures géométriques avait été découverte dès 1943 par Maurice Fréchet [17-18] de façon concomitante à sa découverte de la borne de Fréchet-Darmonis.

Références bibliographiques (1/2)

- [1] Casalis, M.: Familles exponentielles naturelles invariantes par un groupe. Thèse de l'Université Paul Sabatier, Toulouse, France (1990)
- [2] Casalis, M.: Familles exponentielles naturelles sur \mathbb{R}^d invariantes par un groupe. Int. Stat. Rev., 59(2):241–262, 1991.
- [3] Koszul J.L., Domaines bornés homogènes et orbites de groupes de transformations affines, Bull. Soc. Math. France 89, pp. 515-533., 1961
- [4] Koszul J.L. , Ouverts convexes homogènes des espaces affines. Math. Z., 79, 254–259, 1962
- [5] Koszul J.L. , Variétés localement plates et convexité. Osaka. J. Math., 2, 285–290, 1965
- [6] Koszul J.L., Déformations des variétés localement plates, .Ann Inst Fourier, 18 , 103-114, 1968
- [7] Koszul J.L. , Trajectoires Convexes de Groupes Affines Unimodulaires. In Essays on Topology and Related Topics ; Springer : Berlin, Germany, pp. 105–110, 1970
- [8] Alekseevsky D., Vinberg's theory of homogeneous convex cones : developments and applications, Transformation groups 2017. Conference dedicated to Prof. Ernest B. Vinberg on the occasion of his 80th birthday, Moscou, December, 2017, <https://www.mccme.ru/tg2017/slides/alexeevsky.pdf>
- [9] Vinberg E.B., Homogeneous cones, Dokl. Akad. Nauk SSSR., 133, pp. 9–12, 1960 ; Soviet Math. Dokl., 1, pp. 787–790, 1961
- [10] Vinberg E.B., The structure of the group of automorphisms of a convex cone, Trudy Moscov. Mat. Obshch., 13, pp.56–83, 1964 ; Trans. Moscow Math. Soc., 13, 1964
- [11] Souriau, J.M. Structure des systèmes dynamiques; Editions Jacques Gabay: Paris, France, 1970.

Références bibliographiques (1/2)

- [12] Souriau, J.-M., Mécanique statistique, groupes de Lie et cosmologie, Colloques internationaux du CNRS numéro 237, Géométrie symplectique et physique mathématique, 1974, pp. 59–113.
- [13] Koszul, J.L. Introduction to Symplectic Geometry; Science Press: Beijing, China, 1986. (en chinois), traduit en anglais par Springer, avec préface de M. Boyom, F. Barbaresco & C.M. Marle, 2019, <https://www.springer.com/la/book/9789811339868>
- [14] Barbaresco, F. Jean-Louis Koszul and the elementary structures of Information Geometry. In Geometric Structures of Information; Nielsen, F.; Ed.; Springer: Berlin, Germany, 2018; https://link.springer.com/chapter/10.1007/978-3-030-02520-5_12
- [15] Barbaresco, F. Koszul Contemporaneous Lectures: Elementary Structures of Information Geometry and Geometric Heat Theory. In Introduction to Symplectic Geometry; Koszul, J.L., Ed.; Springer: Berlin, Germany, 2018.; <https://www.springer.com/la/book/9789811339868>
- [16] Barbaresco, F. Jean-Louis Koszul et les Structures Élémentaires de la Géométrie de l'Information; Revue SMAI Matapli; SMAI Editor; Volume 116, pp.71-84, Novembre 2018
- [17] Fréchet, M.: Sur l'extension de certaines évaluations statistiques au cas de petits échantillons. Rev. l'Institut Int. Stat. 11(3/4), 182–205, 1943
- [18] Barbaresco, F. Les densités de probabilité « distinguées » et l'équation d'Alexis Clairaut: regards croisés de Maurice Fréchet et de Jean-Louis Koszul, Conférence Histoire de la discipline, GRETSI'17 , Juan-Les-Pins, Septembre 2017 ; <http://gretsi.fr/colloque2017/myGretsi/programme.php>

Gromov question: Are there « entropies » associated to moment maps

Bernoulli Lecture - What is Probability?

- 27 March 2018 - CIB - EPFL - Switzerland
- Lecturer: Mikhail Gromov
- [https://bernoulli.epfl.ch/images/website/What_is_Probability_v2\(2\).mp4](https://bernoulli.epfl.ch/images/website/What_is_Probability_v2(2).mp4)
- <http://forum.cs-dc.org/uploads/files/1525172771489-alternative-probabilities-2018.pdf>

Fisher Metric. Recall (Archimedes, 287-212 BCE) the *real moment map* from the unit sphere $S^n \subset \mathbb{R}^{n+1}$ to the probability simplex $\Delta^n \subset \mathbb{R}^{n+1}$ for

$$(x_0, \dots, x_n) \mapsto (p_0 = x_0^2, \dots, p_n = x_n^2)$$

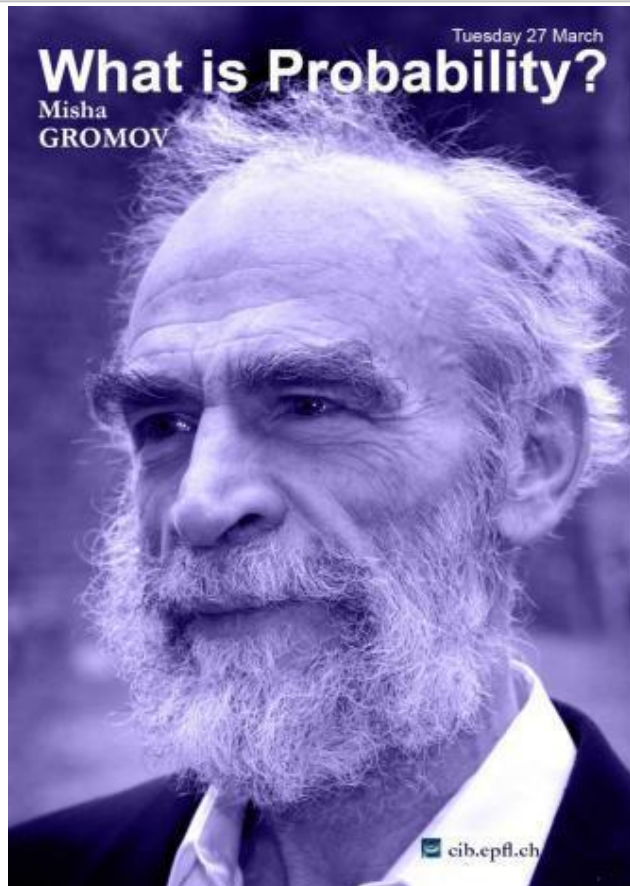
and observe following R. Fisher that the spherical metric (with constant curvature +1) thus transported to Δ^n , call it ds^2 on Δ^n , is equal, up to a scalar multiple, to the *Hessian of the entropy*

$$\text{ent}\{p_0, \dots, p_n\} = -\sum_i p_i \log p_i.$$

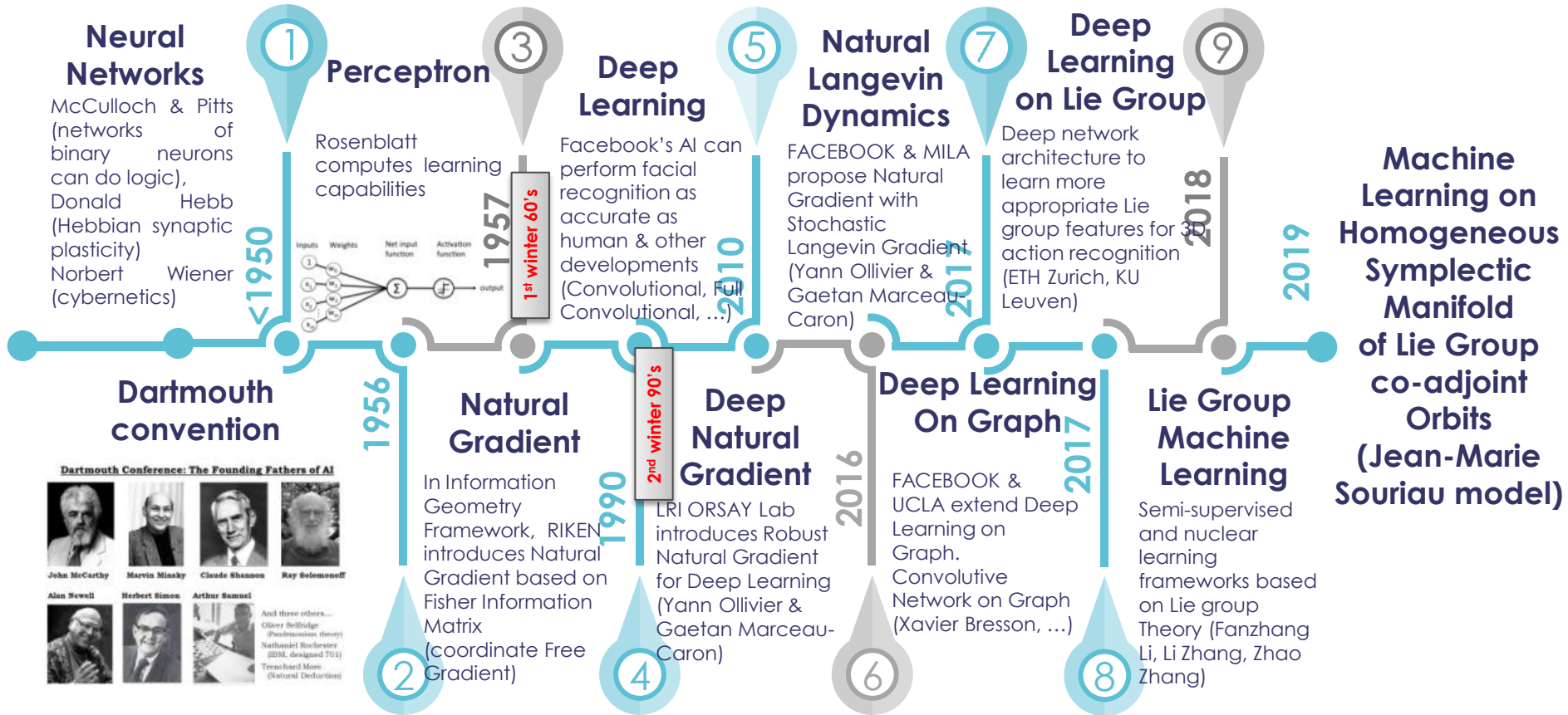
$$ds^2 = \text{const} \frac{\partial^2 \text{ent}(p_i)}{dp_i dp_j}.$$

If, accordingly, we take the "inverse Hessian" – a kind of double integral " $\int \int ds^2$ " for the *definition* of entropy – we arrive at

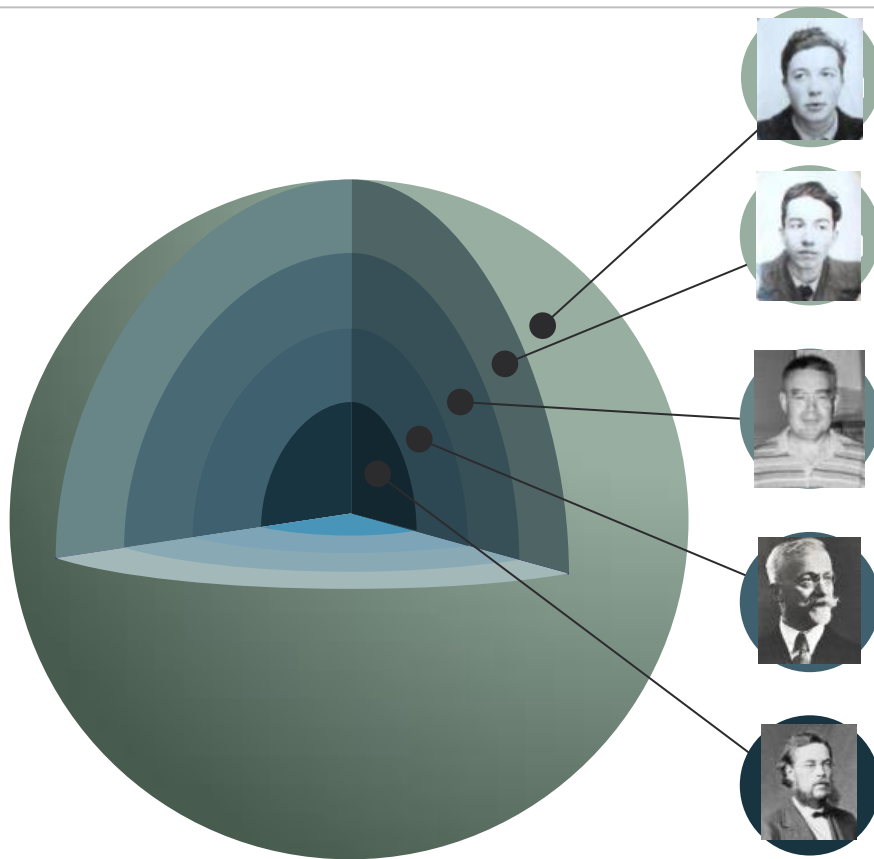
Question 2. Are there *interesting* "entropies" associated to (real and complex) moment maps of general toric varieties? Is there a *meaningful* concept of "generalised probability" grounded in positivity encountered in algebraic geometry?



Evolution/Main Trends of Machine Learning



Lie Groups Tools Development: From Group to Co-adjoint Orbits



Lie Group & Statistical Physics

Jean-Michel Bismut – Random Mechanics

Jean-Marie Souriau – Lie Group Thermodynamics, Souriau Metric

Jean-Louis Koszul – Affine Lie Group & Algebra representation

Harmonic Analysis on Lie Group & Orbits Method

Pierre Torasso & Michèle Vergne – Poisson-Plancherel Formula

Michel Duflo – Extension of Orbits Method, Plancherel & Character

Alexandre Kirillov – Coadjoint Orbits, Kirillov Character

Jacques Dixmier – Unitary representation of nilpotent Group

Lie Group Representation

Bertram Kostant – KKS 2-form, Geometric Quantization

Alexandre Kirillov – Representation Theory, KKS 2-form

Jean-Marie Souriau – Moment Map, KKS 2-form, Souriau Cocycle

Valentine Bargmann – Unitary representation, Central extension

Lie Group Classification

Carl-Ludwig Siegel – Symplectic Group

Hermann Weyl – Conformal Geometry, Symplectic Group

Elie Cartan – Lie algebra classification, Symmetric Spaces

Willem Killing – Cartan-Killing form, Killing Vectors

Group/Lie Group Foundation

Henri Poincaré – Fuchsian Groups

Felix Klein – Erlangen Program (Homogeneous Manifold)

Sophus Lie – Lie Group

Evariste Galois/Louis Joseph Lagange – Substitution Group

OPEN

Jean-Marie Souriau @ ENS Ulm 1942

Jean-Marie Souriau



Souriau Jean Marie

- Structure des systèmes dynamiques

René Deheuvels



Deheuvels René

- Formes quadratiques et groupes classiques
- Tenseurs et spineurs

Jacques Dixmier



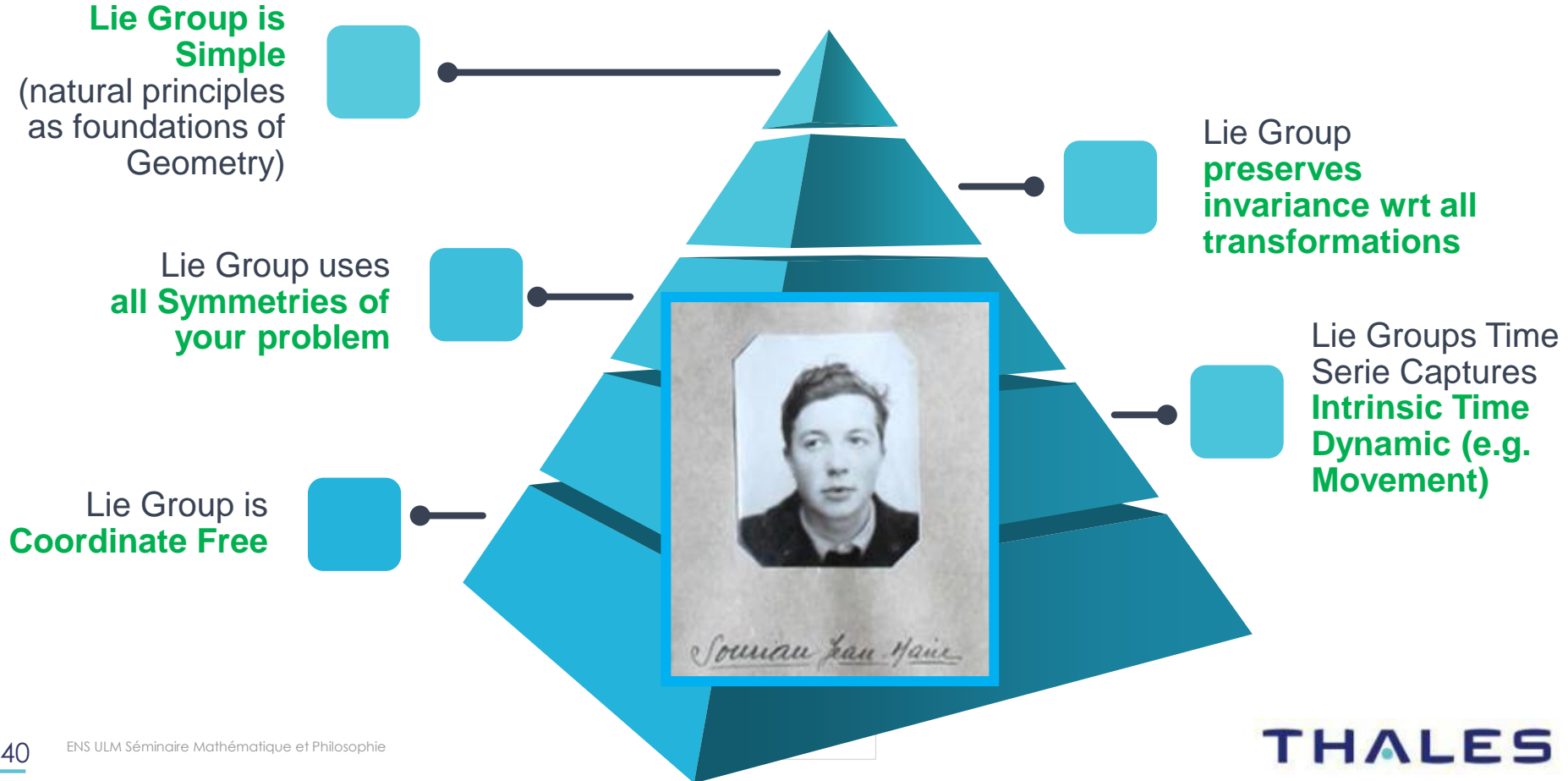
Dixmier Jacques

- Algèbres enveloppantes

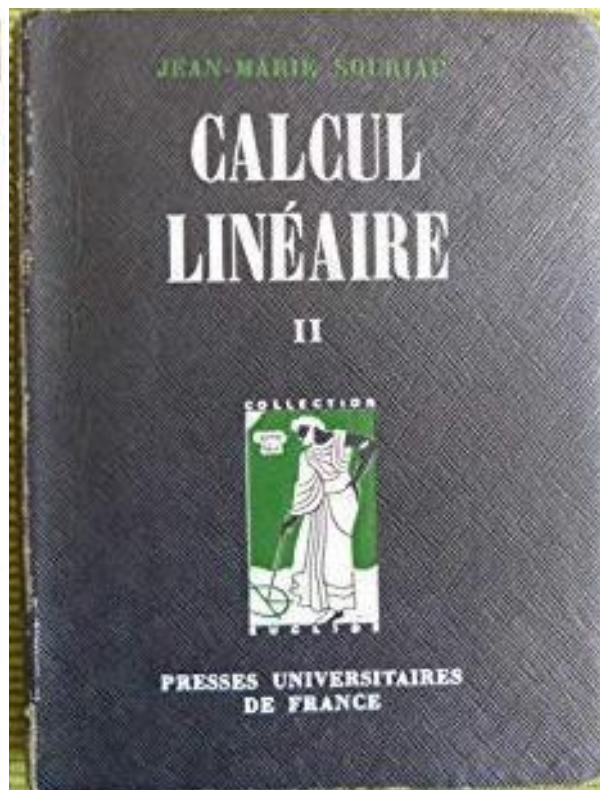
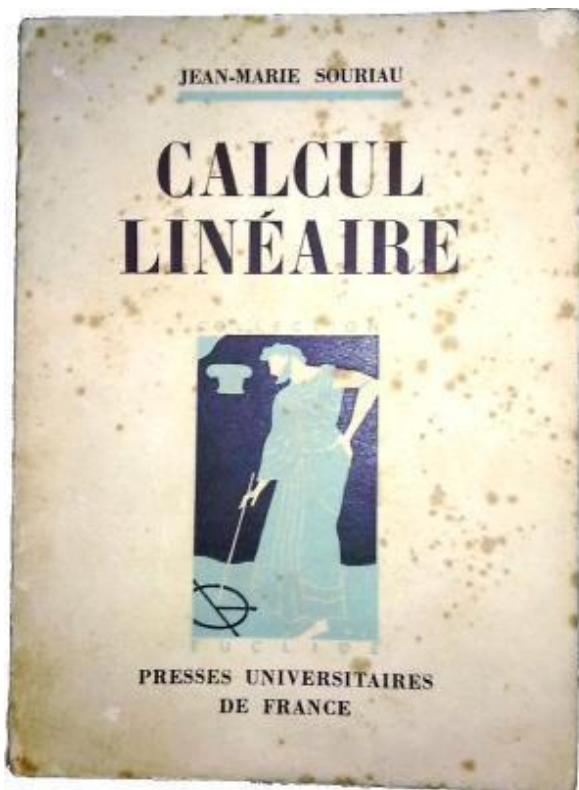
- SCIENCES - 1942



Rational to Use Lie Groups for THALES Machine Learning Applications



Souriau Book on « Calcul Linéaire » & Leverrier-Souriau Algorithm



$$P(\lambda) = \det(\lambda I - A) = \sum_{i=0}^n k_i \lambda^{n-i}$$

$$k_0 = 1 \text{ et } B_0 = I$$

$$\begin{cases} A_i = B_{i-1}A & , \quad k_i = -\frac{1}{i} \operatorname{tr}(A_i), \quad i = 1, \dots, n-1 \\ B_i = A_i + k_i I & \text{ou } B_i = B_{i-1}A - \frac{1}{i} \operatorname{tr}(B_{i-1}A) I \end{cases}$$

$$A_n = B_{n-1}A \text{ et } k_n = -\frac{1}{n} \operatorname{tr}(A_n)$$

Souriau, J.-M.: Une méthode pour la décomposition spectrale et l'inversion des matrices. Comptes-Rendus hebdomadaires des séances de l'Académie des Sciences 227 (2), 1010-1011, Gauthier-Villars, Paris (1948).

Souriau Algorithm for Characteristic Polynomial Computation

Souriau Algorithm (1948)

$$P(\lambda) = \det(\lambda I - A) = \sum_{i=0}^n k_i \lambda^{n-i} \quad Q(\lambda) = \text{adj}(\lambda I - A) = \sum_{i=0}^{n-1} \lambda^{n-i-1} B_i$$

$$k_0 = 1 \quad \text{and} \quad B_0 = I$$

$$\left\{ \begin{array}{l} A_i = B_{i-1}A \quad , \quad k_i = -\frac{1}{i} \text{tr}(A_i), \quad i = 1, \dots, n-1 \end{array} \right.$$

$$\left\{ \begin{array}{l} B_i = A_i + k_i I \quad \text{or} \quad B_i = B_{i-1}A - \frac{1}{i} \text{tr}(B_{i-1}A) I \end{array} \right.$$

$$A_n = B_{n-1}A \quad \text{and} \quad k_n = -\frac{1}{n} \text{tr}(A_n)$$

Souriau Algorithm for Exponential Map Computation

Souriau Extension Algorithm for Exponential Map

$$[\lambda I - A]^{-1} = \frac{Q(\lambda)}{P(\lambda)} \Leftrightarrow [\lambda I - A]Q(\lambda) = P(\lambda)I$$

$$\left[I \frac{d}{dt} - A \right] Q\left(\frac{d}{dt}\right) = P\left(\frac{d}{dt}\right)I \quad P\left(\frac{d}{dt}\right)\gamma = 0$$

$$1) \begin{cases} B_0 = I \text{ and } B_i = B_{i-1}A - \frac{\text{tr}(B_{i-1}A)}{i}I \\ k_0 = 1, \quad k_i = -\frac{\text{tr}(B_{i-1}A)}{i} \quad i = 1, \dots, n \end{cases}$$

$$2) \begin{cases} \gamma \text{ integrated on } [0, h] \text{ such that} \\ k_0\gamma^{(n)} + k_1\gamma^{(n-1)} + \dots + k_{n-1}\gamma^{(1)} + k_n\gamma = 0 \\ \text{with } \gamma(0) = \dots = \gamma^{(n-2)}(0) = 0 \text{ and } \gamma^{(n-1)}(0) = 1 \end{cases}$$

$$3) \text{ Computation of } \Phi(t) = e^{tA} = \sum_{i=0}^{n-1} \gamma^{(n-i-1)}(t)B_i \text{ on } [0, h]$$

$$4) \text{ Extension of Computation on } [0, ph] \text{ by } \Phi(pt) = (\Phi(t))^p$$

$$5) X(t) = \Phi(t)X_0 \text{ with } X_0 = X(0)$$

$$\Phi = Q\left(\frac{d}{dt}\right)\gamma \quad \frac{d\Phi}{dt} = A\Phi(t)$$

Good Candidate due to its high parallelization capability for Exponential Map Computation for « Lie Group Machine Learning »

Souriau algorithm to recover Lie Group Rodrigue's formula

$$SO(3) = \{R / R^{-1} = R^T\} \quad \omega_x = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} = \omega_1 L_1 + \omega_2 L_2 + \omega_3 L_3 \in \mathfrak{so}(3), \omega = (\omega_1, \omega_2, \omega_3) \in \mathbb{R}^3$$

$$e^{\omega_x} = \gamma^{(2)} B_0 + \gamma^{(1)} B_1 + \gamma B_2$$

$$B_0 = I \text{ and } k_0 = 1$$

$$B_1 = I \cdot \omega_x - \frac{\text{Tr}(I \cdot \omega_x)}{1} I = \omega_x \text{ and } k_1 = -\frac{\text{Tr}(I \omega_x)}{1} = 0$$



$$e^{\omega_x} = \gamma^{(2)} I + \gamma^{(1)} \omega_x + \gamma \omega \otimes \omega^T$$

$$B_2 = B_1 \cdot \omega_x - \frac{\text{Tr}(\omega_x \cdot \omega_x)}{2} I = \omega_x \cdot \omega_x + \|\omega\|^2 I \text{ and } k_2 = -\frac{\text{Tr}(\omega_x \cdot \omega_x)}{2} = \|\omega\|^2$$

$$B_2 = \omega_x \cdot \omega_x + \|\omega\|^2 I = \omega \otimes \omega^T \quad k_3 = 0$$

$$\gamma^{(3)}(t) + \|\omega\|^2 \gamma^{(1)}(t) = 0 \text{ with } \gamma^{(2)}(0) = 1, \gamma^{(1)}(0) = 0, \gamma(0) = 0$$

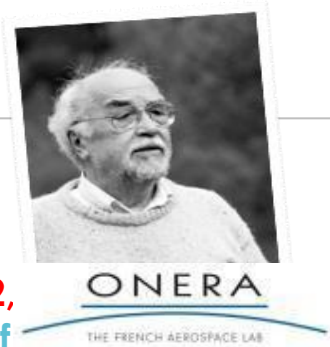


$$\gamma^{(1)}(t) = \frac{1}{\|\omega\|} \sin(\|\omega\|t) \text{ and } \gamma(t) = \frac{1}{\|\omega\|^2} (1 - \cos(\|\omega\|t))$$

$$\omega \otimes \omega^T = \omega_x \cdot \omega_x + \|\omega\|^2 I$$

$$e^{t \cdot \omega_x} = I + \frac{1}{\|\omega\|} \sin(\|\omega\|t) \omega_x + \frac{1 - \cos(\|\omega\|t)}{\|\omega\|^2} \omega_x^2$$

Jean-Marie Souriau



Graduated from ENS ULM (Ecole Normale Supérieure Paris), with Elie Cartan teacher in 1945

Souriau PhD at ONERA: **J.M. Souriau, "Sur la Stabilité des Avions" ONERA Publ., 62, vi+94, 1953** (proof that you can stabilize one aircraft with respect to all positions of engine: Caravelle), supervised by André Lichnerowicz (Collège de France) & Joseph Pérès



Algèbre Multi-Linéaire: **J.M. Souriau, Calcul linéaire, P.U.F., Paris, 1964;**
Le Verrier-Souriau Algorithm (équation des paramètres du polynôme caractéristique)

$$P(\lambda) = \det(\lambda I - A) = k_0 \lambda^n + k_1 \lambda^{n-1} + \dots + k_{n-1} \lambda + k_n$$

$$Q(\lambda) = \text{Adj}(\lambda I - A) = \lambda^{n-1} B_0 + \lambda^{n-2} B_1 + \dots + \lambda B_{n-2} + B_{n-1}$$

$$k_0 = 1 \quad \text{et} \quad B_0 = I$$

$$A_i = B_{i-1} A, \quad k_i = -\frac{1}{i} \text{tr}(A_i), \quad B_i = A_i + k_i I$$

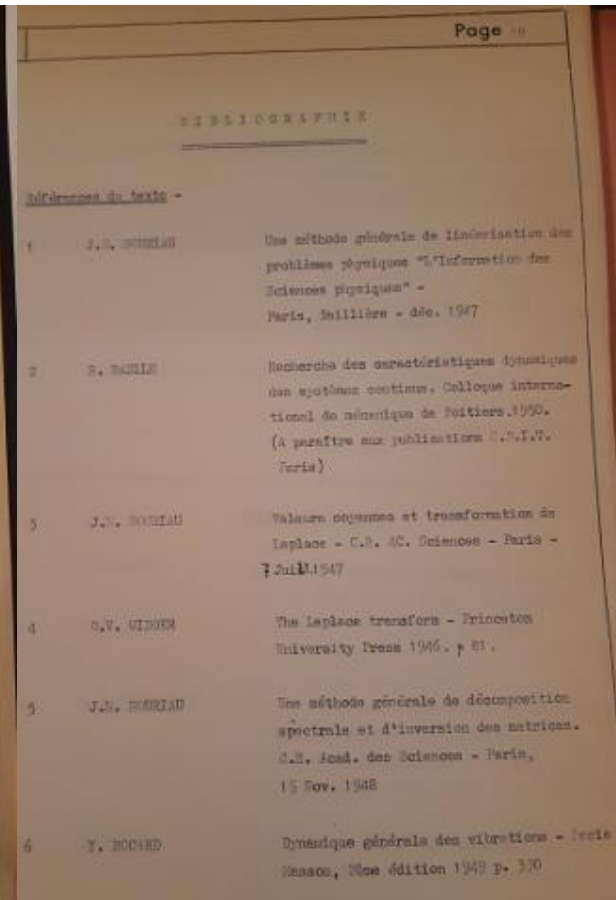
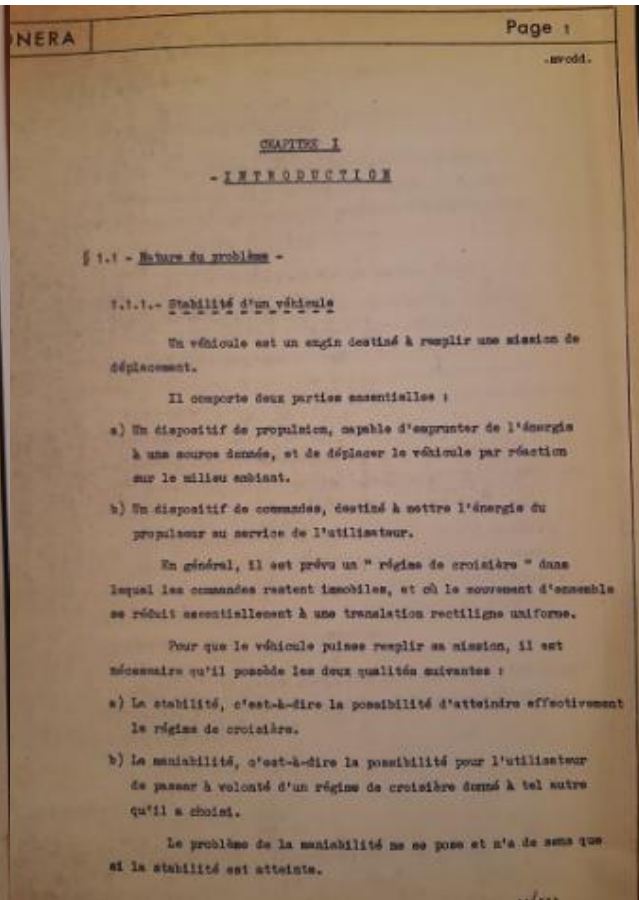
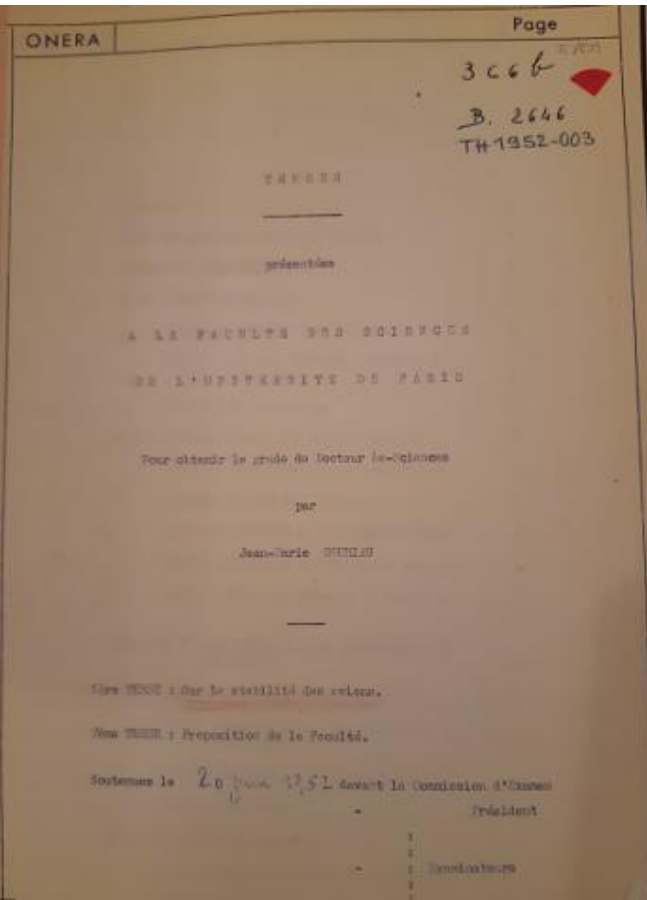
$$A_n = B_{n-1} A \quad \text{et} \quad k_n = -\frac{1}{n} \text{tr}(A_n)$$

Introduction of Symplectic Geometry in Mechanics (seminal Lagrange ideas):
J.M. Souriau, Structure des systèmes dynamiques, Dunod, Paris, 1970

« Ce que Lagrange a vu, que n'a pas vu Laplace, c'était la structure symplectique »

Jean-Marie Souriau PhD at ONERA defended June 20th, 1952: « Sur la stabilité des avions »

Jean-Marie Souriau proved that you can stabilize an airplane whatever the positions of the engines



Jean-Marie Souriau PhD at ONERA defended June 20th, 1952: « Sur la stabilité des avions »

Itinéraire d'un mathématicien Un entretien avec Jean-Marie Souriau propos recueillis par Patrick Iglesias

J.-M. Souriau Ma thèse portait sur la stabilité des avions.

J.-M. Souriau On couple les propriétés élastiques des ailes d'un avion avec la dynamique de l'atmosphère décrite par des équations aux dérivées partielles et une nappe de discontinuités tourbillonnaires. Avec tout ça, on calcule un déterminant complexe et on compte combien il fait de tours autour de l'origine quand varie une pulsation ω . S'il fait le bon nombre de tours, l'avion est stable ; sinon il se mettra à vibrer et il explosera. Et ça marche ! Ça a été utilisé pour des avions comme le Concorde. Il en résultait qu'on pouvait mettre les réacteurs n'importe où, que ça ne changeait rien à la stabilité. A la suite de quoi, on a commencé à mettre les réacteurs sur l'empennage arrière et pendant 25 ans, tous les avions qui avaient des réacteurs à l'arrière ont payé des royalties à la France, mais pas à moi.

Voilà ma vie de scientifique à mes débuts. J'appliquais les mathématiques. J'analysais une situation, j'en donnais un modèle mathématique et, de façon annexe, j'essayais d'en trouver une conséquence pratique. Les problèmes posés dans ma thèse conduisaient à des problèmes de calcul numérique. Nous avions à notre disposition un centre de calcul où les calculatrices fonctionnaient à la manivelle, puis des machines mécanographiques à cartes perforées. Nous étions en pointe à l'ONERA, parce qu'on y était obligés.

C'est comme ça que j'ai fait la première démonstration de calcul scientifique chez IBM. J'avais fait un programme qui, pendant que les invités prenaient l'apéritif, résolvait une équation du troisième degré ; à la fin de l'apéritif, on avait une racine de l'équation. Ça faisait beaucoup de bruit et ça consommait beaucoup de cartes. Peu après je faisais, dans les mêmes conditions, la première démonstration de calcul scientifique chez Bull qui ne voulait pas être en reste.

A ce moment-là, écrire un programme, c'était se mettre devant un tableau et connecter des fils. Après, j'ai vécu tous les stades de l'informatique, j'ai été témoin de l'histoire de l'informatique et des choix stupides qui se sont succédés en France pendant des dizaines d'années : tout ce qu'on a fait dans les écoles, les subventions déguisées à l'informatique française sans se demander si les élèves pourraient en faire quelque chose ! Là, j'étais plutôt spectateur. Non, j'ai quand même inventé un algorithme en 1948 qui a été utilisé sur les premiers ordinateurs aux États Unis pour l'analyse spectrale des matrices (matrices de Leontiev en économie mathématique).

Caravelle Airplane
First flight: 27 May 1955



McD Douglas MD-11
First flight: 30 Dec 1986



Le Souriau Ingénieur inspire le Souriau Physicien Mathématicien

Itinéraire d'un mathématicien Un entretien avec Jean-Marie Souriau propos recueillis par Patrick Iglesias

J.-M. Souriau Ma thèse portait sur la stabilité des avions.

J.-M. Souriau On couple les propriétés élastiques des ailes d'un avion avec la dynamique de l'atmosphère décrite par des équations aux dérivées partielles et une nappe de discontinuités tourbillonnaires. Avec tout ça, on calcule un déterminant complexe et on compte combien il fait de tours autour de l'origine quand varie une pulsation ω . S'il fait le bon nombre de tours, l'avion est stable; sinon il se mettra à vibrer et il explosera. Et ça marche! Ça a été utilisé pour des avions comme le Concorde. Il en résultait qu'on pouvait mettre les réacteurs n'importe où, que ça ne changeait rien à la stabilité. A la suite de quoi, on a commencé à mettre les réacteurs sur l'empennage arrière et pendant 25 ans, tous les avions qui avaient des réacteurs à l'arrière ont payé des royalties à la France, mais pas à moi.

J.-M. Souriau C'est bien plus tard que j'ai compris qu'il était implicite dans Lagrange. L'idée essentielle, c'est que les solutions des équations du mouvement d'un système dynamique constituent une variété symplectique. Et j'ai pensé que ça avait un intérêt d'étudier ce type de variété, comme ça a un intérêt d'étudier les variétés riemanniennes.

Le Journal Uniquement par curiosité?

J.-M. Souriau Non, c'était avec le souvenir de discussions avec des ingénieurs qui se posaient la question suivante : qu'est-ce qui est essentiel en mécanique. Je me rappelle très bien un ingénieur qui m'avait demandé : est-ce que la mécanique c'est simplement le principe de conservation de l'énergie? Ça va bien pour un système à un paramètre, mais dès qu'il y en a deux, ce n'est pas suffisant. J'avais appris bien sûr les équations de Lagrange et tous les principes analytiques de la mécanique, mais tout ça, c'était un livre de recettes; on n'y voyait pas de vrais principes.

Thèse Souriau à l'ONERA: J.M. Souriau, "Sur la Stabilité des Avions" ONERA Publ., 62, vi+94, 1953 (la position des réacteurs ne change rien à la stabilité)



ONERA

THE FRENCH AEROSPACE LAB



AIRBUS project



BOEING project

SOURIAU: Groupe Affine et Thermodynamique

- « Les différentes versions de la science mécanique peuvent se classer par la géométrie que chacune implique pour l'espace et le temps ; géométrie qui se détermine par le groupe de covariance de la théorie. Ainsi la mécanique newtonienne est covariante par le **groupe de Galilée**; la relativité restreinte par le **groupe de Lorentz-Poincaré** ; la relativité générale par le **groupe « lisse »** (le groupe des difféomorphismes de l'espace-temps). Il existe cependant une partie des énoncés de la mécanique dont la covariance appartient à un quatrième groupe – rarement envisagé : **le groupe affine**. Groupe qui figure dans le diagramme d'inclusion suivant :



- Comment se fait-il qu'un point de vue unitaire, (**qui serait nécessairement une véritable Thermodynamique**), ne soit pas encore venu couronner le tableau ?
Mystère... »

Citation de Souriau:

- « Il est évident que l'on ne peut définir de **valeurs moyennes que sur des objets appartenant à un espace vectoriel (ou affine)**; donc - si bourbakiste que puisse sembler cette affirmation - que l'on n'observera et ne mesurera de valeurs moyennes que sur des grandeurs appartenant à un ensemble possédant physiquement une structure affine. Il est clair que cette structure est nécessairement unique - sinon les valeurs moyennes ne seraient pas bien définies. »

Citation de Pascal:

- « C'est sortir de l'humanité que de sortir du milieu. La grandeur de l'âme humaine consiste à savoir s'y tenir. Tant s'en faut que la grandeur soit à en sortir, qu'elle est à n'en point sortir. »

Jean-Marie Souriau à Carthage de 1954 à 1958 (Germination de « structure des systèmes dynamiques »)



Institut des Hautes Etudes de Tunis, 8 rue de Rome

Video : <http://www.ina.fr/video/AFE01000164>

CARTHAGE & MASSILIA: Mediterranean Root of Souriau SSD Book (Institut des hautes études, 8 Rue de Rome, Tunis)



Héméroskopeion Battle between Carthage & Massilia, 490 BJC



En effet, son mari est nommé en 1952 à l'Institut des Hautes Études de Tunis ; leur installation en Tunisie, plus précisément à Carthage, lui apporte la vision d'un monde nouveau

J'allais donc rue de Rome, où était situé l'Institut, et fit la connaissance du secrétaire, Smerly, frère d'un grand poète tunisien. Par la suite, je rencontrai les collègues, les historiens Frezouls, ancien membre de l'École de Rome, Ganiage, historien de l'époque moderne, les juristes Percerou, De Bernis, les scientifiques Diacono, Souriau, etc.

Souriau Esthetism on « structure of motion » by 3 generation of ENS Ulm graduated students

Esthetism of motion

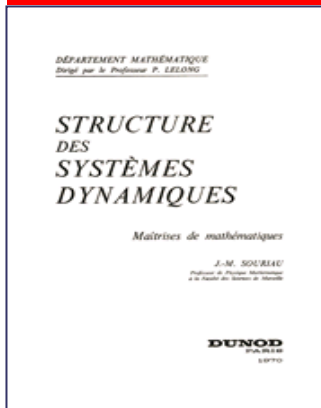


Paul Souriau
ENS Ulm 1873



Souriau Esthetism on « Structure of motion »

Structure of Motion



Jean-Marie Souriau
ENS Ulm 1942

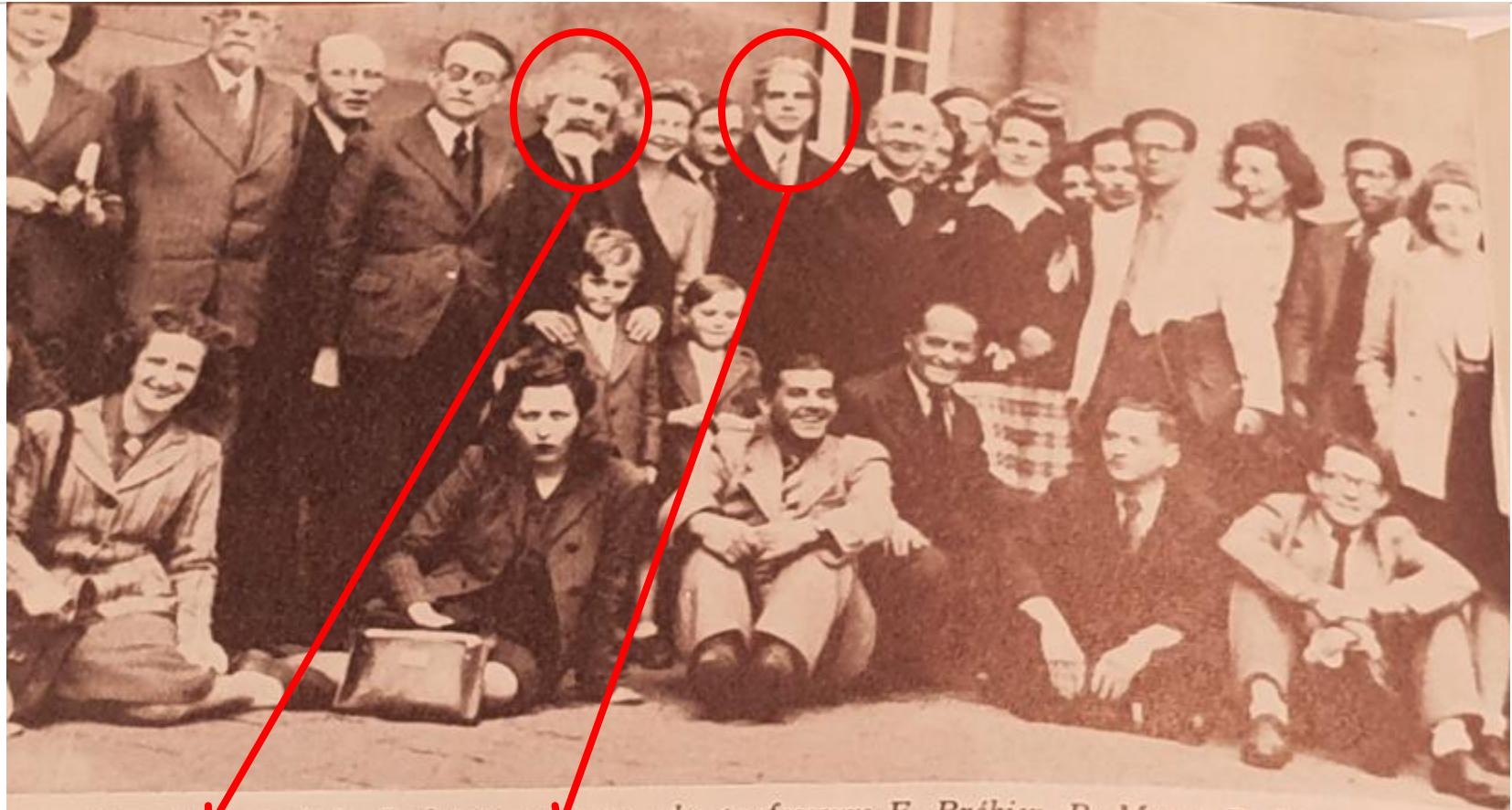
Structure of Esthetism



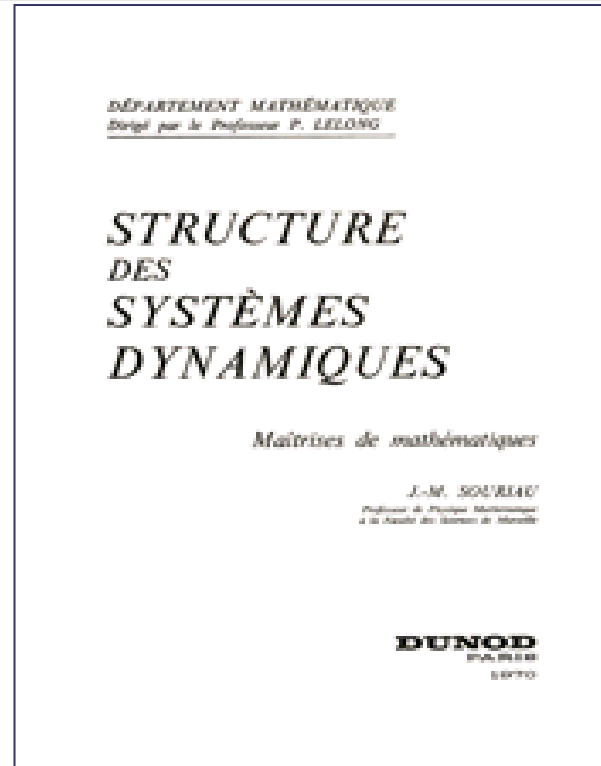
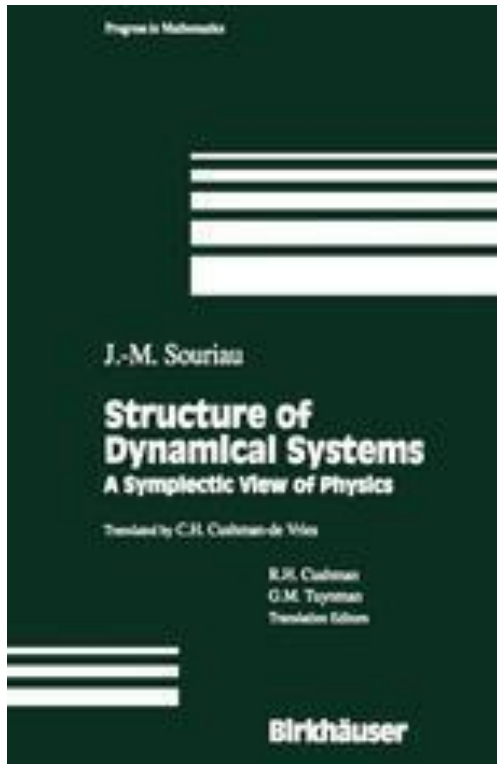
Etienne Souriau
ENS Ulm 1912



Etienne Souriau & Gaston Bachelard at Sorbonne University



Dans la cour de la Sorbonne en 1944 : les professeurs E. Bréhier, P. Mouy, R. Bayer, G. Bachelard, H. Gouhier, E. Souriau, J. Laporte et P. Romeu, bibliothécaire (de g. à d.).



- Introduction de la géométrie symplectique en Mécanique
- Invention de l'application moment
- Géométrisation du théorème de Noether
- Théorème de décomposition barycentrique
- La masse totale d'un système dynamique isolé est la classe de cohomologie du défaut d'équivariance de l'application moment
- Thermodynamique des groupes de Lie (Chapitre IV)

http://www.jmsouriau.com/structure_des_systemes_dynamiques.htm
<http://www.springer.com/us/book/9780817636951>

Souriau SSD Chapter IV: Gibbs Equilibrium is not covariant with respect to Dynamic Groups of Physics

MÉCANIQUE STATISTIQUE COVARIANTE

Le groupe des translations dans le temps (7.9) est un sous-groupe du groupe de Galilée ; mais *ce n'est pas un sous-groupe invariant*, ainsi que le

montre un calcul trivial. Si un système dynamique est *conservatif* dans un repère d'inertie, il en résulte qu'il peut *ne plus être conservatif dans un autre*. La formulation (17.24) du principe de Gibbs doit donc être élargie, pour devenir compatible avec la relativité galiléenne.

Nous proposons donc le principe suivant :

(17.77) [Si un système dynamique est invariant par un sous-groupe de Lie G' du groupe de Galilée, les équilibres naturels du système constituent l'ensemble de Gibbs du groupe dynamique G' .

Soit \mathcal{G}' l'algèbre de Lie G' ; on sait que \mathcal{G}' est une sous-algèbre de Lie de celle de G , notée \mathcal{G} ; un équilibre du système sera caractérisé par un élément Z de \mathcal{G}' , donc de \mathcal{G} ; on pourra écrire

$$(17.78) \quad Z = \begin{bmatrix} j(\omega) & \beta & \gamma \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{bmatrix}$$

en utilisant les notations (13.4) ; Z parcourt l'ensemble Ω défini en (16.219) ; à chaque valeur de Z est associé un élément M du dual \mathcal{G}'^* de \mathcal{G}' , valeur moyenne du moment μ ; on peut appliquer les formules (16.219), (16.220), qui généralisent les relations thermodynamiques (17.26), (17.27), (17.28). On voit que c'est Z (17.78) qui généralise la « température » ; le théorème d'isothermie (17.32) s'étend immédiatement : l'équilibre d'un système composé de plusieurs parties sans interactions s'obtient en attribuant à chaque composante un équilibre correspondant à la même valeur de Z ; l'entropie s , le potentiel de Planck z et le moment moyen M sont *additifs*. W

(17.79)



J.M. Souriau, Structure des systèmes dynamiques, **Chapitre IV « Mécanique Statistique »**



Trompette de Souriau

Lorsque le fait qu'on rencontre est en opposition avec une théorie régnante, il faut accepter le fait et abandonner la théorie, alors même que celle-ci, soutenue par de grands noms, est généralement adoptée

- **Claude Bernard "Introduction à l'Étude de la Médecine Expérimentale"**

Main references for Souriau « Lie Groups Thermodynamics »

1966

SUPPLEMENTO AL NUOVO CIMENTO
VOLUME IV

N. 1, 1966

Définition covariante des équilibres thermodynamiques.

J.-M. SOURIAU

Faculté des Sciences - Marseille

(ricevuto il 5 Novembre 1965)

CONTENTS. — 1. Un problème variationnel. — 2. Mécanique statistique classique. — 3. Équilibres permis par un groupe de Lie. — 4. Exemples. — 5. Localisation de la température vectorielle.

1974

Colloques Internationaux C.N.R.S.
N° 237 — Géométrie symplectique et physique mathématique

MÉCANIQUE STATISTIQUE, GROUPES DE LIE ET COSMOLOGIE

Jean-Marie SOURIAU (1)

Première partie

FORMULATION SYMPLECTIQUE DE LA MÉCANIQUE STATISTIQUE

Référence to Blanc-Lapierre Book in Souriau Book

[7] A. Blanc-Lapierre, P. Casal, and A. Tortrat, *Méthodes mathématiques de la mécanique statistique*, Masson, Paris, 1959.

FORMULATION SYMPLECTIQUE DE LA MECANIQUE STATISTIQUE

1974

RESUME

La notion classique d'ensemble canonique de Gibbs est étendue au cas d'une variété symplectique sur laquelle un groupe de Lie possède une action symplectique ("groupe dynamique").

La définition rigoureuse donnée ici permet d'étendre un certain nombre de propriétés thermodynamiques classiques (la température est ici un élément de l'algèbre de Lie du groupe, la chaleur un élément de son dual), notamment des inégalités de convexité.

Dans le cas de groupes non commutatifs, des propriétés particulières apparaissent : la symétrie est spontanément brisée, certaines relations de type cohomologique sont vérifiées dans l'algèbre de Lie du groupe.

Diverses applications sont abordées (corps tournants, mécanique statistique covariante ou relativiste).

[Ces résultats précisent et complètent une étude publiée dans un ouvrage antérieur ⁽²⁾ qui sera désigné par les initiales S.S.D.].

J.L. Lagrange Symplectic Structure (Lagrange's paper of 1810)

■ Symplectic Structure discovered by J.L. Lagrange

- The concept of a symplectic structure appeared in Mathematics much earlier than the word symplectic, in the works of Joseph Louis Lagrange (1736–1813), first in his paper about the slow changes of the orbital elements of planets in the solar system, then in a following paper a little later, as a fundamental ingredient in the mathematical formulation of any problem in Mechanics.
- J.-M. Souriau has shown that Lagrange's parentheses are the components of the canonical symplectic 2-form on the manifold of motions of the mechanical system, in the chart of that manifold. Lagrange discovered this notion of a symplectic structure more than 100 years before that notion was so named by H. Weyl.

■ See papers:

- J.-M. Souriau, La structure symplectique de la mécanique décrite par Lagrange en 1811, *Mathématiques et sciences humaines*, tome 94 (1986), p. 45–54.
- Charles-Michel Marle, The inception of Symplectic Geometry: the works of Lagrange and Poisson during the years 1808–1810, *Thirty years of bihamiltonian systems*, Be_dlewo, August 3–9, 2008

Lagrange 2-form rediscovered by Jean-Marie Souriau

- Rewriting equations of classical mechanics in phase space

$$\begin{pmatrix} r \\ v \end{pmatrix}$$

$$m \frac{d^2 r}{dt^2} = F \quad \longrightarrow \quad m \frac{dv}{dt} = F \quad \text{et} \quad v = \frac{dr}{dt}$$

- Souriau rediscovered that Lagrange had considered the evolution space: $y = \begin{pmatrix} t \\ r \\ v \end{pmatrix} \in V$

$$\begin{cases} m \delta v - F \delta t = 0 \\ \delta r - v \delta t = 0 \end{cases}$$

- A dynamic system is represented by a foliation. This foliation is determined by an antisymmetric covariant 2nd order tensor σ , called the Lagrange (-Souriau) form, a bilinear operator on the tangent vectors of V .

$$\sigma(\delta y)(\delta' y) = \langle m \delta v - F \delta t, \delta' r - v \delta' t \rangle - \langle m \delta' v - F \delta' t, \delta r - v \delta t \rangle \quad \delta y = \begin{pmatrix} \delta t \\ \delta r \\ \delta v \end{pmatrix} \quad \text{et} \quad \delta' y = \begin{pmatrix} \delta' t \\ \delta' r \\ \delta' v \end{pmatrix}$$

- In the Lagrange-Souriau model, σ is a 2-form on the evolution space V , and the differential equation of motion implies: $\delta y \in \mathcal{E}$

$$\sigma(\delta y)(\delta' y) = 0, \quad \forall \delta' y$$

$$\sigma(\delta y) = 0 \quad \text{ou} \quad \delta y \in \ker(\sigma)$$

François Gallissot Theorem & Poincaré-Cartan Integral Invariant

- **Gallissot Theorem:** There are 3 types of differential forms generating the equations of a material point motion, invariant by the action of the Galileo group

$$A: \begin{cases} s = \frac{1}{2m} \sum_{i=1}^3 (mdv_i - F_i dt)^2 \\ e = \frac{m}{2} \sum_{j=1}^3 (dx_j - v_j dt)^2 \end{cases}$$

F. GALLISSOT, Les formes extérieures en Mécanique (*Thèse*), Durand, Chartres, 1954.

$$B: f = \sum_1^3 \delta_{ij} (dx_i - v_i dt) (mdv_j - F_j dt) \text{ with } \delta_{ij} \text{ krönecker symbol}$$

$$C: \omega = \sum_1^3 \delta_{ij} (mdv_i - F_i dt) \wedge (dx_j - v_j dt)$$

- $d\omega = 0$ constrained the Pfaff form $\delta_{ij} F_i dx_j$ to be closed and to be reduced to the differential of $U: C \Rightarrow \omega = m\delta_{ij} dv_i \wedge dx_j - dH \wedge dt$ with $H = T - U$ and $T = 1/2 \sum_{i=1}^3 m(v_i)^2$

- It proves that ω has an exterior differential $d\omega$ generating **Poincaré-Cartan Integral invariant:**

$$d\omega = \sum_{i=1}^3 mv_i dx_j - H dt$$

François Gallissot Work in 1951 based on Elie and Henri Cartan works

APPLICATION DES FORMES EXTÉRIEURES DU 2^e ORDRE
A LA DYNAMIQUE NEWTONIENNE ET RELATIVISTE
par François GALLISSOT (Grenoble).

1951

Introduction

La mécanique est généralement conçue en définissant des grandeurs cinématiques vitesse, accélération, des grandeurs dynamiques forces, la différentielle du vecteur vitesse étant liée au produit vecteur force par dt par le principe de Newton. En admettant que les phénomènes mécaniques sont susceptibles d'être décrits par des équations différentielles, il est intéressant d'avoir une forme génératrice des équations différentielles complètement invariantes dans les transformations du groupe ponctuel portant sur un ensemble de $2n$ variables de position et vitesse.

A notre point de vue en mécanique Newtonienne, un point matériel de masse m est repéré par 7 variables x^i, t, v^i (i variant de 1 à 3) auquel on associe une forme extérieure construite sur les différentielles dx^i, dt, dv^i .

$$\omega = mk_{ij}dv^i \wedge dx^j - mk_{ij}v^i dv^j \wedge dt + k_{ij}X^i dx^j \wedge dt$$

(k_{ij} symbole de Kronecker, X^i composantes de la force F appliquée au point) ω est invariante dans les transformations du groupe Galiléen et son expression a même forme par rapport à tout repère Galiléen orthonormé. Les équations différentielles du mouvement sont les équations associées à ω :

$$\frac{\partial \omega}{\partial (dx^j)} = -m dv^j + X^j dt = 0 \quad \frac{\partial \omega}{\partial (dv^i)} = m(dx^i - v^i dt) = 0.$$

Il est essentiel de remarquer que ce sont les équations associées à ω qui lient les paramètres v^i aux différentielles des paramètres de position x^i et au temps.

(1) Cf. Communication au Congrès des Sociétés Savantes, Grenoble, 1952, sous presse. Comptes rendus de l'Académie des Sciences, t. 234, p. 2148-2150, 26 mai 1952.

MÉCANIQUE. — *Les formes extérieures et la Mécanique des milieux continus.*

Note de M. François GALLISSOT, présentée par M. Henri Villat.

Le mouvement d'un milieu continu avec déformations est une application φ de classe C^r ($r \geq 2$) de l'espace R^4 dans ρ^3 , en désignant par ρ^3 l'espace numérique à 3 dimensions, par R^4 le produit de l'espace numérique à 3 dimensions par la droite numérique t .

Plus généralement, nous considérons les applications φ de classe C^r de l'espace numérique R^p dans l'espace numérique ρ^n , auxquelles nous associons la variété des jets (1) du premier ordre $J^1(R^p, \rho^n)$. Soient (ξ^1, \dots, ξ^n) les coordonnées d'un point ξ de $\rho^n(x^1, \dots, x^p)$ les coordonnées d'un point x de R^p . Dans $J^1(R^p, \rho^n)$ de dimension $np + n + p$, les coordonnées canoniques d'un point sont $(\xi^1, \dots, \xi^n, a_1^1, \dots, a_1^p, x^1, \dots, x^p)$. Les éléments $(0, \dots, 0; \dots; a_i^1; \dots; 0, \dots, 0)$ constituent le noyau E_{pn} de la variété des jets, de sorte que $J^1(R^p, \rho^n) = R^p E_{pn} \rho^n$.

Les champs \vec{a}_i et les formes $\theta(\vec{a}_i) V_p$. Nous désignons par \vec{a}_i le champ de vecteurs dans J^1 de composantes $(0, \dots, 0; a_1^1, \dots, a_1^p; 0, \dots, 0)$. A la forme volume V_p définie sur R^p qui s'écrit $V_p = dx^1 \wedge \dots \wedge dx^p$ et au champ \vec{a}_i nous faisons correspondre, au moyen de l'opérateur des transformations infinitésimales

François Gallissot Work in 1952 based on Elie and Henri Cartan works

LES FORMES EXTÉRIEURES EN MÉCANIQUE

par F. GALLISSOT.

1952

INTRODUCTION

La mécanique des systèmes paramétriques développée traditionnellement d'après les idées de Lagrange s'est toujours heurtée à des difficultés notables lorsqu'elle a désiré aborder les questions de frottement entre solides (impossibilité et indétermination) ou la notion générale de liaison (asservissement de M. Béghin), d'autre part la forme lagrangienne des équations du mouvement ne nous donne aucune indication sur la nature du problème de l'intégration.

Dans ces célèbres leçons sur les invariants intégraux Elie Cartan a montré que toutes les propriétés des équations différentielles de la dynamique des systèmes holonomes résultaient de l'existence de l'invariant intégral $\int \omega$, $\omega = p_i dq^i - H dt$. Ainsi à tout système holonome dont les forces dérivent d'une fonction de forces est associé une forme ω , les équations du mouvement étant les caractéristiques de la forme extérieure $d\omega$. Au cours de ces dix dernières années, sous l'influence des topologistes s'est édifiée sur des bases qui semblent définitives la théorie des formes extérieures sur les variétés différentiables. Il est alors naturel de se demander si la mécanique classique ne peut pas bénéficier largement de ce courant d'idées, si elle ne peut pas être construite en plaçant à sa base une forme extérieure de degré deux, si grâce à la notion de variétés, la notion de liaison ne peut pas être envisagée sous un angle plus intelligible, si les indéterminations et impossibilités qui paraissent paradoxales dans le cadre lagrangien n'ont pas une explication naturelle, enfin s'il n'est pas possible de considérer sous un jour nouveau le problème de l'intégration des équations du mouvement, ces dernières étant engendrées par une forme Ω de degré deux.

Pour atteindre ces divers objectifs il m'a semblé utile de reprendre dans le chapitre 1 l'étude des bases logiques sur lesquelles est édiflée la mécanique galiléenne. Je montre ainsi dans le § 1 que lorsqu'on se propose de trouver des formes génératrices des équations du mouvement d'un point matériel invariants dans les transformations du groupe galiléen, la forme la plus intéressante est une forme extérieure de degré deux définie sur une variété $V_7 = E_3 \otimes E \otimes T$ (E_3 espace euclidien, T droite numérique temporelle)⁽¹⁾. Dans le § 11 on montre qu'à tout système paramétrique holonome à n degrés de liberté est associé une forme Ω de degré deux de rang $2n$ définie sur une variété différentiable dont les caractéristiques sont les équations du mouvement⁽²⁾. Cette forme s'exprime si l'on veut au moyen de $2n$ formes de Pfaff et de dt , la forme hamiltonienne n'étant qu'un cas particulier simple. Dans le § 3 j'indique sommairement comment on peut s'affranchir de la servitude des coordonnées dans l'étude des systèmes dynamiques et le rôle important joué par l'opérateur $\hat{i}(\)$ antidérivation de M. H. Cartan⁽³⁾, le champ caractéristique E de la forme Ω étant défini par la relation $\hat{i}(E)\Omega = 0$.

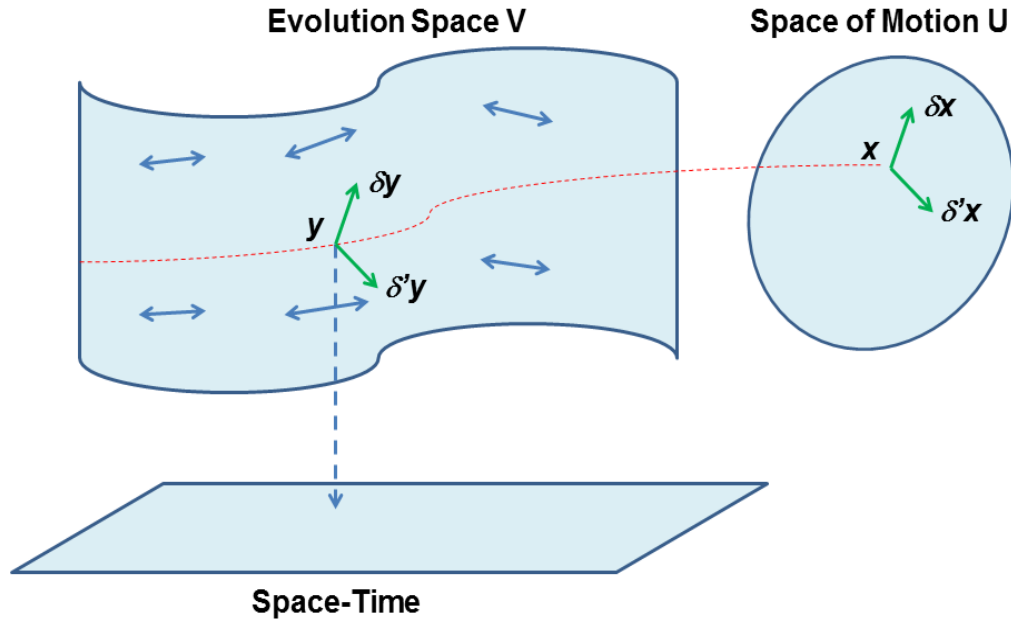
(1) M. KRAVTCHEUKO a présenté cette conception au VIII^e Congrès de Mécanique.

(2) Dès 1946 M. LICHNEROWICZ au *Bulletin des Sciences Mathématiques* tome LXX, p. 90 a déjà introduit les formes extérieures pour la formation des équations des systèmes holonomes et linéairement non holonomes.

(3) M. H. CARTAN, Colloque de Topologie, Bruxelles, 1950. Masson, Paris, 1951.

F. GALLISSOT, Les formes extérieures en Mécanique (*Thèse*), Durand, Chartres, 1954.

Evolution space of Lagrange-Souriau



$$\begin{cases} m\delta v - F\delta t = 0 \\ \delta r - v\delta t = 0 \end{cases}$$

Natural Exponential Families Invariant by a Group: Muriel Casalis & Gérard Letac



Seminal work of Muriel Casalis (Institut Mathématique de Toulouse)

Muriel Casalis PhD at Paul Sabatier Toulouse University supervised by Gérard Letac

Reference of Muriel Casalis

- Casalis, M.: Familles exponentielles naturelles invariantes par un groupe. PhD from Paul Sabatier university, Toulouse, France, 1990
- Casalis, M.: Familles exponentielles naturelles sur \mathbb{R}^d invariantes par un groupe. Int. Stat. Rev., 59(2):241–262, 1991.
- Casalis, M. : Les familles exponentielles à variance quadratique homogène sont des lois de Wishart sur un cône symétrique, C. R. Acad. Sci. Paris Sér. I Math. 312 , p. 537–540., 1991



References of Gérard LETAC

- Letac, G. : A characterization of the Wishart exponential families by an invariance property. J. Theoret. Probab., 2(1):71{86, 1989
- Letac, G.: Lectures on Natural Exponential Families and their Variance Functions, Instituto De Matematica Pura E Aplicada, 1992
- Letac, G. : Les familles exponentielles statistiques invariantes par les groupes du Cône et du paraboloïde de révolution, Journal of Applied Probability, Vol. 31, Studies in Applied Probability, pp. 71-95, 1994
- Letac, G.; Wesolowski, J. : Why Jordan Algebras are Natural in Statistics: Quadratic Regression Implies Wishart Distributions. Bull. Soc. math. France 139 (1), p. 129–144, 2011



Travaux précurseurs de Muriel Casalis

International Statistical Review (1991), 59, 2, pp. 241-262. Printed in Great Britain
© International Statistical Institute

Familles Exponentielles Naturelles sur \mathbb{R}^d Invariantes par un Groupe

Muriel Casalis

Laboratoire de Statistique, Université Paul Sabatier, 118, route de Narbonne, 31062
Toulouse Cedex, France

Résumé

La caractérisation des familles exponentielles naturelles de \mathbb{R}^d préservées par un groupe d'affinités donné est faite dans trois cas: celui d'un groupe compact, en particulier du groupe des rotations, celui de groupe hyperbolique et enfin celui d'un groupe quelconque de translations. La démarche adoptée consiste à traduire la propriété d'invariance de la famille par une propriété portant sur les mesures qui l'engendrent puis à rechercher ces dernières en conséquence.

1 Introduction

Il est courant en statistique d'envisager un modèle $(\Omega, \mathcal{A}, (P_\theta)_{\theta \in \Theta})$ tel qu'il existe un groupe G de permutations de Ω préservant globalement la famille de probabilités $F = \{P_\theta, \theta \in \Theta\}$, c'est-à-dire que pour tout (θ, g) de $\Theta \times G$, l'image $g(P_\theta)$ de P_θ par g est encore dans F (Barndorff-Nielsen parle alors de modèle de transformations). On pourra consulter Barndorff-Nielsen et al. (1982) et plus récemment le livre de Barndorff-Nielsen (1988).

Un exemple célèbre est celui des distributions de Fisher-Von-Mises pour lequel Ω est la sphère unité de l'espace euclidien E ,

$$P_\theta(dx) = L(\theta)^{-1} \exp(\theta, x) \sigma(dx),$$

σ désignant la probabilité uniforme sur Ω et $L(\theta)$ le coefficient de normalisation, et pour lequel G est le groupe des rotations $\mathcal{O}(E)$ de E .

Dans cet exemple, $(P_\theta, \theta \in \Theta)$ est une famille exponentielle naturelle au sens suivant.

Soit E un espace vectoriel de dimension finie, E^* son dual et si (θ, x) est dans $E^* \times E$, (θ, x) désigne le crochet de dualité; soit, de plus, μ une mesure de Radon positive sur E ; on note alors L_μ la transformée de Laplace de μ définie par:

$$L_\mu: E^* \rightarrow [0, \infty]: \theta \mapsto \int_E \exp(\theta, x) \mu(dx);$$

D_μ est l'ensemble $\{\theta \in E^*, L_\mu(\theta) < \infty\}$, $\Theta(\mu)$ son intérieur et k_μ la fonction définie sur $\Theta(\mu)$ par:

$$k_\mu(\theta) = \text{Log } L_\mu(\theta). \quad (1.1)$$

On désigne aussi par $\mathcal{M}(E)$ l'ensemble des mesures de Radon μ positives telles que:

- (i) μ n'est pas concentrée sur un sous-espace affine strict de E ;
- (ii) $\Theta(\mu)$ est non vide.

N° d'ordre - 679

THESE

présentée à

L'UNIVERSITE PAUL SABATIER DE TOULOUSE (SCIENCES)

pour obtenir

DOCTORAT DE L'UNIVERSITE PAUL SABATIER

Spécialité : MATHÉMATIQUES APPLIQUÉES

par

Muriel BONNEFOY - CASALIS

FAMILLES EXPONENTIELLES NATURELLES INVARIANTES PAR UN GROUPE

Soutenue le 11 Juin 1990, devant la Commission d'Examen :

MM.	H. CAUSSINUS	Professeur à l'Université Paul Sabatier
	D. BAKRY	Professeur à l'Université Paul Sabatier
	J. FARAUT	Professeur à l'Université PARIS VI
	Y. GUIVAR'CH	Professeur à l'Université PARIS VI
	G. LETAC	Professeur à l'Université Paul Sabatier

Laboratoire de Statistique et Probabilités
UNIVERSITÉ PAUL SABATIER

THALES

NEF (Natural Exponential Families): Letac & Casalis

Let E a vector space of finite size, E^* its dual. $\langle \theta, x \rangle$ Duality bracket with

$(\theta, x) \in E^* \times E$. μ Positive Radon measure on E , Laplace transform is :

$$L_\mu : E^* \rightarrow [0, \infty] \text{ with } \theta \mapsto L_\mu(\theta) = \int e^{\langle \theta, x \rangle} \mu(dx)$$

Transformation $k_\mu(\theta)$ defined on $\Theta(\mu)$ interior of $D_\mu = \{\theta \in E^*, L_\mu < \infty\}$

$$k_\mu(\theta) = \log L_\mu(\theta)$$

Natural exponential families are given by:

$$F(\mu) = \left\{ P(\theta, \mu)(dx) = e^{\langle \theta, x \rangle - k_\mu(\theta)} \mu(dx), \theta \in \Theta(\mu) \right\}$$

Injective function (domain of means): $k'_\mu(\theta) = \int x P(\theta, \mu) \mu(dx)$

And the inverse function: $\psi_\mu : M_F \rightarrow \Theta(\mu)$ with $M_F = \text{Im}(k'_\mu(\Theta(\mu)))$

Covariance operator: $V_F(m) = k''_\mu(\psi_\mu(m)) = \left(\psi'_\mu(m)\right)^{-1}, m \in M_F$

NEF (Natural Exponential Families): Letac & Casalis

■ Measure generated by a family F :

$$F(\mu) = F(\mu') \Leftrightarrow \exists(a, b) \in E^* \times R, \text{ such that } \mu'(dx) = e^{\langle a, x \rangle + b} \mu(dx)$$

■ Let F an exponential family of E generated by μ and $\varphi : x \mapsto g_\varphi x + v_\varphi$

with $g_\varphi \in GL(E)$ automorphisms of E and $v_\varphi \in E$, then the family

$$\varphi(F) = \{\varphi(P(\theta, \mu)), \theta \in \Theta(\mu)\} \text{ is an exponential family of } E$$

generated by $\varphi(\mu)$

■ Definition: An exponential family F is invariant by a group G (affine group of E), if $\forall \varphi \in G, \varphi(F) = F : \forall \mu, F(\varphi(\mu)) = F(\mu)$

(the contrary could be false)

NEF (Natural Exponential Families): Letac & Casalis

■ Theorem (Casalis): Let $F = F(\mu)$ an exponential family of E and G

affine group of E , then F is invariant by G if and only:

$\exists a : G \rightarrow E^*$, $\exists b : G \rightarrow R$, such that :

$$\forall (\varphi, \varphi') \in G^2, \begin{cases} a(\varphi\varphi') = {}^t g_\varphi^{-1} a(\varphi') + a(\varphi) \\ b(\varphi\varphi') = b(\varphi) + b(\varphi') - \langle a(\varphi'), g_\varphi^{-1} v_\varphi \rangle \end{cases}$$

$$\forall \varphi \in G, \varphi(\mu)(dx) = e^{\langle a(\varphi), x \rangle + b(\varphi)} \mu(dx)$$

■ When G is a linear subgroup, b is a character of G , a could be obtained by the help of **Cohomology of Lie groups**.

NEF (Natural Exponential Families): Letac & Casalis

■ If we define action of G on E^* by: $g.x = {}^t g^{-1}x, g \in G, x \in E^*$

we can verify that: $a(g_1 g_2) = g_1.a(g_2) + a(g_1)$

■ the action a is an inhomogeneous 1-cocycle: $\forall n > 0$, let the set of all functions from G^n to E^* , $\mathfrak{Z}(G^n, E^*)$ called inhomogeneous n -cochains,

then we can define the operators: $d^n : \mathfrak{Z}(G^n, E^*) \rightarrow \mathfrak{Z}(G^{n+1}, E^*)$

$$d^n F(g_1, \dots, g_{n+1}) = g_1.F(g_2, \dots, g_{n+1}) + \sum_{i=1}^n (-1)^i F(g_1, g_2, \dots, g_i g_{i+1}, \dots, g_n) \\ + (-1)^{n+1} F(g_1, g_2, \dots, g_n)$$

NEF (Natural Exponential Families): Letac & Casalis

Let $Z^n(G, E^*) = \text{Ker}(d^n)$, $B(G, E^*) = \text{Im}(d^{n-1})$, with Z^n inhomogeneous n -cocycles, the quotient $H^n(G, E^*) = Z^n(G, E^*) / B^n(G, E^*)$ is the

Cohomology Group of G with value in E^* . We have:

$$d^0 : E^* \rightarrow \mathfrak{Z}(G, E^*) \quad Z^0 = \{x \in E^* ; g.x = x, \forall g \in G\}$$

$$x \mapsto (g \mapsto g.x - x)$$

$$d^1 : \mathfrak{Z}(G, E^*) \rightarrow \mathfrak{Z}(G^2, E^*)$$

$$F \mapsto d^1 F, \quad d^1 F(g_1, g_2) = g_1.F(g_2) - F(g_1 g_2) + F(g_1)$$

$$Z^1 = \{F \in \mathfrak{Z}(G, E^*) ; F(g_1 g_2) = g_1.F(g_2) + F(g_1), \forall (g_1, g_2) \in G^2\}$$

$$B^1 = \{F \in \mathfrak{Z}(G, E^*) ; \exists x \in E^*, F(g) = g.x - x\}$$

NEF (Natural Exponential Families): Letac & Casalis

■ **When the Cohomology Group** $H^1(G, E^*) = 0$ **then** $Z^1(G, E^*) = B^1(G, E^*)$
 $\Rightarrow \exists c \in E^*$, such that $\forall g \in G, a(g) = (I_d - {}^t g^{-1})c$

Then if $F = F(\mu)$ **is an exponential family invariant by** G , μ **verifies**

$$\forall g \in G, g(\mu)(dx) = e^{\langle c, x \rangle - \langle c, g^{-1}x \rangle + b(g)} \mu(dx)$$

$$\forall g \in G, g\left(e^{\langle c, x \rangle} \mu(dx)\right) = e^{b(g)} e^{\langle c, x \rangle} \mu(dx) \text{ with } \mu_0(dx) = e^{\langle c, x \rangle} \mu(dx)$$

■ **For all compact Group,** $H^1(G, E^*) = 0$ **and we can express a**

$$A: G \rightarrow GA(E) \quad \forall (g, g') \in G^2, A_{gg'} = A_g A_{g'}$$

$$g \mapsto A_g, \quad A_g(\theta) = {}^t g^{-1} \theta + a(g) \quad A(G) \text{ compact sub - group of } GA(E)$$

$$\exists \text{fixed point} \Rightarrow \forall g \in G, A_g(c) = {}^t g^{-1} c + a(g) = c \Rightarrow a(g) = (I_d - {}^t g^{-1})c$$

Lie Group Co-adjoint Orbits & Homogeneous Symplectic Manifold



Lie Group

GROUP (Mathematics)

A set equipped with a binary operation with 4 axioms:

- Closure $\forall a, b \in G$ then $a \bullet b \in G$
- Associativity $\forall a, b, c \in G$ then $(a \bullet b) \bullet c = a \bullet (b \bullet c)$
- Identity $\exists e \in G$ such that $e \bullet a = a \bullet e = a$
- invertibility $\forall a \in G, \exists b \in G$ such that $b \bullet a = a \bullet b = e$



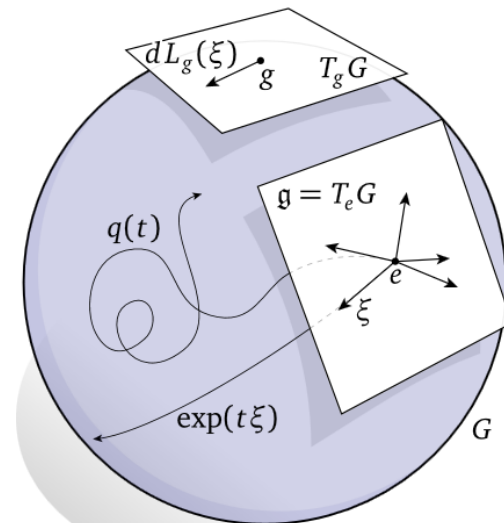
LIE GROUP

- A group that is a differentiable manifold, with the property that the group operations of multiplication and inversion are smooth maps:

$\forall x, y \in G$ then $\phi: G \times G \rightarrow G$ then $\phi(x, y) = x^{-1}y$ is smooth

- A Lie algebra $\mathfrak{g} = T_e G$ is a vector space with a binary operation called the Lie bracket $[\cdot, \cdot]: \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ that satisfies axioms:
 $[ax + by, z] = a[x, z] + b[y, z]$; $[x, x] = 0$; $[x, y] = -[y, x]$

Jacobi Identity: $[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0$



Coadjoint operator and Coadjoint Orbits (Kirillov Representation)

Lie Group Adjoint Representation

- the adjoint representation of a Lie group Ad_g is a way of representing its elements as linear transformations of the Lie algebra, considered as a vector space

$$\Psi : G \rightarrow Aut(G) \quad Ad_g = (d\Psi_g)_e : \mathfrak{g} \rightarrow \mathfrak{g} \quad ad = T_e Ad : T_e G \rightarrow End(T_e G)$$

$$g \mapsto \Psi_g(h) = ghg^{-1} \quad X \mapsto Ad_g(X) = gXg^{-1} \quad X, Y \in T_e G \mapsto ad_X(Y) = [X, Y]$$

Lie Group Co-Adjoint Representation

- the coadjoint representation of a Lie group Ad_g^* , is the dual of the adjoint representation (\mathfrak{g}^* denotes the dual space to \mathfrak{g}):

$$\forall g \in G, Y \in \mathfrak{g}, F \in \mathfrak{g}^*, \text{ then } \langle Ad_g^* F, Y \rangle = \langle F, Ad_{g^{-1}} Y \rangle \quad K = Ad_g^* = (Ad_{g^{-1}})^* \text{ and } K_*(X) = -(ad_X)^*$$

Co-adjoint Orbits as Homogeneous Symplectic Manifold by KKS 2-form

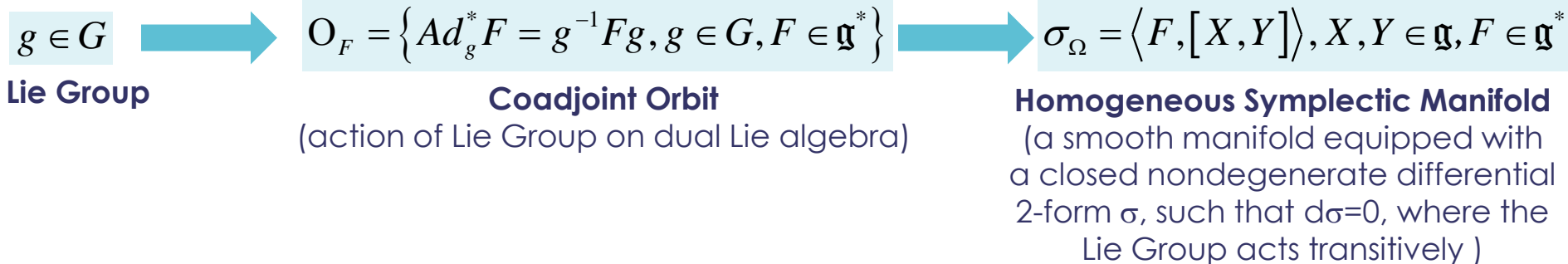
- A coadjoint orbit: $O_F = \{Ad_g^* F = g^{-1} F g, g \in G\}$ subset of $\mathfrak{g}^*, F \in \mathfrak{g}^*$
- carry a natural homogeneous symplectic structure by a closed G-invariant 2-form:

$$\sigma_\Omega(K_{*X} F, K_{*Y} F) = B_F(X, Y) = \langle F, [X, Y] \rangle, X, Y \in \mathfrak{g}. \text{ The coadjoint action on } O_F$$

Souriau Theorem

■ **Souriau Fundamental Theorem** « *Every symplectic manifold is a coadjoint orbit* » is based on classification of symplectic homogeneous Lie group actions by Souriau, Kostant and Kirillov:

- Jean-Marie Souriau. Structure des systèmes dynamiques, Dunod, Paris, 1969.
- Bertram Kostant. Orbits and quantization theory. In Congrès international des mathématiciens 1970.
- Alexandre A. Kirillov. Elements de la théorie des représentations. Editions MIR, Moscou, 1974.



Souriau Invention of « Moment map »: Geometrization of Noether Theorem

As explained in by Thomas Delzant at 2010 CIRM conference “Action Hamiltoniennes: invariants et classification”, organized with Michel Brion:

- “The definition of the moment map is due to Jean-Marie Souriau.... In the book of Souriau, we find a proof of the proposition: the map J is equivariant for an **affine action of G on \mathfrak{g}^*** whose linear part is Ad^* In Souriau's book, we can also find a **study of the non-equivariant case** and its applications to classical and quantum mechanics. In the case of the Galileo group operating in the phase space of space-time, **obstruction to equivariance (a class of cohomology)** is interpreted as the inert mass of the object under study”.
- We can uniquely define the moment map up to an additive constant of integration, that can always be chosen to make the moment map equivariant (a moment map is G -equivariant, when G acts on \mathfrak{g}^* via the coadjoint action) if the group is compact or semi-simple. In 1969, Souriau has considered **the non-equivariant case where the coadjoint action must be modified to make the map equivariant by a 1-cocycle on the group with values in dual Lie algebra \mathfrak{g}^* .**

Interior/Exterior Products and Lie derivative

➤ $i_V \omega$ is the $(p-1)$ -form on X obtained by inserting $V(x)$ as the first argument of ω :

$$\text{Interior product : } i_V \omega(v_2, \dots, v_p) = \omega(V(x), v_2, \dots, v_p)$$

➤ $\theta \wedge \omega$ is the $(p+1)$ -form on X where ω is a p -form and θ is a 1-form on X :

$$\text{Exterior product : } \theta \wedge \omega(v_0, \dots, v_p) = \sum_{i=0}^p (-1)^i \theta(v_i) \omega(v_0, \dots, \hat{v}_i, \dots, v_p)$$

(where the hat indicates a term to be omitted).

➤ $L_V \omega$ is a p -form on X , and $L_V \omega = 0$ if the flow of V consists of symmetries of ω .

$$\text{Lie derivative : } L_V \omega(v_1, \dots, v_p) = \left. \frac{d}{dt} e^{tV^*} \omega(v_1, \dots, v_p) \right|_{t=0}$$

Exterior derivative and E.Cartan, H. Cartan & S. Lie formulas

- $d\omega$ is the $(p+1)$ -form on X defined by taking the ordinary derivative of ω and then antisymmetrizing:

Exterior derivative :
$$d\omega(v_0, \dots, v_p) = \sum_{i=0}^p (-1)^i \frac{\partial \omega}{\partial x} (v_i)(v_0, \dots, \hat{v}_i, \dots, v_p)$$

$$p=0, [d\omega]_i = \partial_i \omega \quad ; \quad p=1, [d\omega]_{ij} = \partial_i \omega_j - \partial_j \omega_i \quad ; \quad p=2, [d\omega]_{ijk} = \partial_i \omega_{jk} + \partial_j \omega_{ki} + \partial_k \omega_{ij}$$

- The properties of the exterior and Lie Derivative are the following:

$$L_V \omega = di_V \omega + i_V d\omega \quad (\text{E. Cartan})$$

$$i_{[U,V]} \omega = i_V L_U \omega - L_U i_V \omega \quad (\text{H. Cartan})$$

$$L_{[U,V]} \omega = L_V L_U \omega - L_U L_V \omega \quad (\text{S. Lie})$$

Souriau Moment Map (1/2)

- Let (M, σ) be a connected symplectic manifold.
- A vector field η on M is called symplectic if its flow preserves the 2-form :

$$L_{\eta}\sigma = 0$$

- If we use Elie Cartan's formula, we can deduce that :

$$L_{\eta}\sigma = di_{\eta}\sigma + i_{\eta}d\sigma = 0$$

- but as $d\sigma = 0$ then $di_{\eta}\sigma = 0$. We observe that the 1-form $i_{\eta}\sigma$ is closed.
- When this 1-form is exact, there is a smooth function $x \mapsto H$ on M with:

$$i_{\eta}\sigma = -dH$$

- This vector field η is called Hamiltonian and could be define as s symplectic gradient :

$$\eta = \nabla_{\text{Symp}} H$$

Souriau Moment Map (2/2)

$$di_{\eta}\sigma = 0$$

$$i_{\eta}\sigma = -dH$$

► We define the Poisson bracket of two functions H, H' by:

$$\{H, H'\} = \sigma(\eta, \eta') = \sigma(\nabla_{\text{Symp}} H', \nabla_{\text{Symp}} H)$$

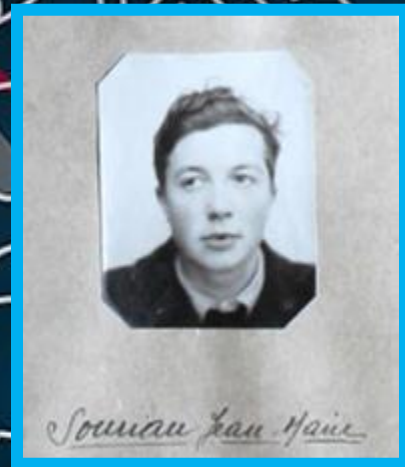
$$\text{with } i_{\eta}\sigma = -dH \text{ and } i_{\eta'}\sigma = -dH'$$

► Let a Lie group G that acts on M and that also preserve σ .

► A moment map exists if these infinitesimal generators are actually hamiltonian, so that a map exists:

$$J : M \rightarrow \mathfrak{g}^* \quad \text{with} \quad i_{Z_x}\sigma = -dH_Z \quad \text{where} \quad H_Z = \langle J(x), Z \rangle$$

Supervised Learning based on Geodesic Natural Gradient on Lie Algebra



Fisher Metric by Misha Gromov (IHES)

M. Gromov, In a Search for a Structure, Part 1: On Entropy. July 6, 2012

➤ <http://www.ihes.fr/~gromov/PDF/structre-serch-entropy-july5-2012.pdf>

Gromov Six Lectures on Probability, Symmetry, Linearity. October 2014, Jussieu, November 6th, 2014

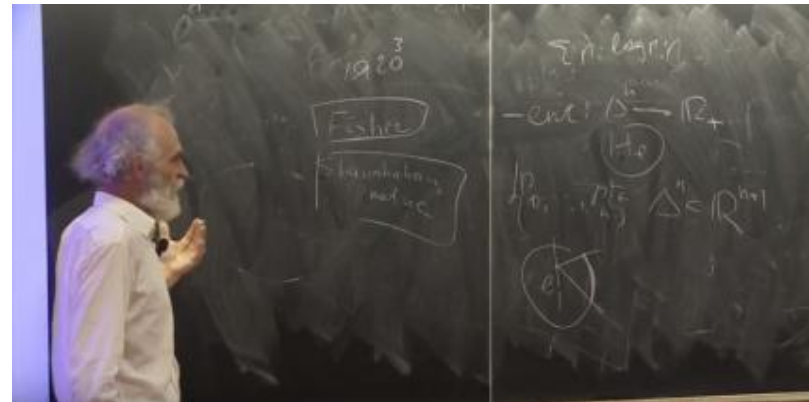
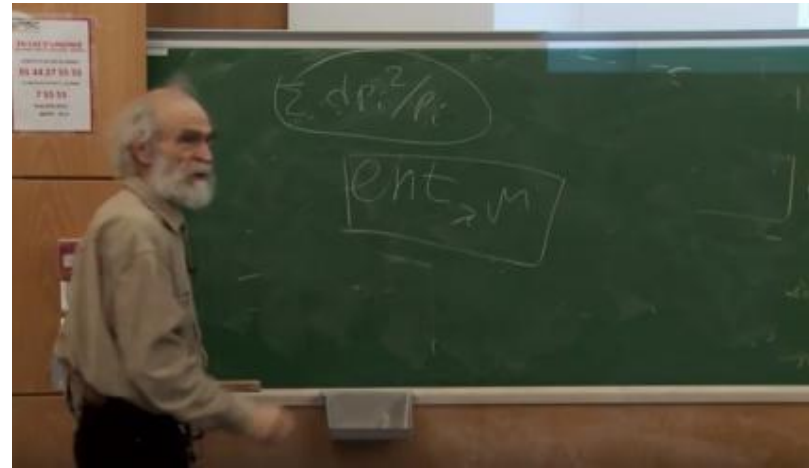
➤ Lecture Slides & video:

<http://www.ihes.fr/~gromov/PDF/probability-huge-Lecture-Nov-2014.pdf>

<https://www.youtube.com/watch?v=hb4D8yMdoV4>

Gromov Four Lectures on Mathematical Structures arising from Genetics and Molecular Biology, IHES, October 2013

<https://www.youtube.com/watch?v=v7QuYuoyLQc&t=5935s>
(at time 01h35min)



Souriau-Fisher Metric and Fréchet-Darmois Bound

■ Cramer-Rao –Fréchet-Darmois Bound has been introduced by Fréchet in 1939 and by Rao in 1945 as inverse of the Fisher Information Matrix: $I(\theta)$

$$R_{\hat{\theta}} = E\left[(\theta - \hat{\theta})(\theta - \hat{\theta})^+\right] \geq I(\theta)^{-1} \quad [I(\theta)]_{i,j} = -E\left[\frac{\partial^2 \log p_{\theta}(z)}{\partial \theta_i \partial \theta_j^*}\right]$$

■ Rao has proposed to introduced an invariant metric in parameter space of density of probabilities (axiomatised by N. Chentsov):

$$ds_{\theta}^2 = \text{Kullback _ Divergence}(p_{\theta}(z), p_{\theta+d\theta}(z))$$

$$ds_{\theta}^2 = -\int p_{\theta}(z) \log \frac{p_{\theta+d\theta}(z)}{p_{\theta}(z)} dz$$

$$ds_{\theta}^2 \underset{\text{Taylor}}{\approx} \sum_{i,j} g_{ij} d\theta_i d\theta_j^* = \sum_{i,j} [I(\theta)]_{i,j} d\theta_i d\theta_j^* = d\theta^+ \cdot I(\theta) \cdot d\theta$$

$$w = W(\theta)$$

$$\Rightarrow ds_w^2 = ds_{\theta}^2$$

Information Geometry & Natural Gradient

- Information geometry has been derived from invariant geometrical structure involved in statistical inference. The Fisher metric defines a Riemannian metric as the Hessian of two dual potential functions, linked to dually coupled affine connections in a manifold of probability distributions. With the Souriau model, this structure is extended preserving the Legendre transform between two dual potential function parametrized in Lie algebra of the group acting transitively on the homogeneous manifold.
- Classically, to optimize the parameter θ of a probabilistic model, based on a sequence of observations y_t , is an online gradient descent with learning rate η_t , and the loss function $l_t = -\log p(y_t / \hat{y}_t)$:

$$\theta_t \leftarrow \theta_{t-1} - \eta_t \frac{\partial l_t(y_t)^T}{\partial \theta}$$

Information Geometry & Natural Gradient

- This simple gradient descent has a first drawback of using the same non-adaptive learning rate for all parameter components, and a second drawback of non invariance with respect to parameter re-encoding inducing different learning rates. Amari has introduced the natural gradient to preserve this invariance to be insensitive to the characteristic scale of each parameter direction. The gradient descent could be corrected by $I(\theta)^{-1}$ where I is the Fisher information matrix with respect to parameter θ , given by:

$$I(\theta) = [g_{ij}]$$

$$\text{with } g_{ij} = \left[-E_{y \sim p(y/\theta)} \left[\frac{\partial^2 \log p(y/\theta)}{\partial \theta_i \partial \theta_j} \right] \right]_{ij} = \left[E_{y \sim p(y/\theta)} \left[\frac{\partial \log p(y/\theta)}{\partial \theta_i} \frac{\partial \log p(y/\theta)}{\partial \theta_j} \right] \right]_{ij}$$

$$\theta_t \leftarrow \theta_{t-1} - \eta_t I(\theta_{t-1})^{-1} \frac{\partial l_t(y_t)^T}{\partial \theta}$$

Natural Gradient & Stochastic Gradient: Natural Langevin Dynamics

- To regularize solution and avoid over-fitting, Stochastic gradient is used, as Langevin Stochastic Gradients
- Yann Ollivier (FACEBOOK FAIR, previously CNRS LRI Orsay) and Gaëtan Marceau-Caron (MILA, previously CNRS LRI Orsay and THALES LAS/ATM & TRT PhD) have proposed to coupled Natural Gradient with Langevin Dynamics: **Natural Langevin Dynamics (Best SMF/SEE GSI'17 paper)**

$$\theta_t \leftarrow \theta_{t-1} - \eta_t I(\theta_{t-1})^{-1} \frac{\partial \left(l_t(y_t)^T - \frac{1}{N} \log \alpha(\theta_{t-1}) \right)}{\partial \theta} + \sqrt{\frac{2\eta_t}{N}} I(\theta_{t-1})^{-1/2} N(0, I_d)$$



- The resulting natural Langevin dynamics combines the advantages of Amari's natural gradient descent and Fisher-preconditioned Langevin dynamics for large neural networks

Dual Entropic Natural Gradient

We can define a natural gradient with dual potential given by Shannon Entropy H (Legendre transform of characteristic function G , logarithm of partition function).

$$\theta_t \leftarrow \theta_{t-1} - \eta_t I(\theta_{t-1})^{-1} \frac{\partial l_t(y_t / \theta)^T}{\partial \theta}$$

$$\theta = \nabla H(\xi) = h(\xi) \quad \xi = \nabla G(\theta) = g(\theta) \quad H(\xi) = \text{Sup}_{\theta \in \Theta} \{ \langle \theta, \xi \rangle - G(\theta) \}$$

$$\xi_t \leftarrow \xi_{t-1} - \alpha_t \frac{\partial l_t(y_t / h(\xi_{t-1}))^T}{\partial \theta}$$

$$\nabla_{\xi} l_t(h(\xi)) = \nabla_{\xi} h(\xi) \nabla_{\theta} l_t(h(\xi)) \Rightarrow \nabla_{\theta} l_t(h(\xi)) = [\nabla_{\xi} h(\xi)]^{-1} \nabla_{\xi} l_t(h(\xi))$$

$$\xi_t \leftarrow \xi_{t-1} - \alpha_t \left[\nabla^2 H(\xi_{t-1}) \right]^{-1} \frac{\partial l_t(y_t / h(\xi_{t-1}))^T}{\partial \xi}$$

Natural Dual Entropic Gradient

Information Geometry, Dual Potentials & Fisher Metric

- Amari has proved that the Riemannian metric in an exponential family is the Fisher information matrix defined by:

$$g_{ij} = - \left[\frac{\partial^2 \Phi}{\partial \theta_i \partial \theta_j} \right]_{ij} \quad \text{with} \quad \Phi(\theta) = - \log \int_{\mathbb{R}} e^{-\langle \theta, y \rangle} dy$$

- and the dual potential, the Shannon entropy, is given by the Legendre transform:

$$S(\eta) = \langle \theta, \eta \rangle - \Phi(\theta) \quad \text{with} \quad \eta_i = \frac{\partial \Phi(\theta)}{\partial \theta_i} \quad \text{and} \quad \theta_i = \frac{\partial S(\eta)}{\partial \eta_i}$$

Koszul-Vinberg Characteristic Function, Koszul Forms

- J.L. Koszul and E. Vinberg have introduced an affinely invariant Hessian metric on a sharp convex cone through its characteristic function

$$\Phi_{\Omega}(\theta) = -\log \int_{\Omega^*} e^{-\langle \theta, y \rangle} dy = -\log \psi_{\Omega}(\theta) \quad \text{with } \theta \in \Omega \text{ sharp convex cone}$$

$$\psi_{\Omega}(\theta) = \int_{\Omega^*} e^{-\langle \theta, y \rangle} dy \quad \text{with Koszul-Vinberg Characteristic function}$$

- 1st Koszul form α : $\alpha = d\Phi_{\Omega}(\theta) = -d \log \psi_{\Omega}(\theta)$

- 2nd Koszul form γ : $\gamma = D\alpha = Dd \log \psi_{\Omega}(\theta)$

$$(Dd \log \psi_{\Omega}(x))(u) = \frac{1}{\psi_{\Omega}(u)^2} \left[\int_{\Omega^*} F(\xi)^2 d\xi \cdot \int_{\Omega^*} G(\xi)^2 d\xi - \left(\int_{\Omega^*} F(\xi) \cdot G(\xi) d\xi \right)^2 \right] > 0 \quad \text{with } F(\xi) = e^{-\frac{1}{2}\langle x, \xi \rangle} \quad \text{and } G(\xi) = e^{-\frac{1}{2}\langle x, \xi \rangle} \langle u, \xi \rangle$$

- Diffeomorphism: $\eta = -\alpha = -d \log \psi_{\Omega}(\theta) = \int_{\Omega^*} \xi p_{\theta}(\xi) d\xi \quad \text{with } p_{\theta}(\xi) = \frac{e^{-\langle \xi, \theta \rangle}}{\int_{\Omega^*} e^{-\langle \xi, \theta \rangle} d\xi}$

- Legendre transform: $S_{\Omega}(\eta) = \langle \theta, \eta \rangle - \Phi_{\Omega}(\theta) \quad \text{with } \eta = d\Phi_{\Omega}(\theta) \quad \text{and } \theta = dS_{\Omega}(\eta)$

Statistical Mechanics, Dual Potentials & Fisher Metric

- In geometric statistical mechanics, Souriau has developed a “**Lie groups thermodynamics**” of dynamical systems where the (maximum entropy) Gibbs density is covariant with respect to the action of the Lie group. In the Souriau model, previous structures of information geometry are preserved:

$$I(\beta) = -\frac{\partial^2 \Phi}{\partial \beta^2} \quad \text{with} \quad \Phi(\beta) = -\log \int_M e^{-\langle \beta, U(\xi) \rangle} d\lambda \quad U : M \rightarrow \mathfrak{g}^*$$

$$S(Q) = \langle \beta, Q \rangle - \Phi(\beta) \quad \text{with} \quad Q = \frac{\partial \Phi(\beta)}{\partial \beta} \in \mathfrak{g}^* \quad \text{and} \quad \beta = \frac{\partial S(Q)}{\partial Q} \in \mathfrak{g}$$

- In the Souriau Lie groups thermodynamics model, β is a “geometric” (Planck) temperature, element of Lie algebra \mathfrak{g} of the group, and Q is a “geometric” heat, element of dual Lie algebra \mathfrak{g}^* of the group.

Statistical Mechanics & Invariant Souriau-Fisher Metric

- In Souriau's Lie groups thermodynamics, the invariance by re-parameterization in information geometry has been replaced by invariance with respect to the action of the group. When an element of the group g acts on the element $\beta \in \mathfrak{g}$ of the Lie algebra, given by adjoint operator Ad_g . Under the action of the group, $Ad_g(\beta)$, **the entropy $S(Q)$ and the Fisher metric $I(\beta)$ are invariant:**

$$\beta \in \mathfrak{g} \rightarrow Ad_g(\beta) \Rightarrow \begin{cases} S[Q(Ad_g(\beta))] = S(Q) \\ I[Ad_g(\beta)] = I(\beta) \end{cases}$$

$$I(\beta) = -\frac{\partial^2 \Phi}{\partial \beta^2} \text{ with } \Phi(\beta) = -\log \int_M e^{-\langle \beta, U(\xi) \rangle} d\lambda$$

$$S(Q) = \langle \beta, Q \rangle - \Phi(\beta) \text{ with } Q = \frac{\partial \Phi(\beta)}{\partial \beta} \in \mathfrak{g}^* \text{ and } \beta = \frac{\partial S(Q)}{\partial Q} \in \mathfrak{g}$$

Statistical Mechanics & Fisher Metric

- Souriau has proposed a Riemannian metric that we have identified as a generalization of the Fisher metric:

$$I(\beta) = [g_\beta] \quad \text{with} \quad g_\beta([\beta, Z_1], [\beta, Z_2]) = \tilde{\Theta}_\beta(Z_1, [\beta, Z_2])$$

$$\text{with} \quad \tilde{\Theta}_\beta(Z_1, Z_2) = \tilde{\Theta}(Z_1, Z_2) + \langle Q, ad_{Z_1}(Z_2) \rangle \quad \text{where} \quad ad_{Z_1}(Z_2) = [Z_1, Z_2]$$

- The tensor $\tilde{\Theta}$ used to define this extended Fisher metric is defined by the moment map $J(x)$, from M (homogeneous symplectic manifold) to the dual Lie algebra \mathfrak{g}^* , given by:

$$\tilde{\Theta}(X, Y) = J_{\downarrow[X, Y]} - \{J_X, J_Y\} \quad \text{with} \quad J(x): M \rightarrow \mathfrak{g}^* \quad \text{such that} \quad J_X(x) = \langle J(x), X \rangle, \quad X \in \mathfrak{g}$$

- This tensor $\tilde{\Theta}$ is also defined in tangent space of the cocycle $\theta(g) \in \mathfrak{g}^*$ (this cocycle appears due to the non-equivariance of the coadjoint operator Ad_g^* , action of the group on the dual lie algebra): $Q(Ad_g(\beta)) = Ad_g^*(Q) + \theta(g)$

$$\tilde{\Theta}(X, Y): \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{R}$$

$$\text{with} \quad \Theta(X) = T_e \theta(X(e))$$

$$X, Y \mapsto \langle \Theta(X), Y \rangle \quad \text{OPEN}$$

Souriau Definition of Souriau Metric as non-equivariant extension of KKS 2 form

■ Souriau definition of Fisher Metric is related to the extension of KKS 2-form (Kostant-Kirillov-Souriau) in case of non-equivariant Cohomology:

Souriau-Fisher Metric

$$I(\beta) = [g_\beta] \quad \text{with} \quad g_\beta([\beta, Z_1], [\beta, Z_2]) = \tilde{\Theta}_\beta(Z_1, [\beta, Z_2])$$

$$\text{with} \quad \tilde{\Theta}_\beta(Z_1, Z_2) = \tilde{\Theta}(Z_1, Z_2) + \langle Q, [Z_1, Z_2] \rangle$$

Non-Equivariant Case: additional term from Souriau Cocycle

Equivariant KKS 2 form

$$\tilde{\Theta}(X, Y) = J_{[X, Y]} - \{J_X, J_Y\} \quad \text{with} \quad J(x) : M \rightarrow \mathfrak{g}^* \quad \text{such that} \quad J_X(x) = \langle J(x), X \rangle, \quad X \in \mathfrak{g}$$

$$\tilde{\Theta}(X, Y) : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{R}$$

$$\text{with} \quad \Theta(X) = T_e \theta(X(e))$$

$$\tilde{\Theta}(\beta, Z) + \langle Q, [\beta, Z] \rangle = 0$$

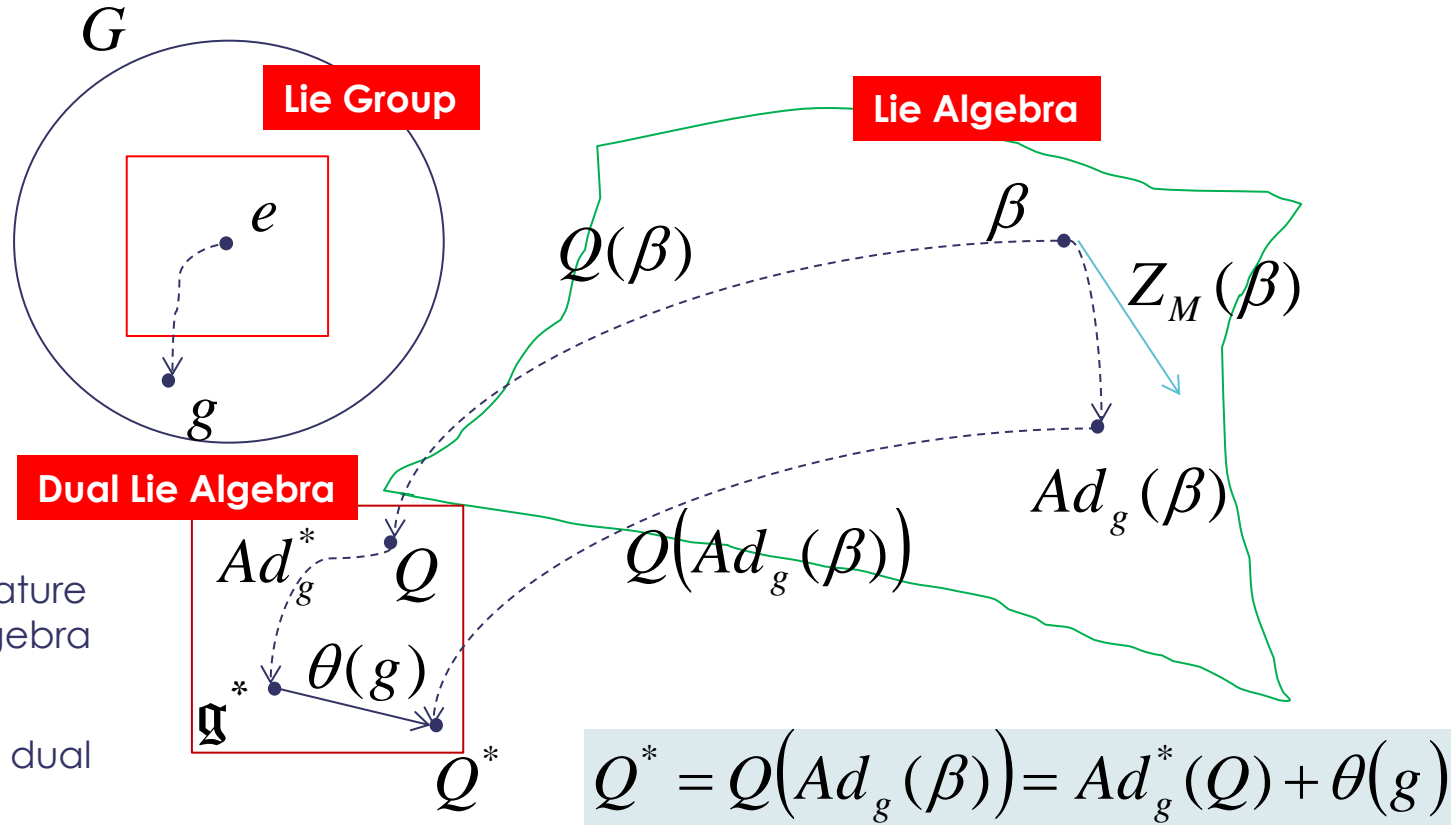
$$X, Y \mapsto \langle \Theta(X), Y \rangle$$

$$\beta \in \text{Ker } \tilde{\Theta}_\beta$$

Souriau Fundamental Equation of Lie Group Thermodynamics

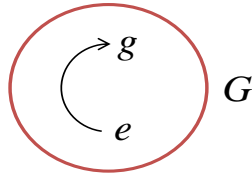
$$Q(Ad_g(\beta)) = Ad_g^*(Q) + \theta(g)$$

Fundamental Souriau Theorem

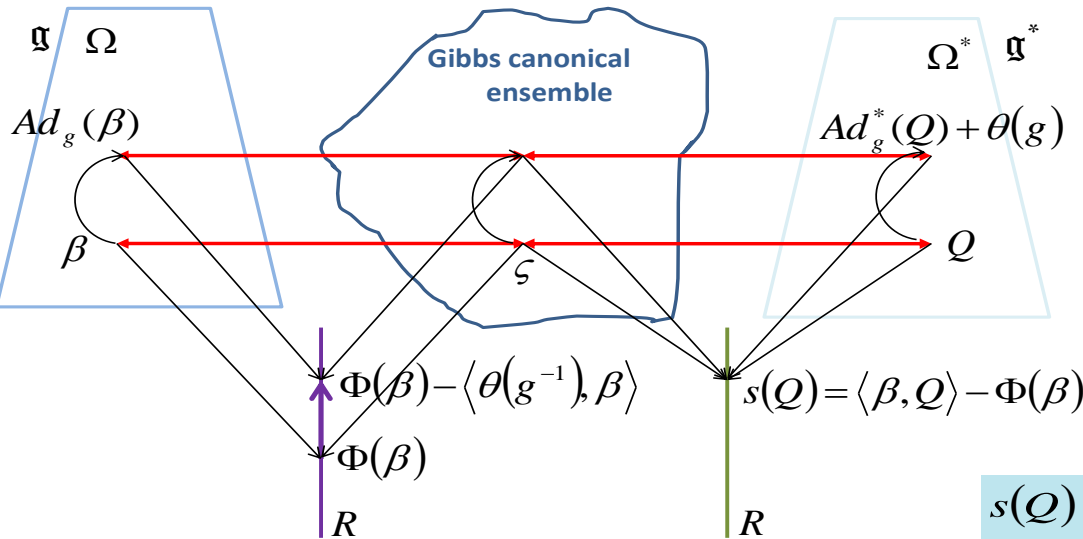


Souriau-Fisher Metric & Souriau Lie Groups Thermodynamics: Bedrock for Lie Group Machine Learning

**TEMPERATURE
In Lie Algebra**



**HEAT
In Dual Lie Algebra**



Fisher Metric

$$g_\beta([\beta, Z_1], [\beta, Z_2]) = \tilde{\Theta}_\beta(Z_1, [\beta, Z_2]) \geq 0$$

$$I(\beta) = I(Ad_g(\beta)) = -\frac{\partial^2 \Phi}{\partial \beta^2} = \frac{\partial^2 \log \int e^{-\langle \beta, U(\xi) \rangle} d\xi}{\partial \beta^2}$$

Entropy invariant under the action of the group

Legendre

$$s(Q) = \langle \beta, Q \rangle - \Phi(\beta) = \langle \Theta^{-1}(Q), Q \rangle - \Phi(\Theta^{-1}(Q))$$

Clairaut

$$Q = \Theta(\beta) = \frac{\partial \Phi}{\partial \beta} \in \mathfrak{g}^*$$

$$\beta = \Theta^{-1}(Q) \in \mathfrak{g}$$

**Logarithm of Partition Function
(Massieu Characteristic Function)**

Entropy

Souriau Lie Groups Thermodynamics: Geometric Calorific Capacity

Nous prenons désormais Z dans C . La valeur moyenne du moment $\psi(x)$ dans l'état de Gibbs est égal à la dérivée

$$Q = z'(Z);$$

$Z \mapsto Q$ est un difféomorphisme analytique de C sur un ouvert convexe de \mathfrak{g}^* ; la transformée de Legendre s de z :

$$s(Q) = QZ - z$$

y est convexe et vérifie $Z = s'(Q)$; la dérivée seconde:

$$K = z''(Z)$$

est un tenseur positif, dont l'inverse est égal à $s''(Q)$.

K munit l'ensemble C d'une structure riemannienne invariante par l'action du groupe; pour cette structure, l'application linéaire $Ad(Z)$ est antihermitienne.

L'application f_Z , définie par:

$$f_Z(Z', Z'') = K([Z, Z'], Z'') \quad \forall Z', Z'' \in \mathfrak{g}$$

est un cocycle symplectique, cohomologue à f [formule (2,7 C)]; son noyau est l'orthogonal de l'orbite adjointe de Z pour la structure riemannienne de C .

Souriau-Fisher Metric is a Geometrization of Thermodynamical «Calorific Capacity» (Pierre Duhem has deeply developed this idea of « generalized capacities »)

■ Dans le cas classique, on ne considère que le groupe de dimension 1 des translations temporelles (qui n'est défini qu'après avoir choisi un référentiel - par exemple celui de la boîte qui contient le gaz). Alors, avec des unités convenables, Z est l'inverse de la TEMPERATURE ABSOLUE; z est le POTENTIEL THERMODYNAMIQUE DE PLANCK; $-s$ est l'ENTROPIE; Q est l'ENERGIE INTERNE; K caractérise la CAPACITE CALORIFIQUE. ■

Souriau Lie Group Thermodynamics: Geometric Calorific Capacity

Il faut bien entendu que cette intégrale soit convergente ; nous définirons l'ensemble canonique de Gibbs Ω comme le plus grand ouvert (dans l'algèbre de Lie) où cette intégrale est localement normalement convergente (en Θ). On montre que Ω est convexe, et que z est une fonction C^∞ sur Ω ; que la dérivée $Q = \frac{\partial z}{\partial \Theta}$ coïncide avec la valeur moyenne de l'énergie E (Q généralise donc la chaleur) ; que le tenseur $\frac{\partial Q}{\partial \Theta}$ est symétrique et positif (il généralise la capacité calorifique). Il en résulte que z est fonction convexe de Θ ; la transformation de Legendre lui associe une fonction concave, à savoir

$$(7.3) \quad Q \mapsto s = z - Q\Theta$$

s est l'entropie.

Souriau-Fisher Metric based on cocycle

pour chaque "température" Θ , définissons un tenseur f_Θ , somme du cocycle f (défini en (3.2)) et du cobord de la chaleur :

$$(7.4) \quad f_\Theta(z, z') = f(z, z') + Q[z, z']$$

f_Θ jouit alors des propriétés suivantes :

- a) f_Θ est un cocycle symplectique ;
- b) $\Theta \in \ker f_\Theta$
- 5) c) Le tenseur symétrique g_Θ , défini sur l'ensemble de valeurs de $\text{ad}(\Theta)$ par

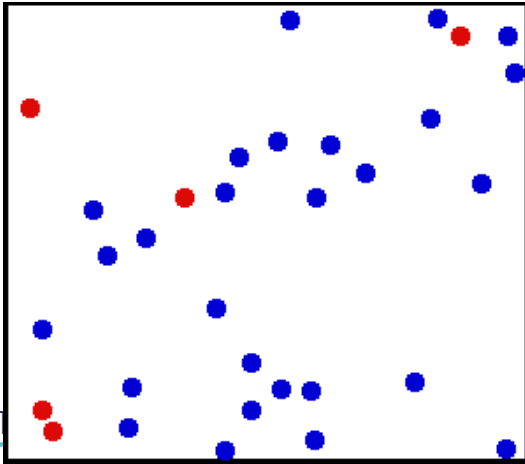
$$g_\Theta([\Theta, z], [\Theta, z']) = f_\Theta(z, [\Theta, z'])$$

est positif (et même défini positif si l'action du groupe est effective).

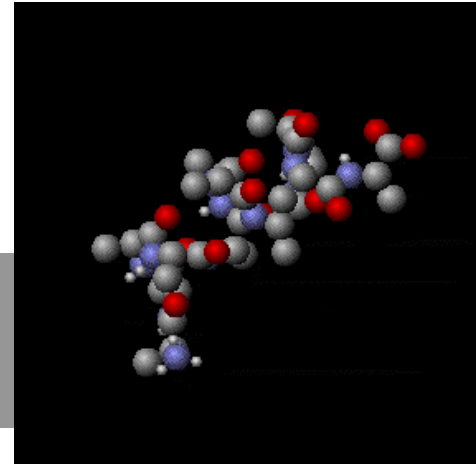
Ces formules sont universelles, en ce sens qu'elles ne mettent pas en jeu la variété symplectique U - mais seulement le groupe G , son cocycle symplectique f et les couples Θ, Q . Peut-être cette "thermodynamique des groupes de Lie" a-t-elle un intérêt mathématique.

Souriau Model of Lie Groups Thermodynamics

- Souriau Geometric (Planck) Temperature is **an element of Lie Algebra** of Dynamical Group (Galileo/Poincaré groups) acting on the system
- Generalized Entropy is **Legendre Transform of minus logarithm of Laplace Transform**
- Fisher(-Souriau) Metric is a **Geometric Calorific Capacity (hessian of Massieu Potential)**
- Higher Order Souriau Lie Groups Thermodynamics is given by **Günther's Poly-Symplectic Model (vector-valued model in non-equivariant case)**



Souriau formalism is fully **covariant**, with no special coordinates (**covariance of Gibbs density wrt Dynamical Groups**)



Massieu Potential versus Gibbs Potentials

GIBBS Potential: Free Energy

$$~~F = E - TS~~$$

Tout mathématicien sait qu'il est impossible de comprendre un cours élémentaire en thermodynamique.

Vladimir Arnold

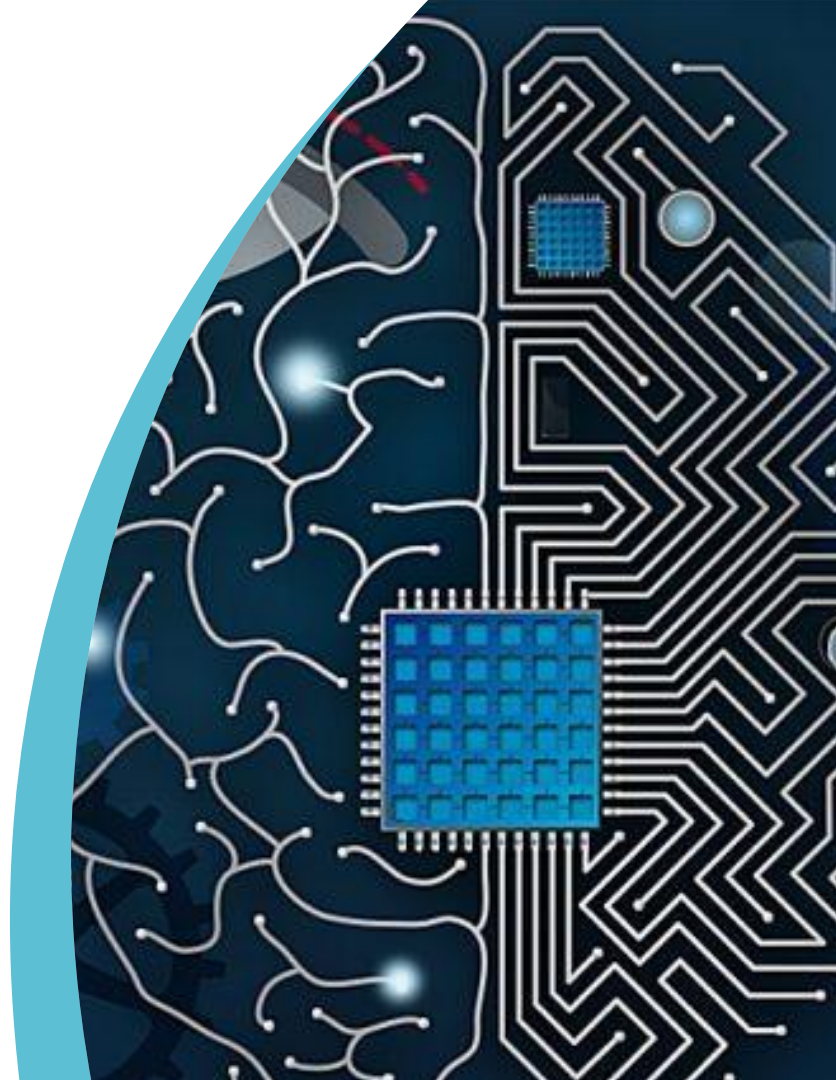
MASSIEU Potential : characteristic function

$$\frac{F}{T} = \frac{1}{T} E - S \quad \Rightarrow \quad \Psi = \left\langle \beta, E \right\rangle - S$$

$\beta = \frac{1}{T}$

Preservation of Legendre Duality

Souriau Verbatim: “Quinta Essentia” momentum-movimentum



Souriau Quinta Essentia (Quinte Essence)

- “Plaçons-nous d’abord dans le cadre de la mécanique classique. Étudions un système mécanique isolé, non dissipatif—nous dirons brièvement une «chose». L’ensemble des mouvements de cette «chose» est une variété symplectique. Pourquoi ? Il suffit de se reporter à la Mécanique Analytique de Lagrange (1811); l’espace des mouvements y est traité comme variété différentiable; les coordonnées covariantes et contravariantes de la forme symplectique y sont écrites (Ce sont les “parenthèses” et “crochets” de Lagrange). Évoquons maintenant la géométrie du 20^{ème} siècle. Soit G un groupe difféologique (par exemple un groupe de Lie); μ un moment de G (un moment, c’est une 1-forme invariante à gauche sur G); alors l’action du groupe sur μ engendre canoniquement un espace symplectique (ces groupes pourront avoir une dimension infinie). Présomption épistémologique: derrière chaque «chose» est caché un groupe G (sa “source”), et les mouvements de la «chose» sont simplement des moments de G (doublet latin mnémotechnique : momentum-movimentum). L’isolement de la «chose» indique alors que le groupe de Poincaré (respectivement de Galilée-Bargmann) est inséré dans G ; voilà l’origine des grandeurs conservées relativistes (respectivement classiques) associées à un mouvement x : elles constituent simplement le moment induit sur le groupe spatio-temporel par le moment-mouvement x .”
- Quantique? Alors c’est géométrie

Souriau Quinta Essentia (Quinte Essence)

- *“Il y a un théorème qui remonte au XXème siècle. Si on prend une orbite coadjointe d'un groupe de Lie, elle est pourvue d'une structure symplectique. Voici un algorithme pour produire des variétés symplectiques : prendre des orbites coadjointes d'un groupe. Donc cela laisse penser que derrière cette structure symplectique de Lagrange, il y avait un groupe caché. Prenons le mouvement classique d'un moment du groupe, alors ce groupe est très «gros» pour avoir tout le système solaire. Mais dans ce groupe est inclus le groupe de Galilée, et tout moment d'un groupe engendre des moments d'un sous-groupe. On va retrouver comme cela les moments du groupe de Galilée, et si on veut de la mécanique relativiste, cela va être celui du groupe de Poincaré. En fait avec le groupe de Galilée, il y a un petit problème, ce ne sont pas les moments du groupe de Galilée qu'on utilise, ce sont les moments d'une extension centrale du groupe de Galilée, qui s'appelle le groupe de Bargmann, et qui est de dimension 11. C'est à cause de cette extension, qu'il y a cette fameuse constante arbitraire figurant dans l'énergie. Par contre quand on fait de la relativité restreinte, on prend le groupe de Poincaré et il n'y a plus de problèmes car parmi les moments il y a la masse et l'énergie c'est mc^2 . Donc le groupe de dimension 11 est un artéfact qui disparaît, quand on fait de la relativité restreinte.”*

Souriau Quinta Essentia (Quinte Essence)

- *“Je me suis dit, à force de rencontrer des groupes, il y a quelque chose de caché là-dessous. La catégorie métaphysique des groupes qui plane dans l’empyrée des mathématiques, que nous découvrons et que nous adorons, elle doit se rattacher à quelque chose de plus proche de nous. En écoutant de nombreux exposés faits par des neurophysiologistes, j’ai fini par apprendre le rôle primitif du déplacement des objets. Nous savons manipuler ces déplacements mentalement avec une très grande virtuosité. Ce qui nous permet de nous manipuler nous-même, de marcher, de courir, de sauter, de nous rattraper quand nous tombons, etc. Ce n’est pas vrai seulement pour nous, c’est vrai aussi pour les singes ; ils sont beaucoup plus adroits que nous pour anticiper les résultats d’un déplacement. Pour certaines opérations élémentaires de «lecture», ils vont même dix fois plus vite que nous. Beaucoup de neurophysiologistes pensent qu’il y a une structure spéciale génétiquement inscrite dans le cerveau, le câblage d’un groupe ... Lorsque il y un tremblement de terre, nous assistons à la mort de l’Espace. ... Nous vivons avec nos habitudes que nous pensons universelles. ... La neuroscience s’occupe rarement de la géométrie ... Pour les singes qui vivent dans les arbres, certaines propriétés du groupe d’Euclide sont mieux câblées dans leurs cerveaux.”*

Koszul Book Introduction to Symplectic Geometry



Jean-Marie Souriau by Jean-Louis Koszul

- Both Souriau and Koszul followed last lectures of Elie Cartan and were inspired to use Affine Representation of Lie Group and Lie Algebra
- « A l'époque où Souriau développait sa théorie, l'establishment avait tendance à ne pas y voir des avancées importantes. Je l'ai entendu exposer ses idées sur la thermodynamique mais je n'ai pas du tout réalisé à l'époque que la géométrie hessienne était en jeu. », J-L. Koszul 2018
- « At the time when Souriau was developing his theory, the establishment tended not to see significant progress. I heard him explaining his ideas on thermodynamics but I did not realize at the time that Hessian geometry was at stake. », J-L. Koszul 2018

Souriau & Koszul at 1953 Conference « Topologie différentielle » in Strasbourg



J.L. Koszul



J.M. Souriau

Koszul Works and links with Souriau Work

- In 1986, "Introduction to symplectic geometry" book following a Chinese Koszul course in China (translated into English by Springer in 2019).
- This book takes up and develops works of Jean-Marie Souriau on homogeneous symplectic manifolds. Chuan Yu Ma writes in a review, on this book in Chinese, that "***This work coincided with developments in the field of analytical mechanics. Many new ideas have also been derived using a wide variety of notions of modern algebra, differential geometry, Lie groups, functional analysis, differentiable manifolds, and representation theory. [Koszul's book] emphasizes the differential-geometric and topological properties of symplectic manifolds. It gives a modern treatment of the subject that is useful for beginners as well as for experts***".

TABLE OF CONTENTS

1 Some Algebra Basics	1
1.1 Skew-symmetric forms	1
1.2 Symplectic vector spaces, symplectic bases	8
1.3 The canonical linear representation of $sl(2, k)$ in the algebra of the skew-symmetric forms on a symplectic vector space	10
1.4 Symplectic groups	15
1.5 Symplectic complex structures	21
2 Symplectic Manifolds	25
2.1 Symplectic structures on manifolds	26
2.2 Operators of the algebra of differential forms on a symplectic manifold	31
2.3 Symplectic coordinates	37
2.4 Hamiltonian vector fields and symplectic vector fields	43
2.5 Poisson brackets under symplectic coordinates	54
2.6 Submanifolds of symplectic manifolds	59
3 Cotangent Bundles	69
3.1 Liouville forms and canonical symplectic structures on cotangent bundles	69
3.2 Symplectic vector fields on a cotangent bundle	74
3.3 Lagrangian submanifolds of a cotangent bundle	83
4 Symplectic G-spaces	91
4.1 Definitions and examples	93
4.2 Hamiltonian g -spaces and moment maps	97
4.3 Equivariance of moment maps	108
5 Poisson Manifolds	113
5.1 The structure of a Poisson manifold	113
5.2 The leaves of a Poisson manifold	119
5.3 Poisson structures on the dual of a Lie algebra	124



iii


Koszul Book on Souriau Work

Jean-Louis Koszul · Yi Ming Zou

Introduction to Symplectic Geometry

Forewords by Michel Nguiffo Boyom, Frédéric Barbaresco and Charles-Michel Marle

This introductory book offers a unique and unified overview of symplectic geometry, highlighting the differential properties of symplectic manifolds. It consists of six chapters: Some Algebra Basics, Symplectic Manifolds, Cotangent Bundles, Symplectic G-spaces, Poisson Manifolds, and A Graded Case, concluding with a discussion of the differential properties of graded symplectic manifolds of dimensions (o, n) . It is a useful reference resource for students and researchers interested in geometry, group theory, analysis and differential equations. This book is also inspiring in the emerging field of Geometric Science of Information, in particular the chapter on Symplectic G-spaces, where Jean-Louis Koszul develops Jean-Marie Souriau's tools related to the non-equivariant case of co-adjoint action on Souriau's moment map through Souriau's Cocycle, opening the door to Lie Group Machine Learning with Souriau-Fisher metric.


$$\begin{aligned}\mu : M &\longrightarrow \mathfrak{g}^* \\ \mu(sx) &= s\mu(x) = \text{Ad}^*(s)\mu(x) + \phi_\mu(s), \quad \forall s \in G, x \in M \\ c_\mu(a, b) &= \{\langle \mu, a \rangle, \langle \mu, b \rangle\} - \langle \mu, [a, b] \rangle = \langle d\phi_\mu(a), b \rangle, \quad \forall a, b \in \mathfrak{g}\end{aligned}$$

Jean-Louis Koszul
Yiming Zou

Introduction to Symplectic Geometry


$$\begin{aligned}\mu : M &\longrightarrow \mathfrak{g}^* \\ \mu(sx) &= s\mu(x) = \text{Ad}^*(s)\mu(x) + \phi_\mu(s), \quad \forall s \in G, x \in M \\ c_\mu(a, b) &= \{\langle \mu, a \rangle, \langle \mu, b \rangle\} - \langle \mu, [a, b] \rangle = \langle d\phi_\mu(a), b \rangle, \quad \forall a, b \in \mathfrak{g}\end{aligned}$$

 Science Press
Beijing

 Springer

<https://link.springer.com/book/10.1007%2F978-981-13-3987-5>

ENS ULM Séminaire Mathématique et Philosophie

OPEN

New Koszul Definition of Souriau- Fisher Metric (1/3)

➤ Notations: $Ad_s a = sas^{-1}$, $s \in G, a \in \mathfrak{g}$, $ad_a b = [a, b]$, $a \in \mathfrak{g}, b \in \mathfrak{g}$

$$Ad_s^* = {}^t Ad_{s^{-1}} \quad , \quad s \in G$$

➤ Properteis: $Ad_{\exp a} = \exp(-ad_a)$, $a \in \mathfrak{g}$ $Ad_{\exp a}^* = \exp^t(ad_a)$, $a \in \mathfrak{g}$

➤ Properties of moment map μ : $\mu: M \rightarrow \mathfrak{g}^*$ $x \mapsto sx$, $x \in M$

$$\langle d\mu(v), a \rangle = \omega(ax, v)$$

$$d\langle Ad_s^* \circ \mu, a \rangle = \langle Ad_s^* d\mu, a \rangle = \langle d\mu, Ad_{s^{-1}} a \rangle$$

$$\langle d\mu(v), Ad_{s^{-1}} a \rangle = \omega(s^{-1}asx, v) = \omega(asx, sv) = \langle d\mu(sv), a \rangle = (d\langle \mu \circ s_M, a \rangle)(v)$$

$$d\langle Ad_s^* \circ \mu, a \rangle = d\langle \mu \circ s_M, a \rangle \Rightarrow d\langle \mu \circ s_M - Ad_s^* \circ \mu, a \rangle = 0$$

New Koszul Definition of Souriau- Fisher Metric (2/3)

► Symplectic Cocycle: $\theta_\mu(s) = \mu(sx) - Ad_s^* \mu(x)$, $s \in G$

$$\theta_\mu(st) = \mu(stx) - Ad_{st}^* \mu(x) = \theta_\mu(s) + Ad_s^* \mu(tx) - Ad_s^* Ad_t^* \mu(x)$$

$$\theta_\mu(st) = \theta_\mu(s) + Ad_s^* \theta_\mu(t)$$

► Group action on dual Lie Algebra : $G \times \mathfrak{g}^* \rightarrow \mathfrak{g}^*$, $(s, \xi) \mapsto s\xi = Ad_s^* \xi + \theta_\mu(s)$

$$\mu(sx) = s\mu(x) = Ad_s^* \mu(x) + \theta_\mu(s) , \quad \forall s \in G, x \in M$$

$$\theta_\mu(s) = \mu(sx) - Ad_s^* \mu(x)$$

► Properties: $G \times \mathfrak{g}^* \rightarrow \mathfrak{g}^*$, $(e, \xi) \mapsto e\xi = Ad_e^* \xi + \theta_\mu(e) = \xi + \mu(x) - \mu(x) = \xi$

$$(s_1 s_2) \xi = Ad_{s_1 s_2}^* \xi + \theta_\mu(s_1 s_2) = Ad_{s_1}^* Ad_{s_2}^* \xi + \theta_\mu(s_1) + Ad_{s_1}^* \theta_\mu(s_2)$$

$$(s_1 s_2) \xi = Ad_{s_1}^* \left(Ad_{s_2}^* \xi + \theta_\mu(s_2) \right) + \theta_\mu(s_1) = s_1 (s_2 \xi) , \quad \forall s_1, s_2 \in G, \xi \in \mathfrak{g}^*$$

New Koszul Definition of Souriau- Fisher Metric (3/3)

► Koszul definition of Souriau 2-cocycle: $c_\mu(a, b) = \langle d\theta_\mu(a), b \rangle$, $a, b \in \mathfrak{g}$

$$d\mu(ax) = {}^t ad_a \mu(x) + d\theta_\mu(a), \quad x \in M, a \in \mathfrak{g}$$

$$\langle d\mu(ax), b \rangle = \langle \mu(x), [a, b] \rangle + \langle d\theta_\mu(a), b \rangle = \{ \langle \mu, a \rangle, \langle \mu, b \rangle \}(x), \quad x \in M, a, b \in \mathfrak{g}$$

$$c_\mu(a, b) = \{ \langle \mu, a \rangle, \langle \mu, b \rangle \} - \langle \mu, [a, b] \rangle = \langle d\theta_\mu(a), b \rangle, \quad a, b \in \mathfrak{g}$$

$$c_\mu([a, b], c) + c_\mu([b, c], a) + c_\mu([c, a], b) = 0, \quad a, b, c \in \mathfrak{g}$$

$$\{ \mu^*(a), \mu^*(b) \} = \{ \langle \mu, a \rangle, \langle \mu, b \rangle \} = \mu^*([a, b] + c_\mu(a, b)) = \mu^*\{a, b\}_{c_\mu}$$

► Property: $\mu' = \mu + \varphi \Rightarrow c_{\mu'}(a, b) = c_\mu(a, b) - \langle \varphi, [a, b] \rangle$

Souriau model in SSD book

$$\tilde{\Theta}(X, Y) = J_{[X, Y]} - \{J_X, J_Y\}$$

$J : M \rightarrow \mathfrak{g}^*$ with $x \mapsto J(x)$

such that $J_X(x) = \langle J(x), X \rangle$, $X \in \mathfrak{g}$

$$\tilde{\Theta}([X, Y], Z) + \tilde{\Theta}([Y, Z], X) + \tilde{\Theta}([Z, X], Y) = 0$$

$$\tilde{\Theta}(X, Y) : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{R}$$

$$X, Y \mapsto \langle \Theta(X), Y \rangle$$

with $\Theta(X) = T_e \theta(X(e))$

$$\tilde{\Theta}_\beta(Z_1, Z_2) = \tilde{\Theta}(Z_1, Z_2) + \langle Q, ad_{Z_1}(Z_2) \rangle$$

with $ad_{Z_1}(Z_2) = [Z_1, Z_2]$

Koszul development in ITSG book

$$c_\mu(a, b) = \{ \langle \mu, a \rangle, \langle \mu, b \rangle \} - \langle \mu, [a, b] \rangle$$

$$c_\mu([a, b], c) + c_\mu([b, c], a) + c_\mu([c, a], b) = 0$$

$$c_\mu(a, b) = \langle d\theta_\mu(a), b \rangle, \quad a, b \in \mathfrak{g}$$

$$\mu' = \mu + \varphi \Rightarrow c_{\mu'}(a, b) = c_\mu(a, b) - \langle \varphi, [a, b] \rangle$$

Souriau Lie Groups Thermodynamics & Covariant Gibbs Density



Gallileo Group & Alebra & V. Bargman Central extensions

■ Symplectic cocycles of the Galilean group: V. Bargmann (Ann. Math. 59, 1954, pp 1–46) has proven that the symplectic cohomology space of the Galilean group is one-dimensional.

■ Gallileo Lie Group & Algebra

$$\begin{cases} \vec{x}' = R \cdot \vec{x} + \vec{u} \cdot t + \vec{w} \\ t' = t + e \end{cases}$$

$$\vec{x}, \vec{u} \text{ and } \vec{w} \in R^3, e \in R^+$$

$$R \in SO(3)$$

■ Bargmann Central extension:

$$\begin{bmatrix} R & \vec{u} & 0 & \vec{w} \\ 0 & 1 & 0 & e \\ -\vec{u}' R & -\frac{\|\vec{u}\|^2}{2} & 1 & f \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Homogeneous Gallileo group

$$\begin{bmatrix} \vec{x}' \\ t' \\ 1 \end{bmatrix} = \begin{bmatrix} R & \vec{u} & \vec{w} \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \vec{x} \\ t \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} \vec{\omega} & \vec{\eta} & \vec{\gamma} \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{bmatrix}, \begin{cases} \vec{\eta} \text{ and } \vec{\gamma} \in R^3, \varepsilon \in R^+ \\ \vec{\omega} \in so(3) : \vec{x} \mapsto \vec{\omega} \times \vec{x} \end{cases}$$

Covariant Gibbs Equilibrium

Jean-Marie Souriau has observed in 1966 in « *Définition covariante des équilibres thermodynamiques* » that **Classical Gibbs Equilibrium is not covariant with respect to Dynamic Groups** (Galilee Group in classical Mechanics or Poincaré Group in Relativity). Classical thermodynamics corresponds to the case of Time translation.

To solve this incoherency, Souriau has extended definition of Canonical Gibbs Ensemble to Symplectic Manifolds on which a Lie Group has a Symplection Action:

- (Planck) Temperature is an element of the Dynamical Group Lie Algebra
- Heat is an element of the Dynamic Group Dual Lie Algebra

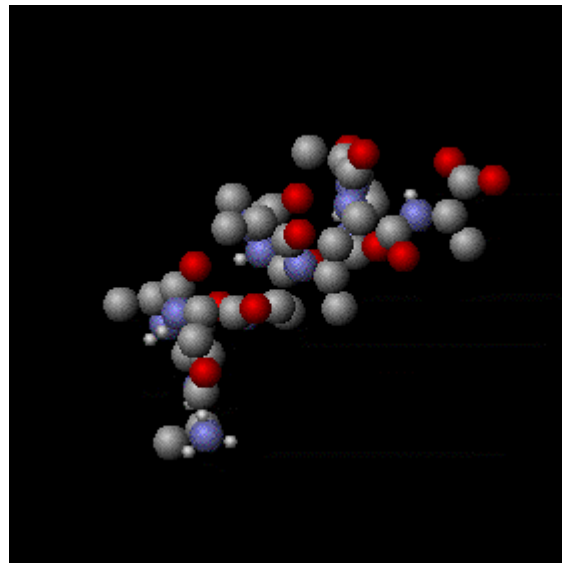
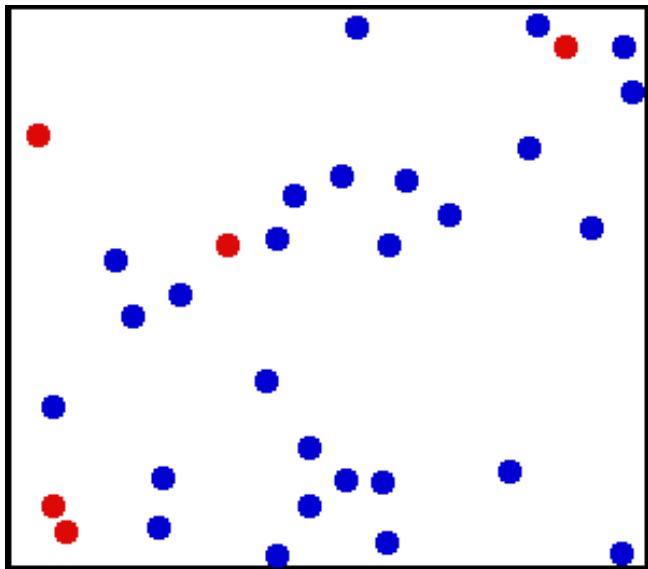
In case of non-commutative groups, specific properties appear: the symmetry is spontaneously broken, some cohomological type of relationships are satisfied in the algebra of the Lie group

Gibbs Canonical Ensemble on Symplectic Manifold

- In statistical mechanics, a canonical ensemble is the statistical ensemble that is used to represent the possible states of a mechanical system that is being maintained in thermodynamic equilibrium.
- Souriau has extended this notion of **Gibbs canonical ensemble on Symplectic manifold M for a Lie group action on M**
- The seminal idea of Lagrange was to consider that a **statistical state is simply a probability measure on the manifold of motions**
- In Jean-Marie Souriau approach, one movement of a dynamical system (classical state) is a point on manifold of movements.
- For statistical mechanics, the movement variable is replaced by a random variable where a statistical state is probability law on this manifold.

What is a Temperature ? Is there an unknown structure behind it ?

« (Planck) Temperature is an element of Lie Algebra of Dynamical Group acting on the system » - Jean-Marie Souriau



Geometric (Planck) Temperature in the Lie Algebra

Let a Group G of a Manifold M with a moment map E , the **Geometric (Planck) Temperature** β is all elements of Lie Algebra \mathfrak{g} of G such that the following integrals converges in a neighborhood of β :

$$I_0(\beta) = \int_M e^{-\langle \beta, U \rangle} d\lambda$$

> $\langle \beta, U \rangle$ notes the duality of \mathfrak{g} and \mathfrak{g}^*

> $d\lambda$ is the Liouville density on M

Theorem: The function I_0 is infinitely differentiable C^∞ in Ω (the largest open proper subset of \mathfrak{g}) and is n^{th} derivative for all $\beta \in \Omega$, the tensor integral is convergent:

$$I_n(\beta) = \int_M e^{-\langle \beta, U \rangle} U^{\otimes n} d\lambda$$

To each temperature β , we can associate probability law on M with distribution function (such that the probability law has a mass equal to 1):

$$e^{\Phi(\beta) - \langle \beta, U(\xi) \rangle} \text{ with } \Phi(\beta) = -\log(I_0) = -\log \int_M e^{-\langle \beta, U(\xi) \rangle} d\lambda \text{ and } Q(\beta) = \int_M e^{\Phi(\beta) - \langle \beta, U(\xi) \rangle} U d\lambda = \frac{I_1}{I_0}$$

> The set of these probabilities law is **Gibbs Ensemble of the Dynamic Group**, Φ is the **Thermodynamic Potential** and Q is the **Geometric Heat** $Q \in \mathfrak{g}^*$

Geometric Fisher Metric: Geometric Heat Capacity

■ We can observe that the Geometric Heat Q is C^∞ function of Geometric Temperature β in Dual Lie Algebra \mathfrak{g}^* :

$$\beta \in \mathfrak{g} \mapsto Q \in \mathfrak{g}^*$$

$$Q(\beta) = \int_M e^{\Phi(\beta) - \langle \beta, U(\xi) \rangle} U d\lambda = \frac{I_1}{I_0}$$

■ We have: $Q = \frac{\partial \Phi}{\partial \beta}$

$$\Phi(\beta) = -\log \int_M e^{-\langle \beta, U(\xi) \rangle} d\lambda$$

■ Its derivative is a 2nd order symmetric tensor: $\frac{\partial Q}{\partial \beta} = \frac{I_2}{I_0} - \frac{I_1 \otimes I_1}{I_0} = \frac{I_2}{I_1} - Q \otimes Q$

$$-\frac{\partial Q}{\partial \beta} = \int_M e^{\Phi(\beta) - \langle \beta, U(\xi) \rangle} [U - Q] \otimes [U - Q] d\lambda$$

$$-\frac{\partial Q}{\partial \beta} \geq 0 \quad -\frac{\partial Q}{\partial \beta} = -\frac{\partial^2 \Phi}{\partial \beta^2}$$

■ This quadratic form is positive, and positive definite for each $x \in M$ unless there exist a non null element $Z \in \mathfrak{g}$ such that $\langle U - Q, Z \rangle = 0$ (means that the moment U varies in an affine sub-manifold of \mathfrak{g}^*)

Distribution of probability by Group action

The distribution density under the action of the Lie Group is given by:

$$\mu^* : e^{\Phi^* - \langle \beta^*, U \rangle}$$

$$\Phi^* = \Phi(\beta^*) = \Phi - \langle \theta(g^{-1}), \beta \rangle$$

$$\Phi^* = \Phi + \langle \theta(g), Ad_g \beta \rangle$$

(**)

$$\beta^* = Ad_g(\beta)$$

$$\theta(g^{-1}) = -Ad_g^* \theta(g)$$

$$\Phi(\beta) = -\log \int_M e^{-\langle \beta, U(\xi) \rangle} d\lambda$$

The set Ω of Geometric Temperature is invariant by the adjoint action of G

$$\Psi_g(\mu_\beta) = \mu_{Ad_g(\beta)}$$

If we use $Q = \frac{\partial \Phi}{\partial \beta}$, we have the constraint $\delta \Phi - \langle Q, \delta \beta \rangle = 0$

By derivation of (**), we have: $\tilde{\Theta}(\beta, Z) + \langle Q, [\beta, Z] \rangle = 0$

$$\tilde{\Theta}(X, Y) : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{R}$$

$$X, Y \mapsto \langle \Theta(X), Y \rangle$$

$$\Theta(X) = T_e \theta(X(e))$$

Geometric (Planck) Temperature

■ We have previously observed that:



■ $\tilde{\Theta}(X, Y)$ is called the **Symplectic Cocycle of Lie algebra** \mathfrak{g} associated to the momentum map J

$$\tilde{\Theta}(X, Y) = J_{[X, Y]} - \{J_X, J_Y\} \text{ with } \{..,..\} \text{ Poisson Bracket and } J \text{ the Moment Map } \mathfrak{g} \rightarrow C^\infty(M, R)$$

► where J_X linear application from \mathfrak{g} to differential function on M :
and the associated differentiable application J , called moment(um) map:

$$X \rightarrow J_X$$

$$J : M \rightarrow \mathfrak{g}^* \text{ with } x \mapsto J(x) \text{ such that } J_X(x) = \langle J(x), X \rangle, X \in \mathfrak{g}$$

■ $\tilde{\Theta}(X, Y)$ is a 2-form of \mathfrak{g} and verify:

$$\tilde{\Theta}([X, Y], Z) + \tilde{\Theta}([Y, Z], X) + \tilde{\Theta}([Z, X], Y) = 0$$

■ If we define: $\tilde{\Theta}_\beta(Z_1, Z_2) = \tilde{\Theta}(Z_1, Z_2) + \langle Q, ad_{Z_1}(Z_2) \rangle$ with $ad_{Z_1}(Z_2) = [Z_1, Z_2]$

■ We can observe that : $\beta \in Ker \tilde{\Theta}_\beta$ $\tilde{\Theta}_\beta(\beta, \beta) = 0$, $\forall \beta \in \mathfrak{g}$

Associated Riemannian Metric: Geometric Fisher Metric

We can compute the image of Geometric Heat by the Lie Group action:

$$Q^* = Ad_g^*(Q) + \theta(g)$$

By tangential derivative to the orbit with respect to $Z \in \mathfrak{g}$ and by using positivity of $-\frac{\partial Q}{\partial \beta} \geq 0$, we find:

$$\tilde{\Theta}_\beta(Z, [\beta, Z]) = \tilde{\Theta}(Z, [\beta, Z]) + \langle Q, [Z, [\beta, Z]] \rangle \geq 0$$

$\tilde{\Theta}_\beta$ is a 2-form of \mathfrak{g} that verifies:

$$\tilde{\Theta}([X, Y], Z) + \tilde{\Theta}([Y, Z], X) + \tilde{\Theta}([Z, X], Y) = 0$$

Then, there exists a **symmetric tensor** g_β defined on $ad_\beta(Z)$

$$g_\beta([\beta, Z_1], [\beta, Z_2]) = \tilde{\Theta}_\beta(Z_1, [\beta, Z_2])$$

With the following invariances:

$$s[Q(Ad_g(\beta))] = s(Q(\beta))$$

$$I(Ad_g(\beta)) = -\frac{\partial^2(\Phi - \langle \theta(g^{-1}), \beta \rangle)}{\partial \beta^2} = -\frac{\partial^2 \Phi}{\partial \beta^2} = I(\beta)$$

Fisher Metric of Souriau Lie Group Thermodynamics

Souriau has introduced the Riemannian metric

$$g_\beta([\beta, Z_1], [\beta, Z_2]) = \tilde{\Theta}_\beta(Z_1, [\beta, Z_2]) \quad \beta \in \text{Ker } \tilde{\Theta}_\beta$$

$$\tilde{\Theta}_\beta(Z_1, Z_2) = \tilde{\Theta}(Z_1, Z_2) + \langle Q, \text{ad}_{Z_1}(Z_2) \rangle \quad \text{with } \text{ad}_{Z_1}(Z_2) = [Z_1, Z_2]$$

This metric is an **extension of Fisher metric, an hessian metric**: If we differentiate the relation $Q(\text{Ad}_g(\beta)) = \text{Ad}_g^*(Q) + \theta(g)$

$$\frac{\partial Q}{\partial \beta}(-[Z_1, \beta], \cdot) = \tilde{\Theta}(Z_1, [\beta, \cdot]) + \langle Q, \text{Ad}_{Z_1}([\beta, \cdot]) \rangle = \tilde{\Theta}_\beta(Z_1, [\beta, \cdot])$$

$$-\frac{\partial Q}{\partial \beta}([Z_1, \beta], Z_2) = \tilde{\Theta}(Z_1, [\beta, Z_2]) + \langle Q, \text{Ad}_{Z_1}([\beta, Z_2]) \rangle = \tilde{\Theta}_\beta(Z_1, [\beta, Z_2])$$

$$\Rightarrow -\frac{\partial^2 \Phi}{\partial \beta^2} = -\frac{\partial Q}{\partial \beta} = g_\beta([\beta, Z_1], [\beta, Z_2]) = \tilde{\Theta}_\beta(Z_1, [\beta, Z_2])$$

The Fisher Metric is then a **generalization of "Heat Capacity"**:

$$\beta = \frac{1}{kT} \quad K = -\frac{\partial Q}{\partial \beta} = -\frac{\partial Q}{\partial T} \left(\frac{\partial(1/kT)}{\partial T} \right)^{-1} = kT^2 \frac{\partial Q}{\partial T} \quad \frac{\partial T}{\partial t} = \frac{\kappa}{C.D} \Delta T \quad \text{with } \frac{\partial Q}{\partial T} = C.D$$

Link with Classical Thermodynamics

■ We have the reciprocal formula:

$$Q = \frac{\partial \Phi}{\partial \beta}$$

$$\beta = \frac{\partial s}{\partial Q}$$

$$s(Q) = \left\langle \frac{\partial \Phi}{\partial \beta}, \beta \right\rangle - \Phi$$

$$\Phi(\beta) = \left\langle Q, \frac{\partial s}{\partial Q} \right\rangle - s$$

■ For Classical Thermodynamics (Time translation only), we recover the definition of Boltzmann-Clausius Entropy:

$$\begin{cases} \beta = \frac{\partial s}{\partial Q} \\ \beta = \frac{1}{T} \end{cases} \Rightarrow ds = \frac{dQ}{T}$$

Continuous Medium Thermodynamics

For Continuous Medium Thermodynamics, « Temperature Vector » is no longer constrained to be in Lie Algebra, but only constrained by phenomenologic equations (e.g. Navier equations, ...).

For Thermodynamic equilibrium, the « Temperature Vector » is a Killing vector of Space-Time.

For each point X , there is a « Temperature Vector » $\beta(X)$, such it is an infinitesimal conformal transform of the metric of the univers g_{ij} :

$$\begin{aligned} \hat{\partial}_i \beta_j + \hat{\partial}_j \beta_i &= \lambda g_{ij} \\ \partial_i \beta_j + \partial_j \beta_i - 2\Gamma_{ij}^k \beta_k &= \lambda g_{ij} \end{aligned} \quad \left\{ \begin{array}{l} \hat{\partial}_i \cdot : \text{covariant derivative} \\ \beta_j : \text{component of Temperature vector} \end{array} \right.$$

$\lambda = 0 \Rightarrow$ Killing Equation

Conservation equations can be deduced for components of Impulsion-Energy tensor T^{ij} and Entropy flux S^j : $\hat{\partial}_i T^{ij} = 0$ $\partial_i S^j = 0$

Poincaré-Cartan Integral Invariant of Lie Group Thermodynamics

Analogies between Geometric Mechanics & Geometric Lie Group Thermodynamics, provides the following similarities of structures:

$$\left\{ \begin{array}{l} \dot{q} \leftrightarrow \beta \\ p \leftrightarrow Q \end{array} \right. \quad \left\{ \begin{array}{l} L(\dot{q}) \leftrightarrow \Phi(\beta) \\ H(p) \leftrightarrow s(Q) \\ H = p.\dot{q} - L \leftrightarrow s = \langle Q, \beta \rangle - \Phi \end{array} \right. \quad \left\{ \begin{array}{l} \dot{q} = \frac{dq}{dt} = \frac{\partial H}{\partial p} \leftrightarrow \beta = \frac{\partial s}{\partial Q} \\ p = \frac{\partial L}{\partial \dot{q}} \leftrightarrow Q = \frac{\partial \Phi}{\partial \beta} \end{array} \right.$$

We can then consider a similar Poincaré-Cartan-Souriau Pfaffian form:

$$\omega = p.dq - H.dt \leftrightarrow \omega = \langle Q, (\beta.dt) \rangle - s.dt = (\langle Q, \beta \rangle - s).dt = \Phi(\beta).dt$$

This analogy provides an associated Poincaré-Cartan Integral Invariant:

$$\int_{C_a} p.dq - H.dt = \int_{C_b} p.dq - H.dt \text{ transforms in } \int_{C_a} \Phi(\beta).dt = \int_{C_b} \Phi(\beta).dt$$

For Thermodynamics, we can then deduce an Euler-Poincaré Variational Principle: The Variational Principle holds on \mathfrak{g} , for variations $\delta\beta = \dot{\eta} + [\beta, \eta]$, where $\eta(t)$ is an arbitrary path that vanishes at the endpoints, $\eta(a) = \eta(b) = 0$:

$$\delta \int_{t_0}^{t_1} \Phi(\beta(t)).dt = 0$$

Souriau Gibbs states for Hamiltonian actions of subgroups of the Galilean group

➤ Galilean Transformation on position and speed:

$$\begin{pmatrix} \vec{r}' & \vec{v}' \\ t' & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} A & \vec{b} & \vec{d} \\ 0 & 1 & e \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \vec{r} & \vec{v} \\ t & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} A\vec{r} + t\vec{b} + \vec{d} & A\vec{v} + \vec{b} \\ t + e & 1 \\ 1 & 0 \end{pmatrix}$$

➤ **Souriau Result:** this action is Hamiltonian, with the map J , defined on the evolution space of the particle, with value in the dual \mathfrak{g}^* of the Lie algebra \mathfrak{G} , as momentum map

$$J(\vec{r}, t, \vec{v}, m) = m \begin{pmatrix} \vec{r} \times \vec{v} & 0 & 0 \\ \vec{r} - t\vec{v} & 0 & 0 \\ \vec{v} & \frac{1}{2}\|\vec{v}\|^2 & 0 \end{pmatrix} = m \left\{ \vec{r} \times \vec{v}, \vec{r} - t\vec{v}, \vec{v}, \frac{1}{2}\|\vec{v}\|^2 \right\} \in \mathfrak{g}^*$$

➤ Coupling formula:

$$\langle J(\vec{r}, t, \vec{v}, m), \beta \rangle = \left\langle m \left\{ \vec{r} \times \vec{v}, \vec{r} - t\vec{v}, \vec{v}, \frac{1}{2}\|\vec{v}\|^2 \right\}, \{ \vec{\omega}, \vec{\alpha}, \vec{\delta}, \varepsilon \} \right\rangle$$

$$Z = \begin{pmatrix} j(\vec{\omega}) & \vec{\alpha} & \vec{\delta} \\ 0 & 1 & \varepsilon \\ 0 & 0 & 0 \end{pmatrix} = \{ \vec{\omega}, \vec{\alpha}, \vec{\delta}, \varepsilon \} \in \mathfrak{g}$$

$$\langle J(\vec{r}, t, \vec{v}, m), \beta \rangle = m \left(\vec{\omega} \cdot \vec{r} \times \vec{v} - (\vec{r} \times \vec{v}) \cdot \vec{\alpha} + \vec{v} \cdot \vec{\delta} - \frac{1}{2}\|\vec{v}\|^2 \varepsilon \right)$$

Souriau Gibbs states for Hamiltonian actions of subgroups of the Galilean group

Souriau Gibbs states for one-parameter subgroups of the Galilean group

- **Souriau Result:** Action of the full Galilean group on the space of motions of an isolated mechanical system is not related to any Equilibrium Gibbs state (the open subset of the Lie algebra, associated to this Gibbs state, is empty)
- The **1-parameter subgroup of the Galilean group** generated by β element of Lie Algebra, is the set of matrices

$$\exp(\tau\beta) = \begin{pmatrix} A(\tau) & \vec{b}(\tau) & \vec{d}(\tau) \\ 0 & 1 & \tau\varepsilon \\ 0 & 0 & 1 \end{pmatrix} \text{ with } \begin{cases} A(\tau) = \exp(\tau j(\vec{\omega})) \text{ and } \vec{b}(\tau) = \left(\sum_{i=1}^{\infty} \frac{\tau^i}{i!} (j(\vec{\omega}))^{i-1} \right) \vec{\alpha} \\ \vec{d}(\tau) = \left(\sum_{i=1}^{\infty} \frac{\tau^i}{i!} (j(\vec{\omega}))^{i-1} \right) \vec{\delta} + \varepsilon \left(\sum_{i=2}^{\infty} \frac{\tau^i}{i!} (j(\vec{\omega}))^{i-2} \right) \vec{\alpha} \end{cases}$$
$$\beta = \begin{pmatrix} j(\vec{\omega}) & \vec{\alpha} & \vec{\delta} \\ 0 & 1 & \varepsilon \\ 0 & 0 & 0 \end{pmatrix} \in \mathfrak{g}$$

Souriau Thermodynamics of butter churn (device used to convert cream into butter) or “La Thermodynamique de la crème”

If we consider the case of the centrifuge

$$\vec{\omega} = \omega \vec{e}_z, \vec{\alpha} = 0 \text{ and } \vec{\delta} = 0$$

$$\text{Rotation speed: } \frac{\omega}{\varepsilon}$$

$$f_i(\vec{r}_{i0}) = -\frac{\omega^2}{2\varepsilon^2} \|\vec{e}_z \times \vec{r}_{i0}\|^2$$

with $\Delta = \|\vec{e}_z \times \vec{r}_{i0}\|$ distance to axis z

$$\rho_i(\beta) = \frac{1}{P_i(\beta)} \exp(-\langle J_i, \beta \rangle) = \text{cst.} \exp\left(-\frac{1}{2m_i kT} \|\vec{p}_{i0}\|^2 + \frac{m_i}{2kT} \left(\frac{\omega}{\varepsilon}\right)^2 \Delta^2\right)$$

- the behaviour of a gas made of point particles of various masses in a centrifuge rotating at a constant angular velocity (the heavier particles concentrate farther from the rotation axis than the lighter ones)



$$\frac{\omega}{\varepsilon}$$

Lie Group Thermodynamics: Centrifuge for Butter, U235 & Ribo acides

- Physiquement, la théorie donne de bons résultats si on l'applique aux divers sous-groupes du groupe de Galilée qui sont caractéristiques des appareils thermodynamiques : ainsi une boîte cylindrique dans laquelle on enferme un fluide lui laisse un sous-groupe d'invariance de dimension 2 : rotations autour de l'axe, translations temporelles. D'où résulte un vecteur température à deux dimensions, que l'on peut "transmettre" au fluide par l'intermédiaire de la boîte, (en la refroidissant, par exemple, et en la faisant tourner) ; les résultats de la théorie sont ceux-là même que l'on exploite dans les centrifugeuses (par exemple pour fabriquer du beurre, de l'uranium 235 ou des acides ribonucléiques).

- On remarquera que le processus par lequel une centrifugeuse réfrigérée transmet son propre vecteur-température à son contenu porte deux noms différents : conduction thermique et viscosité, selon la composante du vecteur-température que l'on considère ; conduction et viscosité devraient donc être unifiées dans une théorie fondamentale des processus irréversibles (théorie qui reste à construire).



Balian Computation of Gibbs Density for Dynamical Centrifuge System



Roger Balian Computation of Gibbs density for centrifuge

- Balian has computed the Boltzmann-Gibbs distribution without knowing Souriau equations. Exercice 7b of :

- **Balian, R. From Microphysics to Macrophysics, 2nd ed.; Springer: Berlin, Germany, 2007; Volume I**

- Balian started by considering the constants of motion that are the energy and the component J_z of the total angular momentum:

$$J = \sum_i (r_i \times p_i)$$

- Balian observed that he must add to the Lagrangian parameter, given by (Planck) temperature β for energy, an additional one associated with J_z . He identifies this additional multiplier with $-\beta\omega$ by evaluating the mean velocity at each point.
- He then introduced the same results also by changing the frame of reference, the Lagrangian and the Hamiltonian in the rotating frame and by writing down the canonical equilibrium in that frame. He uses the resulting distribution to find, through integration, over the momenta, an expression for the particles density as the function of the distance from the cylinder axis.

Roger Balian Computation of Gibbs density for centrifuge

- The fluid carried along by the walls of the rotating vessel acquires a non-vanishing average angular momentum $\langle J_z \rangle$ around the axis of rotation, that is a constant of motion. In order to be able to assign to it a definite value, Balian proposed to associate with it a Lagrangian multiplier λ , in exactly the same way as we classically associate the multiplier β with the energy in canonical equilibrium. The average $\langle J_z \rangle$ will be a function of λ . The Gibbs density for rotating gas is given by Balian as:

$$D = \frac{1}{Z} e^{-\beta H - \lambda J_z} = \frac{1}{Z} \exp \left\{ \sum_i \left[\frac{\beta p_i^2}{2m} + \lambda (x_i p_{y_i} - y_i p_{x_i}) \right] \right\}$$

- With the energy and the average angular momentum given by:

$$U = -\frac{\partial \ln Z}{\partial \beta} = \frac{1}{kT} \quad \langle J_z \rangle = -\frac{\partial \ln Z}{\partial \lambda}$$

Roger Balian Computation of Gibbs density for centrifuge

- The Lagrangian parameter λ has a mechanical nature. To identify this parameter, Balian compared microscopic and macroscopy descriptions of fluid mechanics. He described the single-particle reduced density by:

$$f(r, p) \propto \exp\left\{-\frac{\beta p^2}{2m} - \lambda(xp_y - yp_x)\right\} = \exp\left\{-\frac{\beta}{2m}\left(p + \frac{m}{\beta}[\lambda \times r]\right)^2 + \frac{m\lambda^2}{2\beta}(x^2 + y^2)\right\}$$

- Whence Balian finds the velocity distribution at a point r to be proportional to:

$$\exp\left\{-\frac{m}{2kT}\left(v + \frac{1}{\beta}[\lambda \times r]\right)^2\right\}$$

- The mean velocity of the fluid at the point r is equal to: $\langle v \rangle = -\frac{1}{\beta}[\lambda \times r]$

- and can be identified with the velocity $[\omega \times r]$ in an uniform rotation with angular velocity ω . By comparison, Balian put : $\omega = -\frac{\lambda}{\beta}$

Roger Balian Computation of Gibbs density for centrifuge

Balian made the remarks that *“The angular momentum is imparted to the gas when the molecules collide with the rotating walls, which changes the Maxwell distribution at every point, shifting its origin. The walls play the role of an angular momentum reservoir. Their motion is characterized by a certain angular velocity, and the angular velocities of the fluid and of the walls become equal at equilibrium, exactly like the equalization of the temperature through energy exchanges”*.

Compliance with Symplectic Model of Thermodynamics by Balian-Valentin



Compatible Balian Gauge Theory of Thermodynamics

- Entropy S is an extensive variable $q^0 = S(q^1, \dots, q^n)$ depending on q^i ($i = 1, \dots, n$) n independent extensive/conservative quantities characterizing the system
- The n intensive variables γ_i are defined as the partial derivatives:
$$\gamma_i = \frac{\partial S(q^1, \dots, q^n)}{\partial q^i}$$
- Balian has introduced a non-vanishing gauge variable which multiplies all the intensive variables, defining a new set of variables: $p_i = -p_0 \cdot \gamma_i$, $i = 1, \dots, n$
- The $2n+1$ -dimensional space is thereby extended into a $2n+2$ -dimensional thermodynamic space T spanned by the variables p_i, q^i with $i = 0, 1, \dots, n$, where the physical system is associated with a $n+1$ -dimensional manifold M in T , parameterized for instance by the coordinates q^1, \dots, q^n and p_0 .

Compatible Balian Gauge Theory of Thermodynamics

the contact structure in $2n+1$ dimension: $\tilde{\omega} = dq^0 - \sum_{i=1}^n \gamma_i \cdot dq^i$

is embedded into a symplectic structure in $2n+2$ dimension, with 1-form, as symplectization: $\omega = \sum_{i=0}^n p_i \cdot dq^i$

The $n+1$ -dimensional thermodynamic manifolds M are characterized by $\omega = 0$. The 1-form induces then a symplectic structure on T : $d\omega = \sum_{i=0}^n dp_i \wedge dq^i$

The concavity of the entropy $S(q^1, \dots, q^n)$, as function of the extensive variables, expresses the stability of equilibrium states. It entails the existence of a metric structure in the n -dimensional space q_i :

$$ds^2 = -d^2S = -\sum_{i,j=1}^n \frac{\partial^2 S}{\partial q^i \partial q^j} dq^i dq^j$$

which defines a distance between two neighboring thermodynamic states:

$$d\gamma_i = \sum_{j=1}^n \frac{\partial^2 S}{\partial q^i \partial q^j} dq^j$$

$$ds^2 = -\sum_{i=1}^n d\gamma_i dq_i = \frac{1}{P_0} \sum_{i=0}^n dp_i dq^i$$

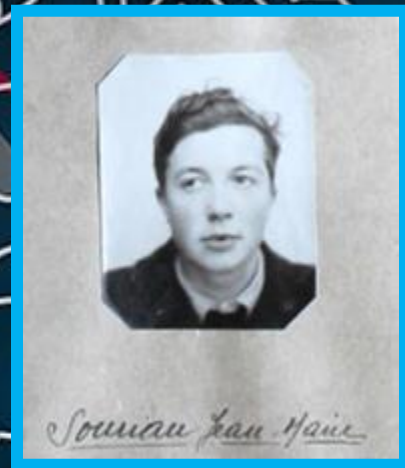
Compatible Balian Gauge Theory of Thermodynamics

- We can observe that this Gauge Theory of Thermodynamics is compatible with Souriau Lie Group Thermodynamics, where we have to consider the Souriau vector :

$$\beta = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix} \quad \text{transformed in a new vector} \quad p_i = -p_0 \cdot \gamma_i$$

$$p = \begin{bmatrix} -p_0 \gamma_1 \\ \vdots \\ -p_0 \gamma_n \end{bmatrix} = -p_0 \cdot \beta$$

Souriau Entropy Definition



Entropy Definition by Jean-Marie Souriau (1/4)

Let E a vector space of finite size, μ a measure on the dual space E^* , then the function given by:

$$\alpha \mapsto \int_{E^*} e^{M\alpha} \mu(M) dM$$

for all $\alpha \in E$ such that the integral is convergent.

This function is called **Laplace Transform**. This transform F of the measure μ is differentiable inside its definition set $def(F)$. Its p -th derivatives are given by the following convergent integrals :

$$F^{(p)}(\alpha) = \int_{E^*} M \otimes M \dots \otimes M \mu(M) dM$$

Souriau Theorem:

- Let E a vector space of finite size, μ a non zero positive measure of its dual space E^* , F its Laplace transform, then:
- F is a semi-definite convex function, $F(\alpha) > 0, \forall \alpha \in \text{def}(F)$
 - $f = \log(F)$ is convex and semi-continuous
 - Let α an interior point of $\text{def}(F)$ then:
 - $D^2(f)(\alpha) \geq 0$
 - $D^2(f)(\alpha) = \int_{E^*} e^{M\alpha} [M - D(f)(\alpha)]^{\otimes 2} \mu(M) dM$
 - $D^2(f)(\alpha)$ inversible \Leftrightarrow affine Enveloppe (support(μ)) = E^*

Entropy Definition by Jean-Marie Souriau (3/4)

Lemme:

- Let X a locally compact space, Let λ a positive measure of X , with X as support, then the following function Φ is convex:

$$\Phi(h) = \log \int_X e^{h(x)} \lambda(x) dx, \quad \forall h \in C(X)$$

such that the integral is convergent.

Proof:

- The integral is strictly positive when its converges, insuring existence of its logarithm
- Epigraph Φ is the set of $\begin{pmatrix} h \\ y \end{pmatrix}$ such that $\int_X e^{h(x)-y} \lambda(x) dx \leq 1$.
- Convexity of exponential prove that this epigraph is convex.

Entropy definition by Jean-Marie Souriau:

➤ (Neg)entropy is given by Legendre transform of:

$$\Phi(h) = \log \int_X e^{h(x)} \lambda(x) dx, \quad \forall h \in C(X)$$

➤ We call “Boltzmann Law” (relatively to λ) all measures μ of X such that the set of real values $\mu(h) - \Phi(h)$, $h \in \text{def}(\Phi)$ and h est μ -integrable.

Example of Multivariate Gaussian Law (real case)

Multivariate Gaussian law parameterized by moments

$$p_{\hat{\xi}}(\xi) = \frac{1}{(2\pi)^{n/2} \det(R)^{1/2}} e^{-\frac{1}{2}(z-m)^T R^{-1}(z-m)}$$

$$\begin{aligned} \frac{1}{2}(z-m)^T R^{-1}(z-m) &= \frac{1}{2} \left[z^T R^{-1} z - m^T R^{-1} z - z^T R^{-1} m + m^T R^{-1} m \right] \\ &= \frac{1}{2} z^T R^{-1} z - m^T R^{-1} z + \frac{1}{2} m^T R^{-1} m \end{aligned}$$

$$p_{\hat{\xi}}(\xi) = \frac{1}{(2\pi)^{n/2} \det(R)^{1/2} e^{\frac{1}{2} m^T R^{-1} m}} e^{-\left[-m^T R^{-1} z + \frac{1}{2} z^T R^{-1} z \right]} = \frac{1}{Z} e^{-\langle \xi, \beta \rangle}$$

$$\xi = \begin{bmatrix} z \\ z z^T \end{bmatrix} \text{ and } \beta = \begin{bmatrix} -R^{-1} m \\ \frac{1}{2} R^{-1} \end{bmatrix} = \begin{bmatrix} a \\ H \end{bmatrix} \text{ with } \langle \xi, \beta \rangle = a^T z + z^T H z = \text{Tr} \left[z a^T + H^T z z^T \right]$$

**Gaussian
Density is
a 1st order
Maximum
Entropy Density !**

Géométrie de l'Information pour une gaussienne multivariée

Pour la densité Gaussienne multivariée de moyenne m et de matrice de covariance R , nous avons pris l'habitude de la paramétrer par:

$$p_{m,R}(z) = \frac{1}{(2\pi)^{n/2} \det(R)^{1/2}} e^{-\frac{1}{2}(z-m)^T R^{-1}(z-m)}$$

Mais la bonne paramétrisation est la suivante:

$$p_{\xi}(z) = \frac{1}{\int_{\Omega^*} e^{-\langle \xi, \beta \rangle} . d\xi} e^{-\langle \xi, \beta \rangle} \quad \text{avec} \quad \xi = \begin{bmatrix} z \\ zz^T \end{bmatrix} \quad \text{et} \quad \beta = \begin{bmatrix} -R^{-1}m \\ \frac{1}{2}R^{-1} \end{bmatrix} = \begin{bmatrix} a \\ H \end{bmatrix}$$

$$\text{Dualité donnée par} \quad \langle \xi, \beta \rangle = a^T z + z^T H z = \text{Tr} \left[z a^T + H^T z z^T \right]$$

$$\log \left(\int_{\Omega^*} e^{-\langle \xi, \beta \rangle} . d\xi \right) = \log(Z) = n \log(2\pi) + \frac{1}{2} \log \det(R) + \frac{1}{2} m^T R^{-1} m$$

Fonction caractéristique de Massieu:

$$\psi_{\Omega}(\beta) = \int_{\Omega^*} e^{-\langle \xi, \beta \rangle} \cdot d\xi \quad \text{et} \quad \beta = \begin{bmatrix} -R^{-1}m \\ \frac{1}{2}R^{-1} \end{bmatrix} = \begin{bmatrix} a \\ H \end{bmatrix}$$

$$\Phi(\beta) = -\log \psi_{\Omega}(\beta) = \frac{1}{2} \left[-\text{Tr}[H^{-1}aa^T] + \log[(2)^n \det H] - n \log(2\pi) \right]$$

Relation de la dérivée donnant le moment

$$\frac{\partial \Phi(\beta)}{\partial \beta} = \int_{\Omega^*} \xi \cdot p_{\bar{\xi}}(\xi) d\xi = \bar{\xi} = \begin{bmatrix} E[z] \\ E[zz^T] \end{bmatrix} = \begin{bmatrix} m \\ R + mm^T \end{bmatrix}$$

avec $\xi = \begin{bmatrix} z \\ zz^T \end{bmatrix}$ et $R = E[(z - m)(z - m)^T] = E[zz^T] - mm^T$

Entropie (de Shannon), transformée de Legendre de la fonction

$$S(\bar{\xi}) = \langle \bar{\xi}, \beta \rangle - \Phi(\beta) \quad \text{avec} \quad \frac{\partial S(\bar{\xi})}{\partial \bar{\xi}} = \beta = \begin{bmatrix} a \\ H \end{bmatrix} = \begin{bmatrix} -R^{-1}m \\ \frac{1}{2}R^{-1} \end{bmatrix}$$

$$S(\bar{\xi}) = - \int_{\Omega^*} \frac{e^{-\langle \xi, \beta \rangle}}{\int_{\Omega^*} e^{-\langle \xi, \beta \rangle} .d\xi} \log \frac{e^{-\langle \xi, \beta \rangle}}{\int_{\Omega^*} e^{-\langle \xi, \beta \rangle} .d\xi} .d\xi = - \int_{\Omega^*} p_{\bar{\xi}}(\xi) \log p_{\bar{\xi}}(\xi) .d\xi$$

Expression de l'Entropie (de Shannon):

$$S(\bar{\xi}) = \frac{1}{2} [\log(2)^n \det[H^{-1}] + n \log(2\pi.e)] = \frac{1}{2} [\log \det[R] + n \log(2\pi.e)]$$

Approche Cohomologique de Koszul et Souriau

Cas des densités multivariées gaussiennes

$$\hat{\xi} = \begin{bmatrix} E[z] \\ E[zz^T] \end{bmatrix} = \begin{bmatrix} m \\ R + mm^T \end{bmatrix}, \beta = \begin{bmatrix} -R^{-1}m \\ \frac{1}{2}R^{-1} \end{bmatrix} \quad Ad_M^* \hat{\xi} = \begin{bmatrix} R + mm^T & -mm^T & R^{1/2}m \\ 0 & 0 & 0 \end{bmatrix}$$

Sont homéomorphes à:

$$\hat{\xi} = \begin{bmatrix} R + mm^T & m \\ 0 & 0 \end{bmatrix} \in \mathfrak{g}, \beta = \begin{bmatrix} \frac{1}{2}R^{-1} & -R^{-1}m \\ 0 & 0 \end{bmatrix} \quad \hat{\xi}(Ad_M(\beta)) = Ad_M^*(\hat{\xi}) + \theta(M)$$

$$Ad_M \beta = M \cdot \beta \cdot M^{-1} = \begin{bmatrix} R^{1/2} & m' \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2}R^{-1} & -R^{-1}m \\ 0 & 0 \end{bmatrix} \begin{bmatrix} R^{1/2} & -R^{1/2}m' \\ 0 & 1 \end{bmatrix}$$

$$Ad_M \beta = \begin{bmatrix} \frac{1}{2}R^{1/2}R^{-1}R^{1/2} & -\frac{1}{2}R^{1/2}R^{-1}R^{1/2}m' - R^{1/2}R^{-1}m \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\Omega^{-1} & -\Omega^{-1}n \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow \hat{\xi}(Ad_M(\beta)) = \begin{bmatrix} \Omega + nn^T & n \\ 0 & 0 \end{bmatrix} \quad \text{with } \Omega = R^{1/2}RR^{1/2} \quad \text{and } n = \left(\frac{1}{2}m' + R^{1/2}m \right)$$

Approche Cohomologique de Koszul et Souriau

Calcul du cocycle:

$$Ad_M \beta = \begin{bmatrix} \Omega^{-1} & -\Omega^{-1}n \\ 0 & 0 \end{bmatrix} \Rightarrow \hat{\xi}(Ad_M(\beta)) = \begin{bmatrix} \Omega + nn^T & n \\ 0 & 0 \end{bmatrix} \quad Ad_M^* \hat{\xi} = \begin{bmatrix} R + mm^T - mm'^T & R^{1/2}m \\ 0 & 0 \end{bmatrix}$$

with $\Omega = 2R^{1/2}RR^{1/2}$ and $n = (m' + 2R^{1/2}m)$

$$\begin{aligned} \hat{\xi}(Ad_M(\beta)) &= Ad_M^*(\hat{\xi}) + \theta(M) \Rightarrow \theta(M) = \hat{\xi}(Ad_M(\beta)) - Ad_M^*\hat{\xi} \\ &= \begin{bmatrix} R^{1/2}RR^{1/2} + \left(\frac{1}{2}m' + R^{1/2}m\right)\left(\frac{1}{2}m' + R^{1/2}m\right)^T & \left(\frac{1}{2}m' + R^{1/2}m\right) \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} R + mm^T - mm'^T & R^{1/2}m \\ 0 & 0 \end{bmatrix} \end{aligned}$$

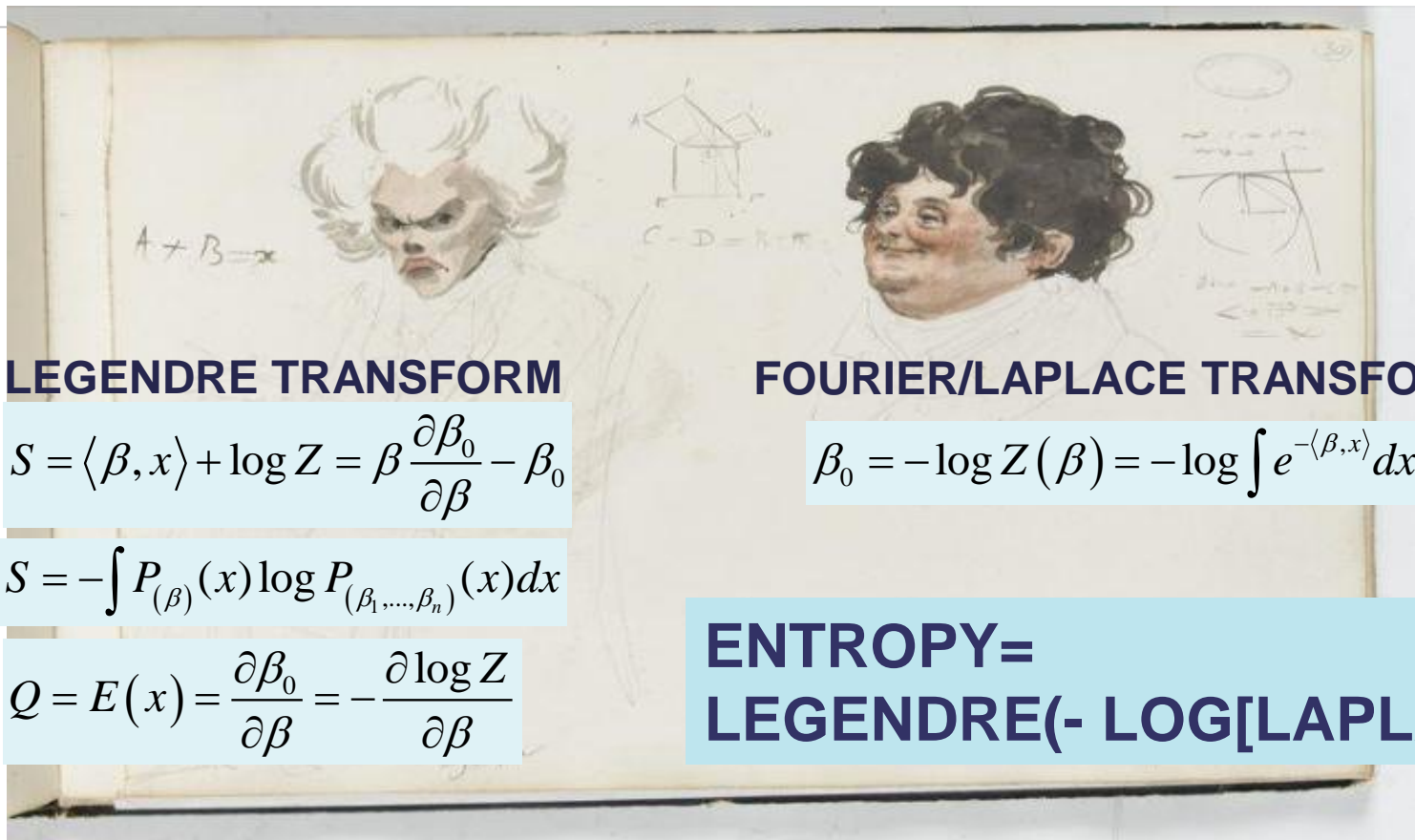
$$g_\beta([\beta, Z_1], [\beta, Z_2]) = \tilde{\Theta}_\beta(Z_1, [\beta, Z_2]) = \tilde{\Theta}(Z_1, [\beta, Z_2]) + \langle \hat{\xi}, ad_{Z_1}([\beta, Z_2]) \rangle$$

$$\tilde{\Theta}(M, Y) = \langle \Theta(M), Y \rangle = \langle T_e \theta(M), Y \rangle$$

Souriau Lie Groups Thermodynamics Extension: Higher Order Lie Groups Thermodynamics based on Günther Polysymplectic Model



Preservation of Legendre Structure by Souriau



LEGENDRE TRANSFORM

$$S = \langle \beta, x \rangle + \log Z = \beta \frac{\partial \beta_0}{\partial \beta} - \beta_0$$

$$S = - \int P_{(\beta)}(x) \log P_{(\beta_1, \dots, \beta_n)}(x) dx$$

$$Q = E(x) = \frac{\partial \beta_0}{\partial \beta} = - \frac{\partial \log Z}{\partial \beta}$$

FOURIER/LAPLACE TRANSFORM

$$\beta_0 = - \log Z(\beta) = - \log \int e^{-\langle \beta, x \rangle} dx$$

**ENTROPY =
LEGENDRE(- LOG[LAPLACE])**

High Order Temperature Model by R.S. Ingarden



High order thermodynamics

➤ High order moments:
$$Q_k = \frac{\partial \Phi(\beta_1, \dots, \beta_n)}{\partial \beta_k} = \frac{\int_M U^k(\xi) \cdot e^{-\sum_{k=1}^n \langle \beta_k, U^k(\xi) \rangle} d\omega}{\int_M e^{-\sum_{k=1}^n \langle \beta_k, U^k(\xi) \rangle} d\omega}$$

➤ High order Massieu characteristic function:
$$\Phi(\beta_1, \dots, \beta_n) = -\log \int_M e^{-\sum_{k=1}^n \langle \beta_k, U^k(\xi) \rangle} d\omega$$

➤ High order temperatures and capacities:
$$\beta_k = \frac{\partial S(Q_1, \dots, Q_n)}{\partial Q_k} \quad K_k = -\frac{\partial Q_k}{\partial \beta_k}$$

➤ Entropy:
$$S(Q_1, \dots, Q_n) = \sum_{k=1}^n \langle \beta_k, Q_k \rangle - \Phi(\beta_1, \dots, \beta_n)$$

➤ High order Gibbs density:
$$p_{Gibbs}(\xi) = e^{\sum_{k=1}^n \langle \beta_k, U^k(\xi) \rangle - \Phi(\beta_1, \dots, \beta_n)} = \frac{e^{-\sum_{k=1}^n \langle \beta_k, U^k(\xi) \rangle}}{\int_M e^{-\sum_{k=1}^n \langle \beta_k, U^k(\xi) \rangle} d\omega}$$

Günther model inspired by J.M. Souriau

Christian Günther work was inspired by the symplectic formulation of classical mechanics by Jean-Marie Souriau:

➤ J. M. Souriau, Structure des systemes dynamiques, Dunod, Paris, 1970

and by the work of Edelen and Rund on a local Hamiltonian formulation of field theory:

➤ D. G. B. Edelen, The invariance group for Hamiltonian systems of partial differential equations, Arch. Rational Mech. Anal. 5 (1961) 95-176.

➤ D. G. B. Edelen, Nonlocal variations and local invariance of fields, American Elsevier, New York,

➤ H. Rund, The Hamilton-Jacobi theory in the calculus of variations, Van Nostrand, Princeton, NJ, 1966

D. G. B. Edelen work is a coordinate version of the local polysymplectic approach of Günther.

Poly-symplectic extension of Souriau Lie Groups Thermodynamics

- This extension defines an action of G over $\mathfrak{g}^* \times \dots \times \mathfrak{g}^*$ ⁽ⁿ⁾ called n-coadjoint action:

$$Ad_g^{*(n)} : G \times \left(\mathfrak{g}^* \times \dots \times \mathfrak{g}^* \right) \rightarrow \mathfrak{g}^* \times \dots \times \mathfrak{g}^*$$

$$g \times \mu_1 \times \dots \times \mu_n \mapsto Ad_g^{*(n)}(\mu_1, \dots, \mu_n) = (Ad_g^* \mu_1, \dots, Ad_g^* \mu_n)$$

- Let $\mu = (\mu_1, \dots, \mu_n)$ a poly-momentum, element of $\mathfrak{g}^* \times \dots \times \mathfrak{g}^*$ ⁽ⁿ⁾, we can define a n-coadjoint orbit $O_\mu = O_{(\mu_1, \dots, \mu_n)}$ at the point μ , for which the canonical projection:

$$Pr_k : \mathfrak{g}^* \times \dots \times \mathfrak{g}^* \rightarrow \mathfrak{g}^* \quad , \quad (\nu_1, \dots, \nu_n) \mapsto \nu_k$$

induces a smooth map between the n-coadjoint orbit O_μ and the coadjoint orbit O_{μ_k} :

$$\pi_k : O_\mu = O_{(\mu_1, \dots, \mu_n)} \rightarrow O_{\mu_k}$$

that is a surjective submersion with $\bigcap_{k=1}^n Ker T\pi_k = \{0\}$.

Poly-symplectic extension of Souriau Lie Groups Thermodynamics

- Extending Souriau approach, equivariance of poly-moment could be studied to prove that there is a unique action $a(\dots)$ of the Lie group G on $\mathfrak{g}^* \times \dots \times \mathfrak{g}^*$ for which the polymoment map with $x \in M$ and $g \in G : J^{(n)} = (J^1, \dots, J^n) : M \rightarrow \mathfrak{g}^* \times \dots \times \mathfrak{g}^*$ that verifies: $J^{(n)}(\Phi_g(x)) = a(g, J^{(n)}(x)) = Ad_g^{*(n)}(J^{(n)}(x)) + \theta^{(n)}(g)$

with $Ad_g^{*(n)}(J^{(n)}(x)) = (Ad_g^* J^1, \dots, Ad_g^* J^n)$ and $\theta^{(n)}(g) = (\theta^1(g), \dots, \theta^n(g))$

- $\theta^{(n)}(g)$ is a **poly-symplectic one-cocycle**
- We can also defined **poly-symplectic two-cocycle**:

$$\tilde{\Theta}^{(n)} = (\tilde{\Theta}^1, \dots, \tilde{\Theta}^n) \text{ with } \tilde{\Theta}^k(X, Y) = \langle \Theta^k(X), Y \rangle = J_{[X, Y]}^k - \{J_X^k, J_Y^k\}$$

where $\Theta^k(X) = T_e \theta^k(X(e))$

- the **poly-symplectic Souriau-Fisher metric** is given by:

$$g_\beta([[\beta, Z_1], Z_2]) = \text{diag} [\tilde{\Theta}_{\beta_k}(Z_1, Z_2)]_k, \quad \forall Z_1 \in \mathfrak{g}, \forall Z_2 \in \text{Im}(ad_\beta(\cdot)), \beta = (\beta_1, \dots, \beta_n)$$

$$\tilde{\Theta}_{\beta_k}(Z_1, Z_2) = -\frac{\partial \Phi(\beta_1, \dots, \beta_n)}{\partial \beta_k} = \tilde{\Theta}^k(Z_1, Z_2) + \langle Q_k, ad_{Z_1}(Z_2) \rangle$$

Poly-symplectic extension of Souriau Lie Groups Thermodynamics

- Compared to Souriau model, heat is replaced by previous polysymplectic model:

$$Q_k = \frac{\partial \Phi(\beta_1, \dots, \beta_n)}{\partial \beta_k} = \frac{\int_M U^{\otimes k}(\xi) \cdot e^{-\sum_{k=1}^n \langle \beta_k, U^{\otimes k}(\xi) \rangle} d\omega}{\int_M e^{-\sum_{k=1}^n \langle \beta_k, U^{\otimes k}(\xi) \rangle} d\omega} \quad \text{with} \quad Q = (Q_1, \dots, Q_n) \in \mathfrak{g}^* \times \dots \times \mathfrak{g}^{*(n)}$$

- with characteristic function: $\Phi(\beta_1, \dots, \beta_n) = -\log \int_M e^{-\sum_{k=1}^n \langle \beta_k, U^{\otimes k}(\xi) \rangle} d\omega$
- We extrapolate Souriau results, who proved that $\int_M U^{\otimes k}(\xi) \cdot e^{-\langle \beta_k, U^{\otimes k}(\xi) \rangle} d\omega$ is locally normally convergent using $\|U^{\otimes k}\| = \text{Sup} \langle E, U \rangle^k$, a M -multi-linear norm and where $U^{\otimes k} = U \otimes U \dots \otimes U$ is defined as a tensorial product.
- Entropy is defined by Legendre transform of Souriau-Massieu characteristic function:

$$S(Q_1, \dots, Q_n) = \sum_{k=1}^n \langle \beta_k, Q_k \rangle - \Phi(\beta_1, \dots, \beta_n) \quad \text{with} \quad \beta_k = \frac{\partial S(Q_1, \dots, Q_n)}{\partial Q_k}$$

Poly-symplectic extension of Souriau Lie Groups Thermodynamics

- The Gibbs density could be then extended with respect to high order temperatures by:

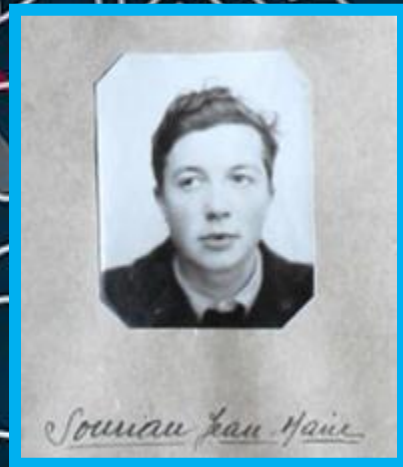
$$p_{Gibbs}(\xi) = e^{\Phi(\beta_1, \dots, \beta_n) - \sum_{k=1}^n \langle \beta_k, U^{\otimes k}(\xi) \rangle} = \frac{e^{-\sum_{k=1}^n \langle \beta_k, U^{\otimes k}(\xi) \rangle}}{\int_M e^{-\sum_{k=1}^n \langle \beta_k, U^{\otimes k}(\xi) \rangle} d\omega}$$

with $U^{\otimes k} = U \otimes^{(k)} U \dots \otimes U$ and $\Phi(\beta_1, \dots, \beta_n) = -\log \int_M e^{-\sum_{k=1}^n \langle \beta_k, U^{\otimes k}(\xi) \rangle} d\omega$

where $S(Q_1, \dots, Q_n) = \sum_{k=1}^n \langle \beta_k, Q_k \rangle - \Phi(\beta_1, \dots, \beta_n)$ with $\beta_k = \frac{\partial S(Q_1, \dots, Q_n)}{\partial Q_k}$

$$\left\{ \begin{array}{l} Q_k = \frac{\partial \Phi(\beta_1, \dots, \beta_n)}{\partial \beta_k} \quad \text{with } Q = (Q_1, \dots, Q_n) \in \mathfrak{g}^* \times \dots \times \mathfrak{g}^{(n)} \\ \beta_k = \frac{\partial S(Q_1, \dots, Q_n)}{\partial Q_k} \quad \text{with } \beta = (\beta_1, \dots, \beta_n) \in \mathfrak{g} \times \dots \times \mathfrak{g}^{(n)} \end{array} \right.$$

Supervised & Non-Supervised Lie Group Machine Learning



Supervised & Non-Supervised Learning on Lie Groups



Geodesic Natural Gradient on Lie Algebra

Extension of Neural Network Natural Gradient from Information Geometry on Lie Algebra for Lie Groups Machine Learning

Souriau-Fisher Metric on Coadjoint Orbits

Extension of Fisher Metric for Lie Group through homogeneous Symplectic Manifolds on Lie Group Co-Adjoint Orbits



Souriau Exponential Map on Lie Algebra

Exponential Map for Geodesic Natural Gradient on Lie Algebra based on Souriau Algorithm for Matrix Characteristic Polynomial



Souriau Maximum Entropy Density on Co-Adjoint Orbits

Covariant Maximum Entropy Probability Density for Lie Groups defined with Souriau Moment Map, Co-Adjoint Orbits Method & Kirillov Representation Theory

Fréchet Geodesic Barycenter by Hermann Karcher Flow

Extension of Mean/Median on Lie Group by Fréchet Definition of Geodesic Barycenter on Souriau-Fisher Metric Space, solved by Karcher Flow



Symplectic Integrator preserving Moment Map

Extension of Neural Network Natural Gradient to Geometric Integrators as Symplectic integrators that preserve moment map

Mean-Shift on Lie Groups with Souriau-Fisher Distance

Extension of Mean-Shift for Homogeneous Symplectic Manifold and Souriau-Fisher Metric Space



LIE GROUP SUPERVISED LEARNING

LIE GROUP NON-SUPERVISED LEARNING

Example of MATRIX LIE GROUPS & THALES APPLICATIONS

- Matrix Lie Group for Gaussian Density

$$\begin{bmatrix} Y \\ 1 \end{bmatrix} = \begin{bmatrix} R^{1/2} & m \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ 1 \end{bmatrix}, \quad \begin{cases} R^{1/2} \in T_n^+ \\ (R^{1/2} : \text{Cholesky of } R) \\ m \in R^n \end{cases}$$

Statistical
Radar/Sonar
Processing

- Matrix Lie Group $SO(3)$

$$SO(3) = \left\{ \Omega / \Omega^T \Omega = \Omega \Omega^T = I, \det^2 \Omega = 1 \right\}$$

Inertial Central
Processing (IMU) +
Video (Pose
Recognition)

- Matrix Lie Group $SE(3)$

$$\begin{bmatrix} Z' \\ 1 \end{bmatrix} = \begin{bmatrix} \Omega & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Z \\ 1 \end{bmatrix}, \quad \begin{cases} \Omega \in SO(3) \\ t \in R^3 \end{cases}$$

3D Intrinsic
TRACKING

- Matrix Affine Group $Aff(1)$

$$\begin{bmatrix} X' \\ 1 \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ 1 \end{bmatrix}, \quad a \in R_+^* \text{ et } b \in R$$

High Resolution
Doppler & Array
Processing

Example of MATRIX LIE GROUPS & THALES APPLICATIONS

- Special unitary Group $SU(2)$, and $SU(1,1)$

$$SU(2) = \left\{ \begin{pmatrix} a & -b^* \\ b & a^* \end{pmatrix} / a, b \in \mathbb{C}, |a|^2 + |b|^2 = 1 \right\}$$

$$SU(1,1) = \left\{ \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} / \alpha, \beta \in \mathbb{C}, |\alpha|^2 - |\beta|^2 = 1 \right\}$$

- Symplectic Group

$$Sp(n, \mathbb{C}) = \left\{ S / S^T \Omega S = \Omega, \Omega = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \right\}$$

- Special Euclidean Group

$$\begin{bmatrix} Z' \\ 1 \end{bmatrix} = \begin{bmatrix} \Omega & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Z \\ 1 \end{bmatrix}, \quad \begin{cases} \Omega \in SO(2) \\ t \in \mathbb{R}^2 \end{cases}$$

Stationnary and
Non-stationary
Micro-Doppler

STAP
Space-Time
Adaptive
Processing

EO/IR Image
Processing
(directional
gradient)

Radar Doppler Data & SU(1,1) Lie Group (1/2)

Covariance Matrix of Stationary Doppler Signal (order 1)

$$\begin{cases} Z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \\ R = E[ZZ^+] \end{cases} \Rightarrow R = \begin{bmatrix} E[|z_1|^2] & E[z_1 z_2^*] \\ E[z_2 z_1^*] & E[|z_2|^2] \end{bmatrix} \xrightarrow{\text{stationary}} \begin{cases} R = \begin{bmatrix} r_0 & r_1^* \\ r_1 & r_0 \end{bmatrix}, \det R > 0 \\ r_0 = E[|z_k|^2], r_1 = E[z_k z_{k-1}^*] \end{cases}$$

$$R = \begin{bmatrix} r_0 & r_1^* \\ r_1 & r_0 \end{bmatrix} = \begin{bmatrix} h & a - ib \\ a + ib & h \end{bmatrix} = h \begin{bmatrix} 1 & \mu^* \\ \mu & 1 \end{bmatrix} \text{ with } \begin{cases} h \in \mathbb{R}^{+*} \\ \mu \in D = \{w \in \mathbb{C} / |w| < 1\} \end{cases}$$

$$\Rightarrow R \in \mathbb{R}^{+*} \times D, \mu = \frac{a + ib}{h} \text{ such that } |\mu| < 1 \text{ with } \det R = h^2 - (a^2 + b^2) > 0$$

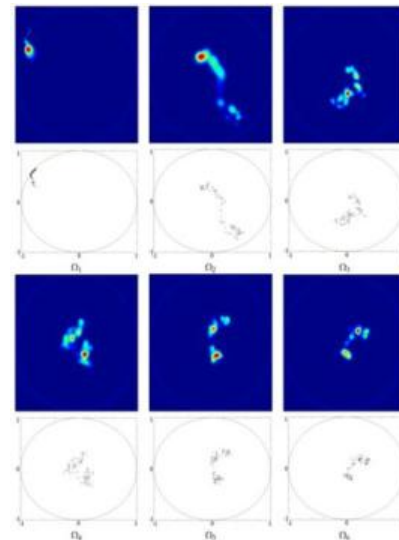
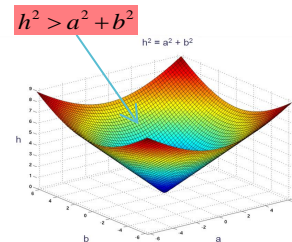
Covariance Matrix of Stationary Doppler Signal (order n)

➤ Verblunsky-Trench Theorem: Covariance matrix could be coded in product space of positive real line with **Poincaré unit Polydisk**

$$\varphi: THDP(n) \rightarrow R_+^* \times D^{n-1}$$

$$R_n = (h, \mu_1, \dots, \mu_{n-1}) \in \mathbb{R}^{+*} \times D^{n-1}$$

$$R_n \mapsto (P_0, \mu_1, \dots, \mu_{n-1}) \text{ with } (D^{n-1} = D \times \dots \times D)$$



Radar Doppler Data & $SU(1,1)$ Lie Group (2/2)

- The group of complex unimodular pseudo-unitary matrices $SU(1,1)$, are element u such that:

$$uMu^+ = M \quad \text{with} \quad M = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}$$

- We can show that the most general matrix u belongs to the Lie group given by:

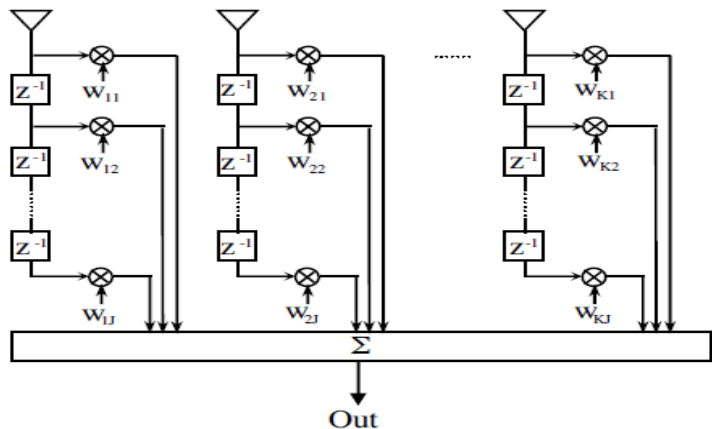
$$G = SU(1,1) = \left\{ \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} / |\alpha|^2 - |\beta|^2 = 1, \alpha, \beta \in \mathbb{C} \right\}$$

- $SU(1,1)$ is associated to group of holomorphic automorphisms of the Poincaré unit disk in the complex plane: $D = \{z = x + iy \in \mathbb{C} / |z| < 1\}$ by considering its action on the disk as : $g(z) = (\alpha z + \beta) / (\beta^* z + \alpha^*)$

- The following measure on Unit disk is invariant under the action of $SU(1,1)$:

$$d\mu_0(z, z^*) = \frac{1}{2\pi i} \frac{dz \wedge dz^*}{(1 - |z|^2)^2}$$

Radar Space-Time Processing & Symplectic Lie Group



$$Z = \begin{bmatrix} z_{1,1} \\ \vdots \\ z_{N,1} \\ \vdots \\ z_{1,M} \\ \vdots \\ z_{N,M} \end{bmatrix} \Rightarrow R = E[ZZ^+] = \begin{bmatrix} R_0 & R_1 & \cdots & R_n \\ R_1^+ & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1 \\ R_n^+ & \cdots & R_1^+ & R_0 \end{bmatrix}$$

Matrix Extension of Trench/Verblunsky Theorem: Existence of diffeomorphism φ and Siegel Polydisk (matrix extension of Poincaré Disk)

$$\varphi : TBTHPD_{n \times n} \rightarrow THPD_n \times SD^{n-1}$$

$$R \mapsto (R_0, A_1^1, \dots, A_{n-1}^{n-1})$$

$$\text{with } SD = \{Z \in Herm(n) / ZZ^+ < I_n\}$$

$$ds_{\text{Siegel}}^2 = Tr \left[(I - ZZ^+)^{-1} dZ (I - Z^+Z)^{-1} dZ^+ \right]$$

$$ds_{\text{Poincaré}}^2 = \frac{|dz|^2}{(1 - |z|^2)^2} = (1 - zz^*) dz (1 - z^*z) dz^*$$

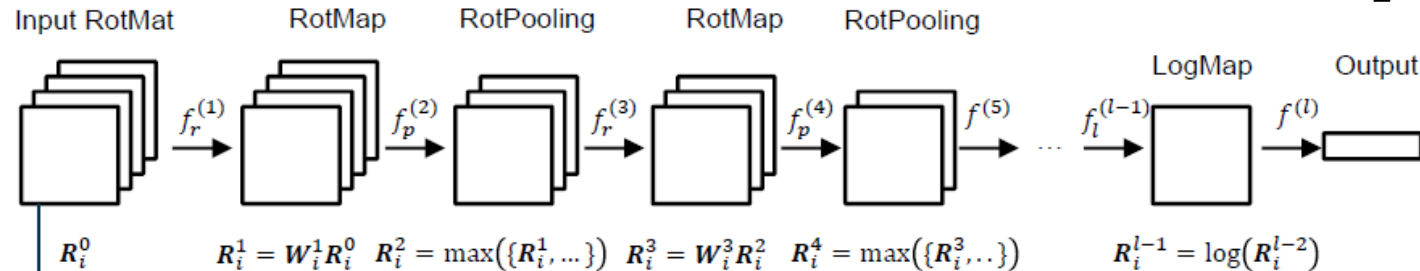
Time series of SO(3) Lie Group to learn articulated movement

SO(3) for articulated Movement

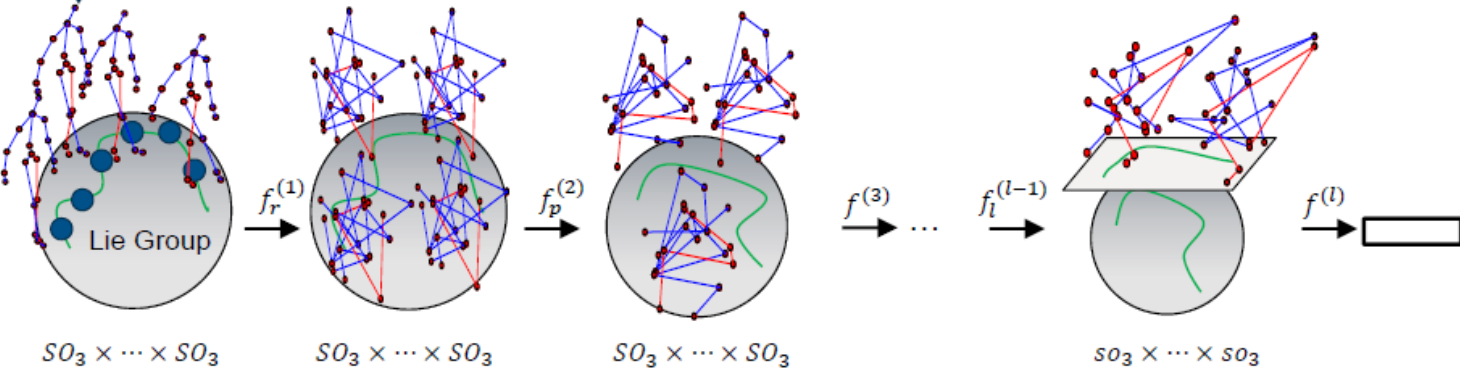
$$SO(3) = \left\{ \Omega / \Omega^T \Omega = \Omega \Omega^T = I, \det^2 \Omega = 1 \right\}$$

Vector of $SO(3)$

$$\begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \vdots \\ \Omega_p \end{bmatrix} \in SO(3) \times \dots \times SO(3)$$



$$R_i^0 \quad R_i^1 = W_i^1 R_i^0 \quad R_i^2 = \max(\{R_i^1, \dots\}) \quad R_i^3 = W_i^3 R_i^2 \quad R_i^4 = \max(\{R_i^3, \dots\}) \quad R_i^{l-1} = \log(R_i^{l-2})$$



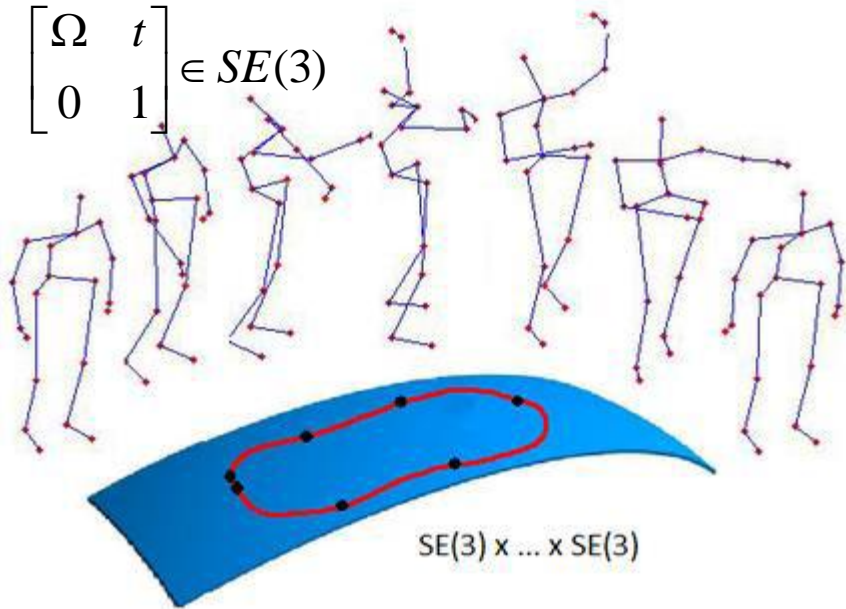
Vectors of SE(3) Lie Group to learn articulated Posture

SE(3) for articulated Posture

$$SO(3) = \{ \Omega / \Omega^T \Omega = \Omega \Omega^T = I, \det^2 \Omega = 1 \}$$

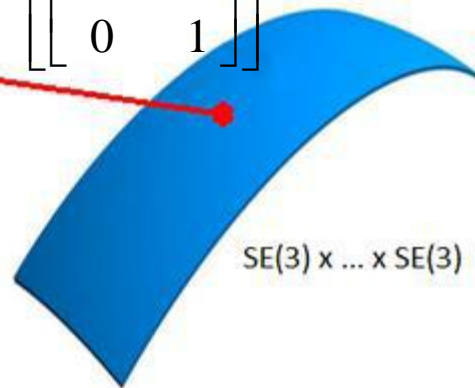
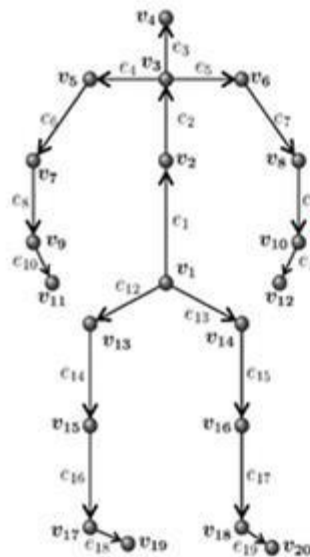
$$\begin{bmatrix} Z' \\ 1 \end{bmatrix} = \begin{bmatrix} \Omega & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Z \\ 1 \end{bmatrix}, \quad \begin{cases} \Omega \in SO(3) \\ t \in \mathbb{R}^3 \end{cases}$$

$$\begin{bmatrix} \Omega & t \\ 0 & 1 \end{bmatrix} \in SE(3)$$



Vector of $SE(3)$

$$\begin{bmatrix} \begin{bmatrix} \Omega_1 & t_1 \\ 0 & 1 \end{bmatrix} \\ \begin{bmatrix} \Omega_2 & t_2 \\ 0 & 1 \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \Omega_m & t_m \\ 0 & 1 \end{bmatrix} \end{bmatrix} \in SE(3) \times \dots \times SE(3)$$



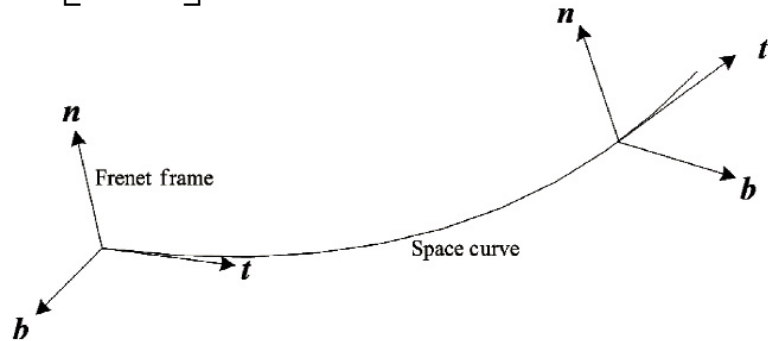
Time series of SE(3) Lie Group to learn continuous movement

SE(3) for continuous movement

$$SO(3) = \left\{ \Omega / \Omega^T \Omega = \Omega \Omega^T = I, \det^2 \Omega = 1 \right\}$$

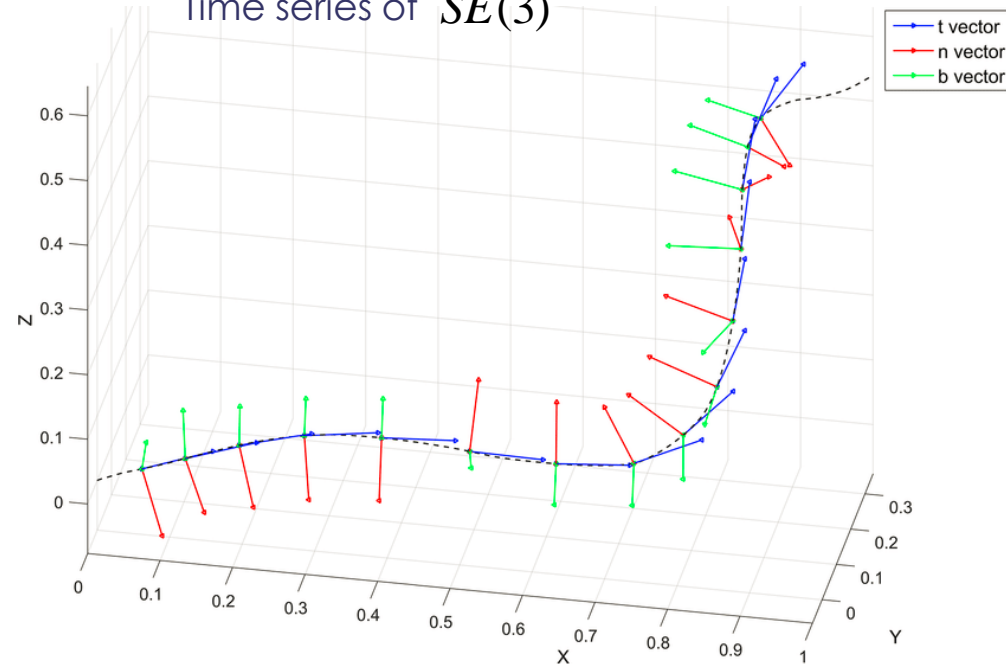
$$\begin{bmatrix} Z' \\ 1 \end{bmatrix} = \begin{bmatrix} \Omega & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Z \\ 1 \end{bmatrix}, \quad \begin{cases} \Omega \in SO(3) \\ t \in \mathbb{R}^3 \end{cases}$$

$$\begin{bmatrix} \Omega & t \\ 0 & 1 \end{bmatrix} \in SE(3)$$

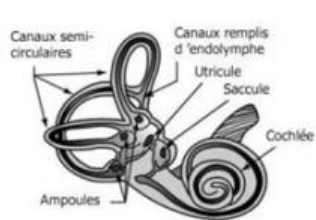


$$\left\{ \begin{bmatrix} \Omega_1 & t_1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \Omega_2 & t_2 \\ 0 & 1 \end{bmatrix}, \dots, \begin{bmatrix} \Omega_n & t_n \\ 0 & 1 \end{bmatrix} \right\}$$

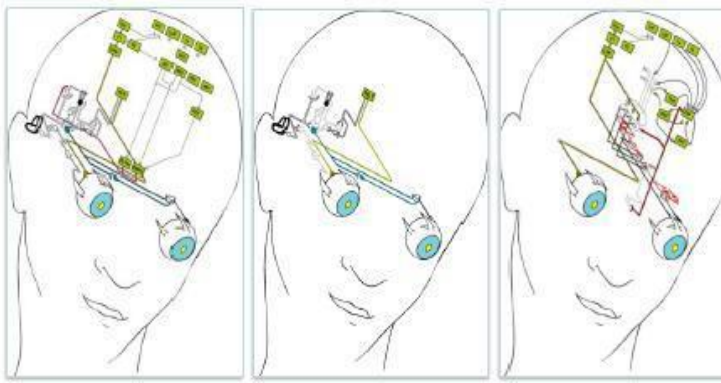
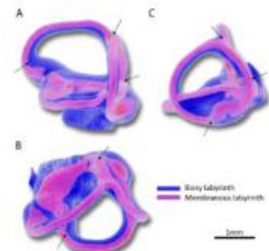
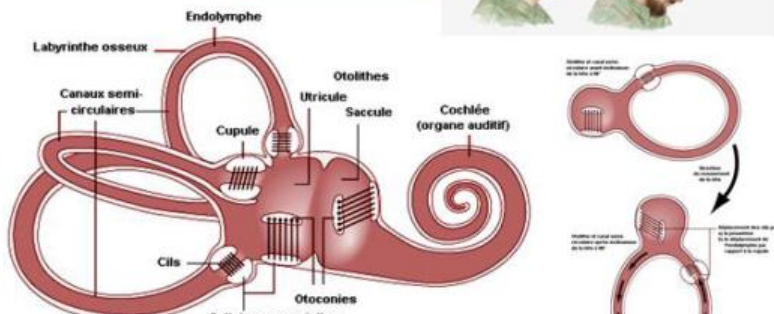
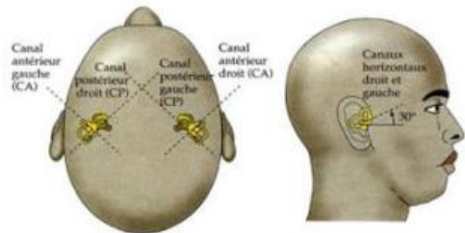
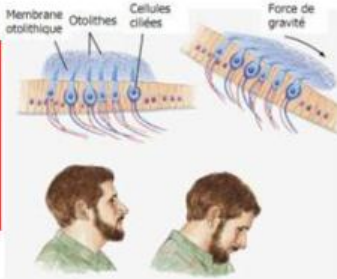
Time series of SE(3)



Geolocalization and Navigation by SE(3) Lie Group: Visio-Inertial SLAM: Human Visio-Vestibular Brain System



Coding of Homogeneous Galileo Group By Vestibular System and Otoliths



VINet: Visual-Inertial Odometry as a Sequence-to-Sequence Learning Problem

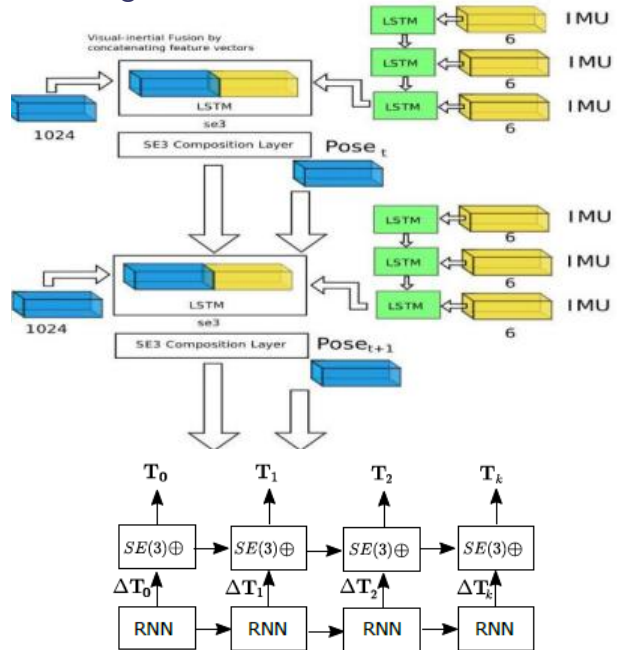


Illustration of the SE(3) composition layer - a parameter-free layer which concatenates transformations between frames on SE(3).

Image Processing by SE(2) Group: hypoelliptic diffusion in SE(2): analogy with Brain Orientation Maps V1

Image Processing with oriented gradient by SE(2)

$$\begin{bmatrix} Z' \\ 1 \end{bmatrix} = \begin{bmatrix} \Omega & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Z \\ 1 \end{bmatrix}, \quad \begin{cases} \Omega \in SO(2) \\ t \in \mathbb{R}^2 \end{cases}$$

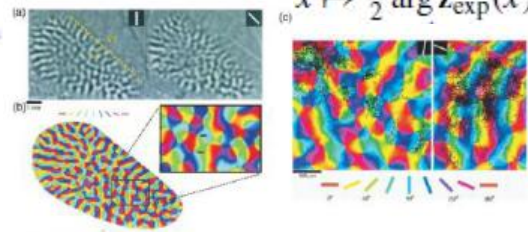
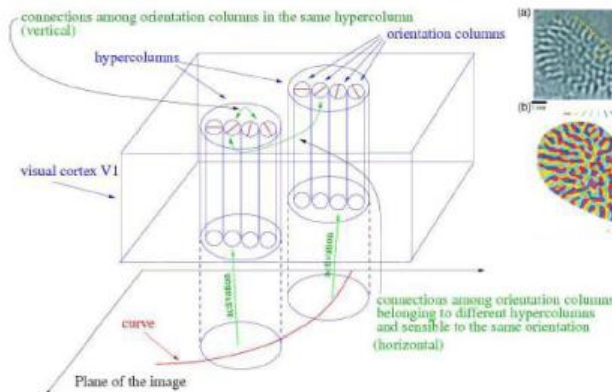
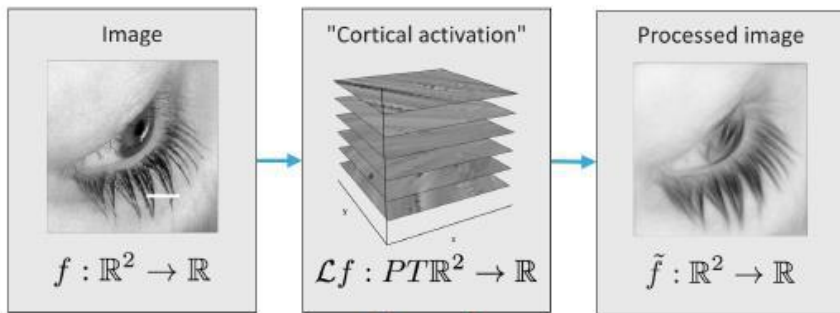
$$Z : \mathbb{R}^2 \rightarrow \mathbb{C}, \quad a = \rho e^{i\theta} \mapsto r(a) e^{i\varphi(a)}$$

$$\omega_S = -\sin(\theta) dx + \cos(\theta) dy$$

$$\cos(\theta) (dy - p dx) = \cos(\theta) \omega.$$

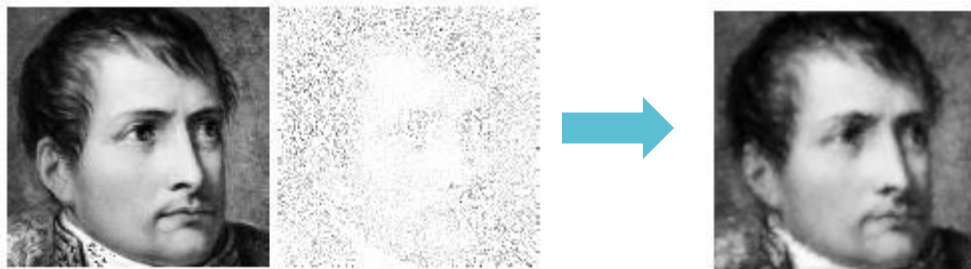
$$z_{\text{exp}}(x_0) = \sum_{k=1}^N a_{\vartheta_k}(x_0) e^{2i\vartheta_k}$$

$$x \mapsto \frac{1}{2} \arg z_{\text{exp}}(x)$$

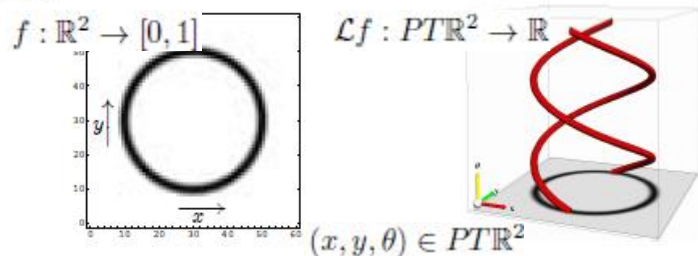


Pinwheel structure of V1

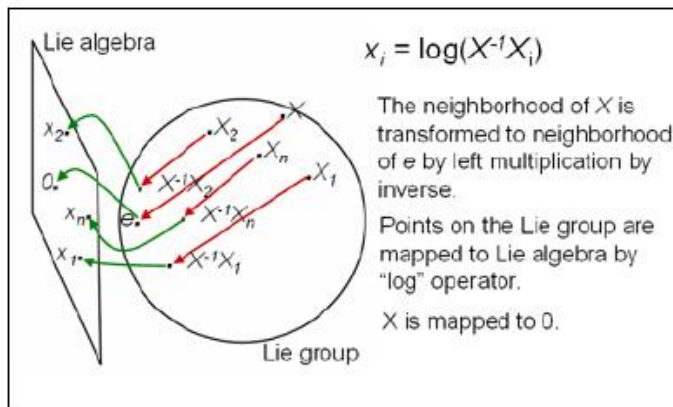
SE(2) double covering of $P\mathbb{T}\mathbb{R}^2$



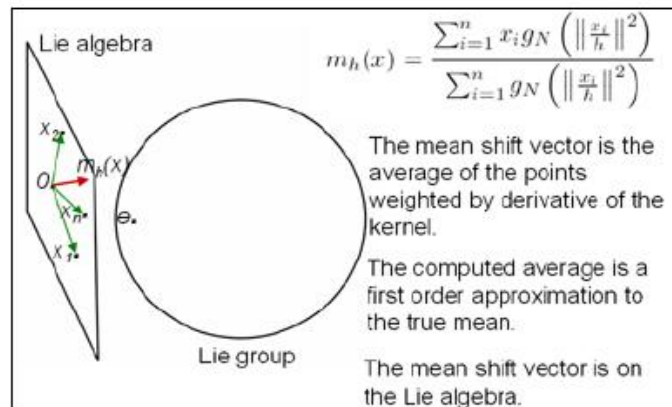
A scheme of the primary visual cortex V1



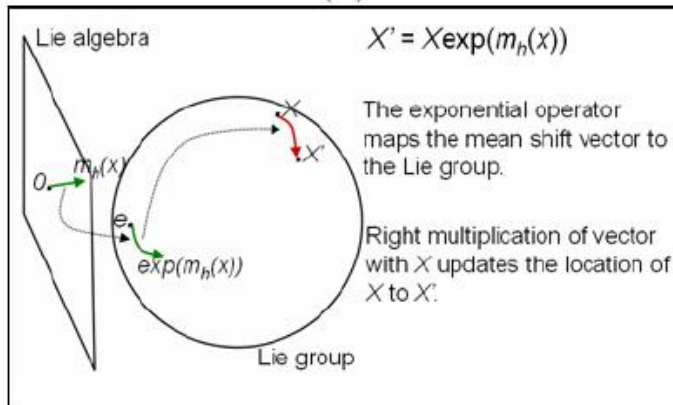
Extension of Mean-Shift for Lie Group (e.g. with SO(3))



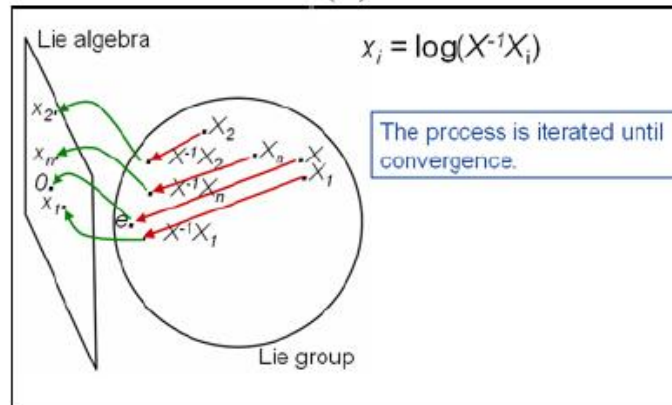
(1)



(2)



(3)



(4)

$$d(X, Y) = \|\log(X^{-1}Y)\|$$

$$k_N(s) = e^{-\frac{1}{2}s}$$

$$\hat{f}(X) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^n k_N \left(\left\| \frac{\log(X^{-1}X_i)}{h} \right\|^2 \right)$$

$$x_i = \log(X^{-1}X_i)$$

$$m_h(x) = \frac{\sum_{i=1}^n x_i g_N \left(\left\| \frac{x_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g_N \left(\left\| \frac{x_i}{h} \right\|^2 \right)}$$

Algorithm: MEAN SHIFT ON LIE GROUPS

Given: Data points on Lie group $\{X_j\}_{j=1..n}$

for $j \leftarrow 1..n$

$X \leftarrow X_j$

repeat

for all data points

$$x_i \leftarrow \log(X^{-1}X_i)$$

$$m_h(x) \leftarrow \frac{\sum_{i=1}^n x_i g_N \left(\left\| \frac{x_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g_N \left(\left\| \frac{x_i}{h} \right\|^2 \right)}$$

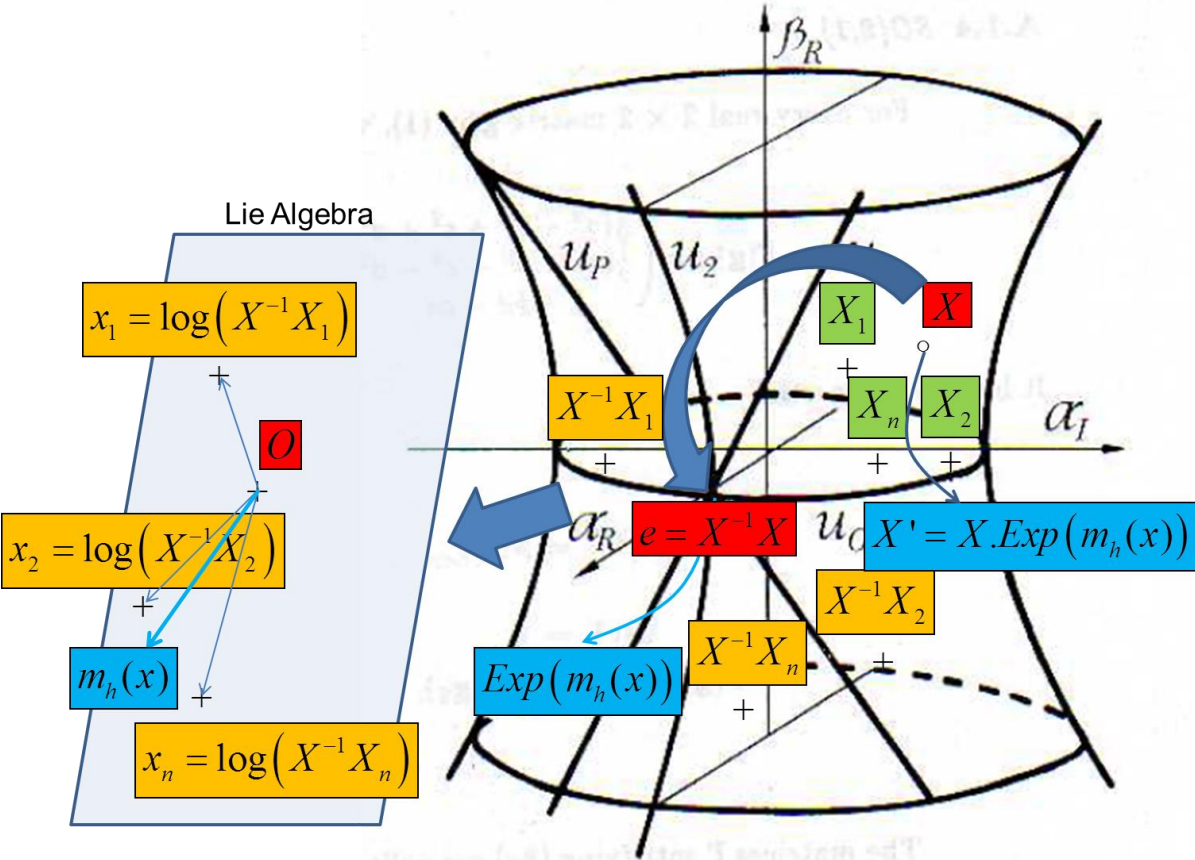
$$X \leftarrow X \exp(m_h(x))$$

until $m_h(x) < \epsilon$

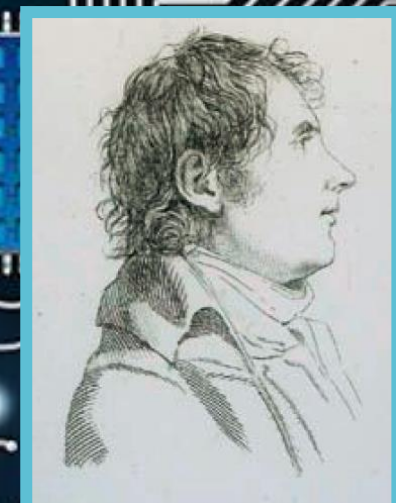
Store X as a local mode.

Report distinct modes.

SU(1,1) Group Manifold



Representation Theory & Orbits Methods: Notion Extension of « Gauss Density » for Lie Group by Maximum Entropy Density and Fourier Transform for Non- Commutative Harmonic analysis



Souriau-Gibbs density covariant by a group

■ Souriau has introduced a Gibbs density on Lie Algebra that is covariant with respect to action of the group

$$p_{Gibbs}(\xi) = e^{\Phi(\beta) - \langle \beta, U(\xi) \rangle} = \frac{e^{-\langle \beta, U(\xi) \rangle}}{\int_M e^{-\langle \beta, U(\xi) \rangle} d\omega}, \quad \text{with } \Phi(\beta) = -\log \int_M e^{-\langle \beta, U(\xi) \rangle} d\omega$$

$$Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \frac{\int_M U(\xi) e^{-\langle \beta, U(\xi) \rangle} d\omega}{\int_M e^{-\langle \beta, U(\xi) \rangle} d\omega} = \int_M U(\xi) p(\xi) d\omega$$

■ We will illustrate and compute this density for Lie Group $SU(1,1)$ for Poincaré Unit Disk considered as an homogeneous symplectic manifold.

Multivariate Gaussian Density as 1st order Maximum Entropy in Souriau Book (Chapter IV)

Exemple : (loi normale) :

Prenons le cas $V = R^n$, $\lambda =$ mesure de Lebesgue; $\Psi(x) \equiv \begin{pmatrix} x \\ x \otimes x \end{pmatrix}$;

un élément Z du dual de E peut se définir par la formule

$$Z(\Psi(x)) \equiv \bar{a} \cdot x + \frac{1}{2} \bar{x} \cdot H \cdot x$$

[$a \in R^n$; $H =$ matrice symétrique]. On vérifie que la convergence de l'intégrale I_0 a lieu si la matrice H est positive ⁽¹⁾; dans ce cas la loi de Gibbs s'appelle *loi normale de Gauss*; on calcule facilement I_0 en faisant le changement de variable $x^* = H^{1/2} x + H^{-1/2} a$ ⁽²⁾; il vient

$$z = \frac{1}{2} [\bar{a} \cdot H^{-1} \cdot a - \log(\det(H)) + n \log(2\pi)]$$

alors la convergence de I_1 a lieu également; on peut donc calculer M , qui est défini par les moments du premier et du second ordre de la loi (16.196); le calcul montre que le moment du premier ordre est égal à $-H^{-1} \cdot a$ et que les composantes du tenseur *variance* (16.196) sont égales aux éléments de la matrice H^{-1} ; le moment du second ordre s'en déduit immédiatement.

La formule (16.200) donne l'entropie :

$$s = \frac{n}{2} \log(2\pi e) - \frac{1}{2} \log(\det(H)) ;$$

⁽¹⁾ Voir *Calcul linéaire*, tome II.

⁽²⁾ C'est-à-dire en recherchant l'image de la loi par l'application $x \mapsto x^*$.

DÉPARTEMENT MATHÉMATIQUE
Dirigé par le Professeur P. LÉLONG

STRUCTURE DES SYSTÈMES DYNAMIQUES

Maîtrises de mathématiques

J.-M. SOURIAU
Professeur de Physique Mathématique
à la Faculté des Sciences de Montréal

DUNOD
S.A.S. 1988
93770

Multivariate Gaussian law parameterized by moments

$$p_{\hat{\xi}}(\xi) = \frac{1}{(2\pi)^{n/2} \det(R)^{1/2}} e^{-\frac{1}{2}(z-m)^T R^{-1}(z-m)}$$

$$\begin{aligned} \frac{1}{2}(z-m)^T R^{-1}(z-m) &= \frac{1}{2} \left[z^T R^{-1} z - m^T R^{-1} z - z^T R^{-1} m + m^T R^{-1} m \right] \\ &= \frac{1}{2} z^T R^{-1} z - m^T R^{-1} z + \frac{1}{2} m^T R^{-1} m \end{aligned}$$

$$p_{\hat{\xi}}(\xi) = \frac{1}{(2\pi)^{n/2} \det(R)^{1/2} e^{\frac{1}{2} m^T R^{-1} m}} e^{-\left[-m^T R^{-1} z + \frac{1}{2} z^T R^{-1} z \right]} = \frac{1}{Z} e^{-\langle \xi, \beta \rangle}$$

$$\xi = \begin{bmatrix} z \\ zz^T \end{bmatrix} \text{ and } \beta = \begin{bmatrix} -R^{-1} m \\ \frac{1}{2} R^{-1} \end{bmatrix} = \begin{bmatrix} a \\ H \end{bmatrix} \text{ with } \langle \xi, \beta \rangle = a^T z + z^T H z = \text{Tr} \left[z a^T + H^T z z^T \right]$$

Gaussian Density is a 1st order Maximum Entropy Density !

Representation theory

- a branch of mathematics that studies abstract algebraic structures by representing their elements as linear transformations of vector spaces
- a representation makes an abstract algebraic object (groups, associative algebras and Lie algebras) more concrete by describing its elements by matrices and the algebraic operations in terms of matrix addition and matrix multiplication.
- The most prominent and first of these is the representation theory of groups, in which elements of a group are represented by invertible matrices in such a way that the group operation is matrix multiplication.
- Representation theory is a useful method because it reduces problems in abstract algebra to problems in linear algebra.
- Representation theory is pervasive across fields of mathematics : **generalizes Fourier analysis via harmonic analysis**, is **connected to geometry via invariant theory** and the Erlangen program, has an impact in number theory via automorphic forms and the Langlands program.

Classical Commutative Harmonic Analysis

$G = \mathbb{T}^n = \mathbb{R}^n / \mathbb{Z}^n$: Fourier series, $G = \mathbb{R}^n$: Fourier Transform

G group character (linked to e^{ikx}):
$$\begin{cases} \chi : G \rightarrow U \\ U = \{z \in \mathbb{C} / |z| = 1\} \end{cases}$$

$$\hat{G} = \{ \chi / \chi_1 \cdot \chi_2(g) = \chi_1(g) \chi_2(g) \}$$

Fourier Transform

$$\varphi : G \rightarrow \mathbb{C}$$

$$\hat{\varphi} : \hat{G} \rightarrow \mathbb{C}$$

$$g \mapsto \varphi(g) = \int_{\hat{G}} \hat{\varphi}(\chi) \chi(g)^{-1} d\chi$$

$$\chi \mapsto \hat{\varphi}(\chi) = \int_G \varphi(g) \chi(g) dg$$

Fourier-Plancherel formula

$$\varphi(e) = \int_{\hat{G}} \hat{\varphi}(\chi) d\chi$$

Non-Commutative Harmonic Analysis

- Unitary Irreducible Representation: $U : G \rightarrow U(H)$ with H Hilbert space
- Character: $\chi_U(g) = \text{tr} U_g$

Fourier Transform for Non-Commutative Group

- Fourier Transform: $U_\varphi = \int_G \varphi(g) U_g dg$
- Character: $\chi_U(g) = \text{tr} U_\varphi$

Representation theory

- All representations are G -equivariante (covariante under the action of G)
- We can consider $f \in \mathfrak{g}^*$, its coadjoint orbits $O_f = \{Ad_g^* f = g^{-1}fg, g \in G\}$ or $G.f$
- Duflo representation: $T_f : X_f^{irr} \rightarrow \hat{G}$

$$\tau \mapsto U = T_f(\tau)$$

Dixmier/Kirillov/Duflo/Vergne Representation Theory

Character-Distribution

➤ (Schwarz) Distribution on G : $\chi_U(g) = \text{tr} U_\varphi$ with $U_\varphi = \int_G \varphi(g) U_g dg$

Character Formula

➤ With only Fourier transform on Lie algebra. We should use Exponential map

$$U_\psi = \int_{\mathfrak{g}} \psi(X) U_{\exp(X)} dX$$

➤ Character formula:

$$\text{tr} U_\psi = \dim \tau \cdot \mu_{G,f} \left(\hat{\psi \cdot j^{-1}} \right) \begin{cases} \hat{\psi \cdot j^{-1}} : \text{Fourier Transf. of } \psi \cdot j^{-1} \\ \hat{\psi \cdot j^{-1}} : \mathfrak{g} \rightarrow \mathfrak{g}^* \end{cases} \begin{matrix} \mu_{G,f} : \text{Liouville measure on } O = G \cdot f, f \in \mathfrak{g}^* \\ \mu_{G,f} \left(\hat{\psi \cdot j^{-1}} \right) : \text{Integral of } \hat{\psi \cdot j^{-1}} \text{ wrt } \mu_{G,f} \text{ (Radon measure)} \end{matrix}$$

➤ Kirillov function:

$$j(X) = \left(\det s(ad_X) \right)^{1/2} \quad \text{with} \quad s(x) = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} \left(\frac{x}{2} \right)^{2n} = \text{sh} \left(\frac{x}{2} \right) / \left(\frac{x}{2} \right)$$

$$ad_X Y = [X, Y] = XY - YX$$

Kirillov Character formula:

$$\chi_U(\exp(X)) = \text{tr} U_{\exp(X)} = j(X)^{-1} \hat{\mu}_0(X)$$

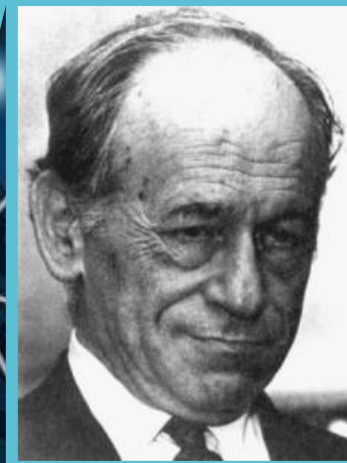
$$\chi_U(\exp(X)) = j(X)^{-1} \int_0 \! \! \! \int e^{i\langle f, X \rangle} d\mu_0(f)$$

Fourier Transform:

$$\int_0 \! \! \! \int e^{i\langle f, X \rangle} d\mu_0(f) = j(X) \text{tr} U_{\exp(X)}$$

$$j(X) = \left(\det \left(\frac{e^{ad_X/2} - e^{-ad_X/2}}{ad_X/2} \right) \right)^{1/2}$$

Gibbs Density on Poincaré Unit Disk from Souriau Lie Groups Thermodynamics and $SU(1,1)$ Coadjoint Orbits



Distance Between Gaussian Density with Fisher Metric

Fisher Matrix for Gaussian Densities:

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

► Fisher matrix induced the following differential metric :

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

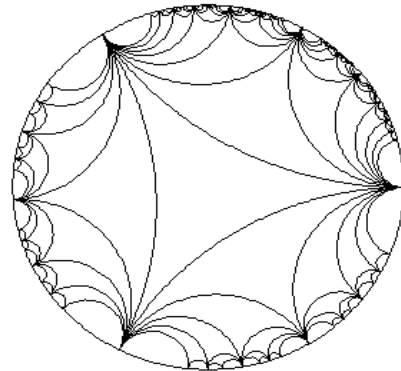
► Poincaré Model of upper half-plane and unit disk

$$z = \frac{m}{\sqrt{2}} + i \cdot \sigma \quad \omega = \frac{z - i}{z + i} \quad (|\omega| < 1)$$

$$\Rightarrow ds^2 = 8 \cdot \frac{|d\omega|^2}{(1 - |\omega|^2)^2}$$

La géométrie paramètres
des gaussiennes est la
géométrie du disque de

OPÉ Poincaré



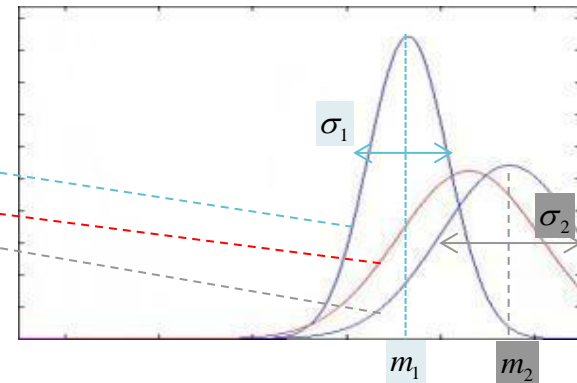
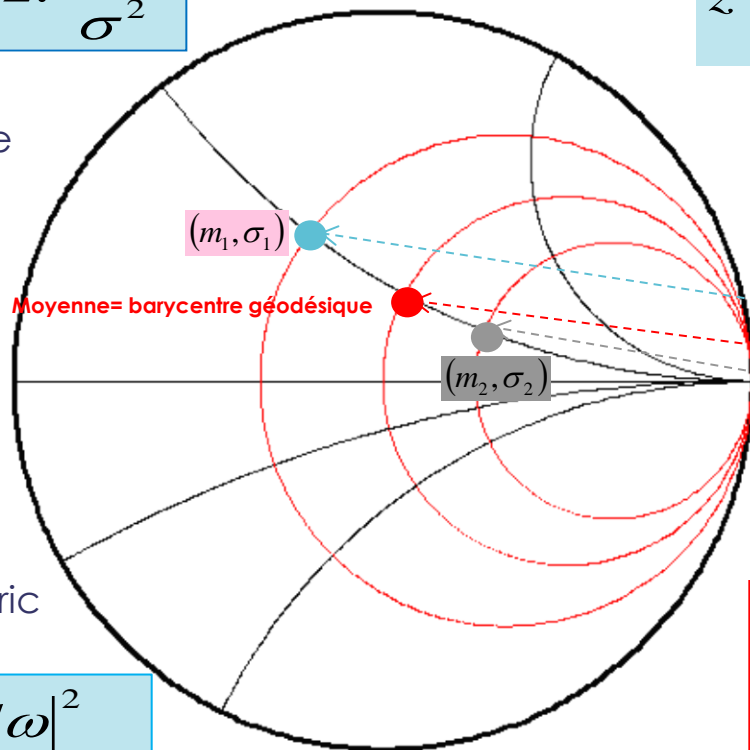
1 monovariate gaussian = 1 point in Poincaré unit disk

$$ds^2 = \frac{dm^2}{\sigma^2} + 2 \cdot \frac{d\sigma^2}{\sigma^2}$$

Fisher Metric in Poincaré Half-Plane

$$z = \frac{m}{\sqrt{2}} + i \cdot \sigma$$

$$\omega = \frac{z - i}{z + i} \quad (|\omega| < 1)$$



Poincaré-Fisher metric In Unit Disk

$$ds^2 = 8 \cdot \frac{|d\omega|^2}{(1 - |\omega|^2)^2}$$

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

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

- What is Gaussian density (density of maximum entropy) of Gaussian densities ?
- What is a « Gaussian density (density of maximum entropy) » in Poincaré Unit Disk ?

Covariance Matrix of Stationary Doppler Signal (order 1)

$$\mathbf{Z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \Rightarrow \mathbf{R} = E[\mathbf{Z}\mathbf{Z}^+] = \begin{bmatrix} E[|z_1|^2] & E[z_1 z_2^*] \\ E[z_2 z_1^*] & E[|z_2|^2] \end{bmatrix} \xrightarrow{\text{stationary}} \begin{cases} \mathbf{R} = \begin{bmatrix} r_0 & r_1^* \\ r_1 & r_0 \end{bmatrix} \text{ with } \det R > 0 \\ r_0 = E[|z_k|^2], r_1 = E[z_k z_{k-1}^*] \end{cases}$$

$$\mathbf{R} = \begin{bmatrix} r_0 & r_1^* \\ r_1 & r_0 \end{bmatrix} = \begin{bmatrix} h & a - ib \\ a + ib & h \end{bmatrix} = h \begin{bmatrix} 1 & \mu^* \\ \mu & 1 \end{bmatrix} \text{ with } \begin{cases} h \in \mathbb{R}^{+*} \\ \mu \in D = \{w \in \mathbb{C} / |w| < 1\} \end{cases} \Rightarrow \mathbf{R} \in \mathbb{R}^{+*} \times D$$

Covariance Matrix of Stationary Doppler Signal (order n)

➤ Verblunsky-Trench Theorem: Covariance matrix could be coded in product space of positive real line with Poincaré unit Polydisk

$$\mathbf{R}_n = (h, \mu_1, \dots, \mu_{n-1}) \in \mathbb{R}^{+*} \times D^{n-1} \text{ with } (D^{n-1} = D \times \dots \times D)$$

Poincaré Unit Disk and $SU(1,1)$ Lie Group

- The group of complex unimodular pseudo-unitary matrices $SU(1,1)$, are element u such that:

$$uMu^+ = M \quad \text{with} \quad M = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}$$

- We can show that the most general matrix u belongs to the Lie group given by:

$$G = SU(1,1) = \left\{ \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} / |a|^2 - |b|^2 = 1, a, b \in \mathbb{C} \right\}$$

- Its Cartan decomposition is given by:

$$\begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} = |a| \begin{pmatrix} 1 & z \\ z^* & 1 \end{pmatrix} \begin{pmatrix} \frac{a}{|a|} & 0 \\ 0 & \frac{a^*}{|a|} \end{pmatrix} \quad \text{with} \quad z = b(a^*)^{-1}, |a| = \left(1 - |z|^2\right)^{-1/2}$$

Poincaré Unit Disk and $SU(1,1)$ Lie Group

► We observe that:

$$\begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} \begin{pmatrix} 1 & z \\ z^* & 1 \end{pmatrix} = |a'| \begin{pmatrix} 1 & z' \\ z'^* & 1 \end{pmatrix} \begin{pmatrix} \frac{a'}{|a'|} & 0 \\ 0 & \frac{a'^*}{|a'|} \end{pmatrix} \quad \text{with} \quad \begin{cases} a' = bz'^* + a \\ z' = \frac{az + b}{b^*z + a^*} \end{cases}$$

► $SU(1,1)$ is associated to group of holomorphic automorphisms of the Poincaré unit disk in the complex plane: $D = \{z = x + iy \in \mathbb{C} / |z| < 1\}$ by considering its action on the disk as : $g(z) = (az + b) / (b^*z + a^*)$

► The following measure on Unit disk $d\mu_0(z, z^*) = \frac{1}{2\pi i} \frac{dz \wedge dz^*}{(1 - |z|^2)^2}$ is invariant

under the action of $SU(1,1)$ captured by the fractional holomorphic transformation:

$$\frac{dz' \wedge dz'^*}{(1 - |z'|^2)^2} = \frac{dz \wedge dz^*}{(1 - |z|^2)^2}$$

Poincaré Unit Disk and $SU(1,1)$ Lie Group

➤ The complex unit disk admits a Kähler structure determined by potential function:

$$\Phi(z', z^*) = -\log(1 - z' z^*)$$

➤ The invariant 2-form, which is closed $d\Omega = 0$, is:

$$\Omega = \frac{1}{i} \frac{\partial \log \Phi(z, z^*)}{\partial z \partial z^*} dz \wedge dz^* = \frac{1}{i} \frac{dz \wedge dz^*}{(1 - |z|^2)^2}$$

Poincaré Unit Disk and $SU(1,1)$ Lie Group

- The group $SU(1,1)$ is isomorphic to the group $SL(2, \mathbb{R})$ as a real Lie group, and the Lie algebra $\mathfrak{g} = \mathfrak{su}(1,1)$ is given by:

$$\mathfrak{g} = \left\{ \begin{pmatrix} -ir & \eta \\ \eta^* & ir \end{pmatrix} / r \in \mathbb{R}, \eta \in \mathbb{C} \right\}$$

with the following bases $(u_1, u_2, u_3) \in \mathfrak{g}$:

$$u_1 = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, u_2 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, u_3 = \frac{1}{2} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$$

with the commutation relation:

$$[u_3, u_2] = u_1, [u_3, u_1] = u_2, [u_2, u_1] = -u_3$$

Poincaré Unit Disk and $SU(1,1)$ Lie Group

➤ Dual base on dual Lie algebra is named

$$(u_1^*, u_2^*, u_3^*) \in \mathfrak{g}^*$$

➤ The dual vector space $\mathfrak{g}^* = \mathfrak{su}^*(1,1)$ can be identified with the subspace of $\mathfrak{sl}(2, \mathbb{C})$ of the form:

$$\mathfrak{g}^* = \left\{ \begin{pmatrix} z & x+iy \\ -x+iy & -z \end{pmatrix} = x \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} + z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} / x, y, z \in \mathbb{R} \right\}$$

➤ Coadjoint action of $g \in G$ on dual Lie algebra $\xi \in \mathfrak{g}^*$ is written $g \cdot \xi$

Coadjoint Orbit of $SU(1,1)$ and Souriau Moment Map

- We will use results of C. Cishahayo and S. de Bièvre and B. Cahen [8,9] for computation of moment map of $SU(1,1)$:
- Cishahayo C., de Bièvre S. : On the contraction of the discrete series of $SU(1;1)$, Annales de l'institut Fourier, tome 43, no 2, p. 551-567, (1993).
 - Cahen B. : Contraction de $SU(1,1)$ vers le groupe de Heisenberg, Travaux mathématiques, Fascicule XV, pp.19-43, (2004).
- Let $r \in \mathbb{R}^{*+}$, orbit $O(ru_3^*)$ of ru_3^* for the coadjoint action of $g \in G$ could be identified with the upper half sheet $x_3 > 0$ of :
- $$\left\{ \xi = x_1 u_1^* + x_2 u_2^* + x_3 u_3^* / -x_1^2 - x_2^2 + x_3^2 = r^2 \right\}$$

the two-sheet hyperboloid. The stabilizer of ru_3^* for the coadjoint action of G is torus :

$$K = \left\{ \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}, \theta \in \mathbb{R} \right\}$$

Coadjoint Orbit of $SU(1,1)$ and Souriau Moment Map

- The torus $K = \left\{ \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}, \theta \in \mathbb{R} \right\}$ induces rotations of the unit disk
- K leaves 0 invariant. The stabilizer for the origin 0 of unit disk is maximal compact subgroup K of $SU(1,1)$.
- B. Cahen has observed that $O(ru_3^*) \simeq G / K$ and is diffeomorphic to the unit disk $D = \{z \in \mathbb{C} / |z| < 1\}$
- Then by composition, the **moment map** is given by:

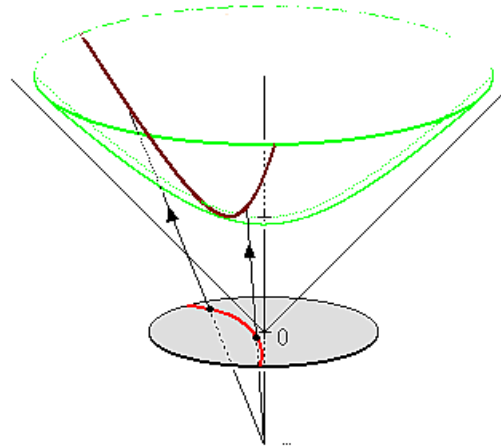
$$J : D \rightarrow O(ru_3^*)$$

$$z \mapsto J(z) = r \left(\frac{z + z^*}{(1 - |z|^2)} u_1^* + \frac{z - z^*}{i(1 - |z|^2)} u_2^* + \frac{1 + |z|^2}{(1 - |z|^2)} u_3^* \right)$$

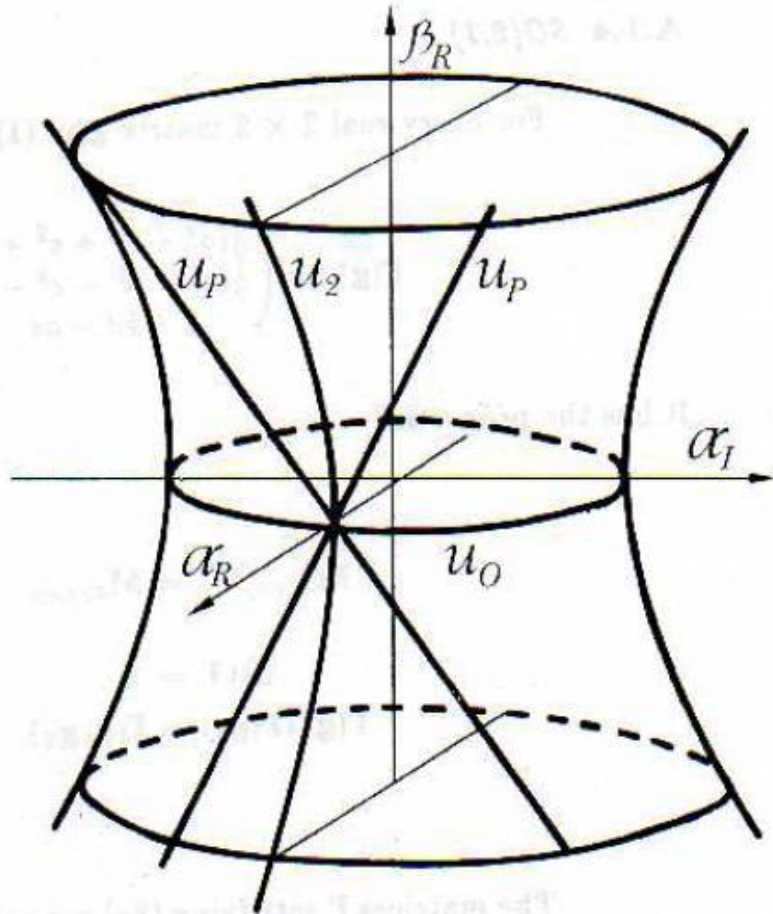
Coadjoint Orbit of $SU(1,1)$ and Souriau Moment Map

$$J(z) = r \left(\frac{z + z^*}{(1 - |z|^2)} u_1^* + \frac{z - z^*}{i(1 - |z|^2)} u_2^* + \frac{1 + |z|^2}{(1 - |z|^2)} u_3^* \right) \in \mathcal{O}(ru_3^*), z \in D$$

- J is linked to the natural action of G on D (by fractional linear transforms) but also the coadjoint action of G on $\mathcal{O}(ru_3^*) \simeq G/K$
- J^{-1} could be interpreted as the stereo-graphic projection from the two-sphere S^2 onto $\mathbb{C} \cup \infty$:



Group Manifold of $SU(1,1)$: one-sheeted equilateral hyperboloid



The $SU(1,1)$ group manifold is three-dimensional, connected and infinitely connected. It is R^3 pierced by a one-sheeted equilateral hyperboloid. Three representation of one-parameter subgroups are drawn:

- 1) Circle
- 2) One branch of hyperbola
- 3) Straight lines

Coadjoint Orbit of $SU(1,1)$ and Souriau Moment Map

➤ In case $r = \frac{n}{2}$ where $n \in \mathbb{N}^+, n \geq 2$ then the coadjoint orbit is given by

$$\mathbf{O}_n = \mathbf{O}(\xi_n)$$

➤ with $\xi_n = \frac{n}{2}u_3^* \in \mathfrak{g}^*$ with stabilizer of ξ_n for coadjoint action the torus

$$K = \left\{ \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}, \theta \in \mathbb{R} \right\}$$

➤ with Lie algebra $\mathbb{R}u_3$.

➤ $\mathbf{O}_n = \mathbf{O}(\xi_n)$ is associated with a holomorphic discrete series representation π_n of G by the KKS (Kirillov-Kostant-Souriau) method of orbits.

Coadjoint Orbit of $SU(1,1)$ and Souriau Moment Map

$$J : D \rightarrow \mathcal{O}_n$$

$$z \mapsto J(z) = \frac{n}{2} \left(\frac{z + z^*}{(1 - |z|^2)} u_1^* + \frac{z - z^*}{i(1 - |z|^2)} u_2^* + \frac{1 + |z|^2}{(1 - |z|^2)} u_3^* \right)$$

- Group G act on D by homography: $g.z = \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix}.z = \frac{az + b}{a^*z + b^*}$
- This action corresponds with coadjoint action of G on \mathcal{O}_n .

- The Kirillov-Kostant-Souriau 2-form of \mathcal{O}_n is given by:

$$\Omega_n(\zeta)(X(\zeta), Y(\zeta)) = \langle \zeta, [X, Y]^n \rangle, \quad X, Y \in \mathfrak{g} \text{ and } \zeta \in \mathcal{O}_n$$

- and is associated in the frame by ψ_n with:

$$\{f, g\} = i \left(1 - |z|^2\right)^2 \left(\frac{\partial f}{\partial z} \frac{\partial g}{\partial z^*} - \frac{\partial f}{\partial z^*} \frac{\partial g}{\partial z} \right) \quad \omega_n = \frac{in}{\left(1 - |z|^2\right)^2} dz \wedge dz^*$$

Coadjoint Orbit of $SU(1,1)$ and Souriau Moment Map

- It has been also observed that there are 3 basic observables generating the $SU(1,1)$ symmetry on classical level:

$$\left\{ \begin{array}{l} D \rightarrow \mathbb{R} \\ z \mapsto k_0(z) = \frac{1+|z|^2}{1-|z|^2} \end{array} \right\}, \left\{ \begin{array}{l} D \rightarrow \mathbb{R} \\ z \mapsto k_1(z) = \frac{1}{i} \frac{z-z^*}{1-|z|^2} \end{array} \right\}, \left\{ \begin{array}{l} D \rightarrow \mathbb{R} \\ z \mapsto k_2(z) = \frac{z+z^*}{1-|z|^2} \end{array} \right\}$$

with the Poisson commutation rule $\{k_0, k_1\} = k_2, \{k_0, k_2\} = -k_1, \{k_1, k_2\} = -k_0$

- (k_0, k_1, k_2) vector points to the upper sheet of the two-sheeted hyperboloid in \mathbb{R}^3 given by $k_0^2 - k_1^2 - k_2^2 = 1$, whose the stereographic projection onto the open unit disk is:

$$\left\{ \begin{array}{l} (k_0, k_1, k_2) \in H^+ \rightarrow D \\ z = \frac{k_2 + ik_1}{1 + k_0} = \sqrt{\frac{k_0 - 1}{k_0 + 1}} e^{i \arg z} \end{array} \right.$$

Coadjoint Orbit of $SU(1,1)$ and Souriau Moment Map

► Under the action of :

$$g \in G = SU(1,1) = \left\{ \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} / |a|^2 - |b|^2 = 1, a, b \in \mathbb{C} \right\}$$

$$\begin{pmatrix} k_- & k_0 \\ k_0 & k_+ \end{pmatrix} = \begin{pmatrix} k_2 + ik_1 & k_0 \\ k_0 & k_2 - ik_1 \end{pmatrix} = \frac{1}{1 - |z|^2} \begin{pmatrix} 2z & 1 + |z|^2 \\ 1 + |z|^2 & 2z^* \end{pmatrix}$$

is transform in:

$$\begin{pmatrix} k'_- & k'_0 \\ k'_0 & k'_+ \end{pmatrix} = \begin{pmatrix} k_- (g^{-1} \cdot z) & k_0 (g^{-1} \cdot z) \\ k_0 (g^{-1} \cdot z) & k_+ (g^{-1} \cdot z) \end{pmatrix} = g^{-1} \begin{pmatrix} k_- & k_0 \\ k_0 & k_+ \end{pmatrix} (g^{-1})^t$$

► This transform can be viewed as the co-adjoint action of $SU(1,1)$ on the coadjoint orbit identified with one-sheet hyperboloid $k_0^2 - k_1^2 - k_2^2 = 1$

Fourier/Laplace Transform and Representation Theory

- Fourier analysis, named after Joseph Fourier, who showed that representing a function as a sum of trigonometric functions greatly simplifies the study of heat transfer and addresses classically **commutative harmonic analysis**. Classical commutative harmonic analysis is restricted to functions defined on a topological locally compact and Abelian group G (Fourier series when $G = \mathbb{R}^n/\mathbb{Z}^n$, Fourier transform when $G = \mathbb{R}^n$, discrete Fourier transform when G is a finite Abelian group).
- The modern development of Fourier analysis during XXth century has explored the **generalization of Fourier and Fourier-Plancherel formula for non-commutative harmonic analysis**, applied to locally compact non-Abelian groups.
- This has been solved by geometric approaches based on “orbits methods” (Fourier-Plancherel formula for G is given by coadjoint representation of G in dual vector space of its Lie algebra) with many contributors (Dixmier, Kirillov, Bernat, Arnold, Berezin, Kostant, Souriau, Duflo, Guichardet, Torasso, Vergne, Paradan, etc.)

Fourier/Laplace Transform and Representation Theory



Editorial

Joseph Fourier 250th Birthday: Modern Fourier Analysis and Fourier Heat Equation in Information Sciences for the XXIst Century

Frédéric Barbaresco ^{1,*} and Jean-Pierre Gazeau ²

- ¹ Key Technology Domain PCC (Processing, Control & Cognition) Representative, Thales Land & Air Systems, Voie Pierre-Gilles de Gennes, F91470 Limours, France
 - ² APC (UMR 7164), Department of Physics, Université Paris-Diderot, F75205 Paris, France; gazeau@apc.in2p3.fr
- * Correspondence: frederic.barbaresco@thalesgroup.com; Tel: +1-501-450-5839

Received: 19 February 2019; Accepted: 27 February 2019; Published: 6 March 2019



Abstract: For the 250th birthday of Joseph Fourier, born in 1768 at Auxerre in France, this MDPI special issue will explore modern topics related to Fourier analysis and Fourier Heat Equation. Fourier analysis, named after Joseph Fourier, addresses classically commutative harmonic analysis. The modern development of Fourier analysis during XXth century has explored the generalization of Fourier and Fourier-Plancherel formula for non-commutative harmonic analysis, applied to locally compact non-Abelian groups. In parallel, the theory of coherent states and wavelets has been generalized over Lie groups (by associating coherent states to group representations that are square integrable over a homogeneous space). The name of Joseph Fourier is also inseparable from the study of mathematics of heat. Modern research on Heat equation explores geometric extension of classical diffusion equation on Riemannian, sub-Riemannian manifolds, and Lie groups. The heat equation for a general volume form that not necessarily coincides with the Riemannian one is useful in sub-Riemannian geometry, where a canonical volume only exists in certain cases. A new geometric theory of heat is emerging by applying geometric mechanics tools extended for statistical mechanics, for example, the Lie groups thermodynamics.

Keywords: harmonic analysis on abstract space; heat equation on manifolds and Lie Groups



Jean-Baptiste Joseph Fourier
(1768–1830)

OPEN

https://www.mdpi.com/journal/entropy/special_issues/fourier

THALES

Covariant Gibbs Density by Souriau Thermodynamics

- We can use Kirillov representation theory and his character formula to compute Souriau covariant Gibbs density in the unit Poincaré disk.
- For any Lie group G , a coadjoint orbit $\mathbf{O} \subset \mathfrak{g}^*$ has a canonical symplectic form given by KKS 2-form $\omega_{\mathbf{O}}$.
- If \mathfrak{g} is finite dimensional, the corresponding volume element defines a G -invariant measure supported on \mathbf{O} , which can be interpreted as a tempered distribution.
- The Fourier transform :
$$\mathfrak{I}(x) = \int_{\mathbf{O} \subset \mathfrak{g}^*} e^{-i\langle x, \lambda \rangle} \frac{1}{d!} d\omega^d \quad \text{with } \lambda \in \mathfrak{g}^* \text{ and } x \in \mathfrak{g}$$
- is Ad G -invariant. When $\mathbf{O} \subset \mathfrak{g}^*$ is an integral coadjoint orbit, Kirillov has proved that this Fourier transform is related to Kirillov character $\chi_{\mathbf{O}}$ by:

$$\mathfrak{I}(x) = j(x) \chi_{\mathbf{O}}(e^x) \quad \text{where } j(x) = \det^{1/2} \left(\frac{\sinh(ad(x/2))}{ad(x/2)} \right)$$

Covariant Gibbs Density by Souriau Thermodynamics

➤ \mathcal{X}_0 is called Kirillov character of a unitary representation associated to the orbit.

We will consider the universal covering of $PSU(1,1)$, the Lie algebra is:

$$\mathfrak{g}^* = \mathfrak{su}(1,1)^* = \left\{ \begin{pmatrix} iE & p^* \\ p & -iE \end{pmatrix} / E \in \mathbb{R}, p \in \mathbb{C} \right\}$$

➤ the Ad-invariant form $m^2 = E^2 - |p|^2$ allows to identify the following operator Ad and Ad^* , m could be considered analogously as rest mass, E as energy, and $p = p_1 + ip_2$ as the momentum vector.

➤ The coadjoint orbits are the rest mass shells. Let $D = \{w \in \mathbb{C} / |w| < 1\}$ Poincaré unit disk, for any $m > 0$, there is a corresponding action of the universal covering of $PSU(1,1)$ on $\kappa^{m/2}$ (with κ the holomorphic cotangent bundle of unit disk), with the invariant symplectic form:

$$\omega = \text{curv}(\kappa) = -i\partial\bar{\partial}^* \log |dw|^2 = 2i \frac{dw \wedge dw^*}{(1 - |w|^2)^2}$$

Covariant Gibbs Density by Souriau Thermodynamics

➤ The moment map is equivariant isomorphism (\mathbf{O}_m^+ coadjoint orbit for $m^2 > 0, E > 0$)

$$J : w \in \left(D, \text{curv} \left(\kappa^{m/2} \right) \right) \mapsto (p, E) = \frac{m}{(1 - |w|^2)} \left(2iw, 1 + |w|^2 \right) \in \mathbf{O}_m^+$$

➤ In case $m > 1$, the Kirillov character formula is given by:

$$\chi_m \left(\exp \left(\begin{pmatrix} x & \cdot \\ \cdot & -x \end{pmatrix} \right) \right) = j(x)^{-1} \int_{\mathbf{O}_{m-1}^+} e^{-i \left\langle \begin{pmatrix} x & \cdot \\ \cdot & -x \end{pmatrix}, \begin{pmatrix} iE & p^* \\ p & -iE \end{pmatrix} \right\rangle} \omega_{\mathbf{O}_{m-1}^+}$$

where

$$j(x) = \det^{1/2} \left[\sinh \left(\text{ad} \begin{pmatrix} x/2 & \\ & -x/2 \end{pmatrix} \right) / \text{ad} \begin{pmatrix} x/2 & \\ & -x/2 \end{pmatrix} \right] = \frac{\sinh(x)}{x}$$

which reduces to :

$$\frac{e^{mx}}{1 - e^{2x}} j(x) = \int_D e^{(m-1)x \frac{1+|w|^2}{1-|w|^2}} \frac{1}{(1 - |w|^2)^2} dw \wedge dw^*$$

Covariant Gibbs Density by Souriau Thermodynamics

➤ Souriau-Gibbs covariant density is given by:

$$P_{Gibbs}(w) = \frac{e^{-\left\langle \begin{pmatrix} ix & -\eta \\ -\eta^* & -ix \end{pmatrix}, \begin{pmatrix} im\frac{1+|w|^2}{1-|w|^2} & 2m\frac{w}{1-|w|^2} \\ 2m\frac{w}{1-|w|^2} & -im\frac{1+|w|^2}{1-|w|^2} \end{pmatrix} \right\rangle}}{j(x)\chi_m\left(e^{\begin{pmatrix} x & i\eta \\ i\eta^* & -x \end{pmatrix}}\right)} = \frac{e^{2m\left(x\frac{1+|w|^2}{1-|w|^2} + \frac{w(\eta+\eta^*)}{1-|w|^2}\right)}}{j(x)\chi_m\left(e^{\begin{pmatrix} x & i\eta \\ i\eta^* & -x \end{pmatrix}}\right)}$$

Bargmann parameterization of $SU(1,1)$

➤ $SU(1,1)$ is isomorphic to $SL(2, \mathbb{R}) = Sp(2, \mathbb{R})$ through the complex unitary matrix W :

$$SL(2, \mathbb{R}) = \left\{ g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} / \det g = ad - bc = 1 \right\}$$

$$Sp(2, \mathbb{R}) = \left\{ g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} / gJg^T = J, J = \begin{pmatrix} 0 & +1 \\ -1 & 0 \end{pmatrix} \right\}$$

$$W = \frac{1}{\sqrt{2}} \begin{pmatrix} \omega^{-1} & \omega^{-1} \\ -\omega & \omega \end{pmatrix} = (W^+)^{-1} \quad \text{with} \quad \omega = e^{i\pi/4} = \frac{1}{\sqrt{2}}(1+i)$$

Bargmann parameterization of $SU(1,1)$

► If we observe that $W^{-1}JW = -iM$, the isomorphism is given explicitly by:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = g(u) = WuW^{-1} = \begin{pmatrix} \operatorname{Re}(\alpha + \beta) & -\operatorname{Im}(\alpha - \beta) \\ \operatorname{Im}(\alpha + \beta) & \operatorname{Re}(\alpha - \beta) \end{pmatrix}$$
$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} = u(g) = W^{-1}gW = \frac{1}{2} \begin{pmatrix} (a+d) - i(b-c) & (a-d) + i(b+c) \\ (a-d) - i(b+c) & (a+d) + i(b-c) \end{pmatrix}$$

$$W = \frac{1}{\sqrt{2}} \begin{pmatrix} \omega^{-1} & \omega^{-1} \\ -\omega & \omega \end{pmatrix} = (W^+)^{-1} \quad \text{with} \quad \omega = e^{i\pi/4} = \frac{1}{\sqrt{2}}(1+i)$$

Bargmann parameterization of $SU(1,1)$

► We can also make also a link with $SO(2,1)$ of “1+2” pseudo-orthogonal matrices:

$$SO(2,1) = \left\{ \Gamma \in GL(3,3) / \det(\Gamma) = 1, \Gamma K \Gamma^T = \Gamma, K = \begin{pmatrix} +1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \right\}$$

$$\Gamma(g) = \begin{pmatrix} \frac{1}{2}(a^2 + b^2 + c^2 + d^2) & \frac{1}{2}(a^2 - b^2 + c^2 - d^2) & -cd - ab \\ \frac{1}{2}(a^2 + b^2 - c^2 - d^2) & \frac{1}{2}(a^2 - b^2 - c^2 + d^2) & cd - ab \\ -bd - ac & bd - ac & ad + bc \end{pmatrix}$$

$$\Gamma(g_1)\Gamma(g_2) = \Gamma(g_1g_2), \Gamma(I) = I, \Gamma(g^{-1}) = \Gamma(g)^{-1}$$

Bargmann parameterization of $SU(1,1)$

► The matrix $SO(2,1)$ corresponds to any $SU(1,1)$:

$$\Gamma(u) = \begin{pmatrix} |\alpha|^2 + |\beta|^2 & 2\operatorname{Re} \alpha\beta^* & 2\operatorname{Im} \alpha\beta^* \\ 2\operatorname{Re} \alpha\beta & \operatorname{Re}(\alpha^2 + \beta^2) & \operatorname{Im}(\alpha^2 - \beta^2) \\ -2\operatorname{Im} \alpha\beta & -\operatorname{Im}(\alpha^2 + \beta^2) & \operatorname{Re}(\alpha^2 - \beta^2) \end{pmatrix}$$

$$\alpha = \pm \sqrt{\frac{1}{2}(\Gamma_{11} + \Gamma_{12}) + i(\Gamma_{12} - \Gamma_{21})}, \quad \beta = \frac{1}{2\alpha}(\Gamma_{10} - i\Gamma_{20})$$

Bargmann parameterization of $SU(1,1)$

➤ The properties of connectivity of $Sp(2, \mathbb{R})$ is described by its isomorphy with $SU(1,1)$

➤ Using unimodular condition:

$$|\alpha|^2 - |\beta|^2 = 1 \Rightarrow \alpha_R^2 + \alpha_I^2 - \beta_R^2 = 1 + \beta_I^2 \geq 1$$

with $\alpha = \alpha_R + i\alpha_I$ and $\beta = \beta_R + i\beta_I$

➤ If β_I is fixed, $(\alpha_R, \alpha_I, \beta_R)$ are constrained to define a one-sheeted revolution hyperboloid, with its circular waist in the α plane.

Bargmann parameterization of $SU(1,1)$

- To $SU(1,1)$, we can associate the simply-connected universal covering group, using the maximal compact subgroup $U(1)$ and corresponding to the Iwasawa decomposition (factorization of a noncompact semisimple group into its maximal compact subgroup times a solvable subgroup).

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} = \begin{pmatrix} e^{i\omega} & 0 \\ 0 & e^{i\omega} \end{pmatrix} \begin{pmatrix} \lambda & \mu \\ \mu^* & \lambda \end{pmatrix} \text{ with } \begin{cases} \omega = \arg \alpha = \frac{1}{2} i \ln(\alpha^* \alpha^{-1}) \\ \lambda = |\alpha| > 0 \\ \mu = e^{-i\omega} \beta = \sqrt{\frac{\alpha^*}{\alpha}} \beta \end{cases}$$

$$\beta = e^{i\omega} \mu, |\alpha|^2 - |\beta|^2 = \lambda^2 - |\mu|^2 = 1 \text{ so } |\mu| < \lambda$$

Bargmann parameterization of $SU(1,1)$

➤ Bargmann has generalized this parameterization for $Sp(2N, \mathbb{R})$, more convenient but difficult to generalize to N dimensions.

➤ For $SU(1,1)$, Bargmann has used (ω, γ) :

$$\gamma = \frac{\mu}{\lambda} = \frac{\beta}{\alpha} \quad (|\gamma| < 1), \quad \lambda = \frac{1}{\sqrt{1-|\gamma|^2}}, \quad \mu = \frac{\gamma}{\sqrt{1-|\gamma|^2}}$$

➤ For $SL(2, \mathbb{R}) = Sp(2, \mathbb{R})$, the Bargman, parameterization is given by this decomposition of a non-singular matrix into the product of an orthogonal and a positive definite symmetric matrix:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \cos \omega & -\sin \omega \\ \sin \omega & \cos \omega \end{pmatrix} \begin{pmatrix} \lambda + \operatorname{Re} \mu & \operatorname{Im} \mu \\ \operatorname{Im} \mu & \lambda - \operatorname{Re} \mu \end{pmatrix}$$

$$\omega = \arg[(a+d) - i(b-c)], \quad \mu = e^{-i\omega} [(a-d) + i(b+c)]$$

➤ $SU(1,1)$ and $SL(2, \mathbb{R}) = Sp(2, \mathbb{R})$ are described when ω is counted modulo 2π

$$\omega \equiv \omega \pmod{2\pi}$$

Same approach for other Lie Groups



Group Everywhere, also in Information Geometry

➤ Information Geometry (gaussian densities)

$$\begin{bmatrix} Y \\ 1 \end{bmatrix} = \begin{bmatrix} R^{1/2} & m \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ 1 \end{bmatrix}, \quad \begin{cases} R^{1/2} \in T_n^+ \\ (R^{1/2} : \text{Cholesky de } R) \\ m \in R^n \end{cases}$$

+
Affine Group for
Exponential
Families

➤ Lie Group Thermodynamics (Galileo Group)

$$\begin{bmatrix} \vec{x}' \\ t' \\ 1 \end{bmatrix} = \begin{bmatrix} R & \vec{u} & \vec{w} \\ 0 & 1 & e \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \vec{x} \\ t \\ 1 \end{bmatrix}, \quad R \in SO(3), \vec{u}, \vec{w} \in R^3, e \in R$$

+
Poincaré Group in
Relativity

➤ Mobile object/ Robotics (Invariant Kalman Filter)

$$\begin{bmatrix} Z' \\ 1 \end{bmatrix} = \begin{bmatrix} \Omega & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Z \\ 1 \end{bmatrix}, \quad \begin{cases} \Omega \in SO(3) \\ t \in R^3 \end{cases}$$

+
SIM(3) Group
In Robotics

➤ Homogeneous Bounded Domains (Poincaré Half Plane)

$$\begin{bmatrix} X' \\ 1 \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ 1 \end{bmatrix}, \quad a \in R_+^* \text{ et } b \in R$$

SE(3) Lie Group

➤ SE(3) Lie Group

$$Z' = RZ + t \Rightarrow \begin{bmatrix} Z' \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Z \\ 1 \end{bmatrix}, \quad \begin{cases} R \in SO(3) \\ t \in \mathbb{R}^3 \end{cases} \Rightarrow \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \in SE(3), \quad G = SE(3)$$

➤ SE(3): se(3) Lie Algebra

$$\begin{bmatrix} j(\alpha) & \gamma \\ 0 & 0 \end{bmatrix} \in se(3), \quad Z = (j(\alpha), \gamma) \in \mathbb{R}^6, \quad se(3) \in \mathfrak{g} \cong \mathbb{R}^6$$

➤ Dual Lie Algebra

$$x = (l, p) \in \mathbb{R}^6, \quad x \in \mathfrak{g}^*$$

➤ Duality Bracket

$$\begin{cases} Z = (\alpha, \gamma) \in \mathfrak{g} \\ x = (l, p) \in \mathfrak{g}^* \end{cases} \Rightarrow \langle x, Z \rangle = \langle l, \alpha \rangle + \langle p, \gamma \rangle$$

SE(3) Lie Group

➤ Adjoint Operator

$$g \in SE(3) \Rightarrow g = \begin{pmatrix} I & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}$$

- 1st term:

$$\begin{aligned} \left\langle \begin{pmatrix} I & c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} l \\ p \end{pmatrix}, \begin{pmatrix} j(\alpha) & \gamma \\ 0 & 0 \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} l \\ p \end{pmatrix}, \begin{pmatrix} I & -c \\ 0 & 1 \end{pmatrix} \begin{pmatrix} j(\alpha) & \gamma \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I & c \\ 0 & 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} l \\ p \end{pmatrix}, \begin{pmatrix} j(\alpha) & \alpha \times c + \gamma \\ 0 & 0 \end{pmatrix} \right\rangle \\ &= \langle l, \alpha \rangle + \langle p, \alpha \times c + \gamma \rangle = \langle l, \alpha \rangle + \langle c \times p, \alpha \rangle + \langle p, \gamma \rangle = \langle l + c \times p, \alpha \rangle + \langle p, \gamma \rangle \end{aligned}$$

- 2nd terme:

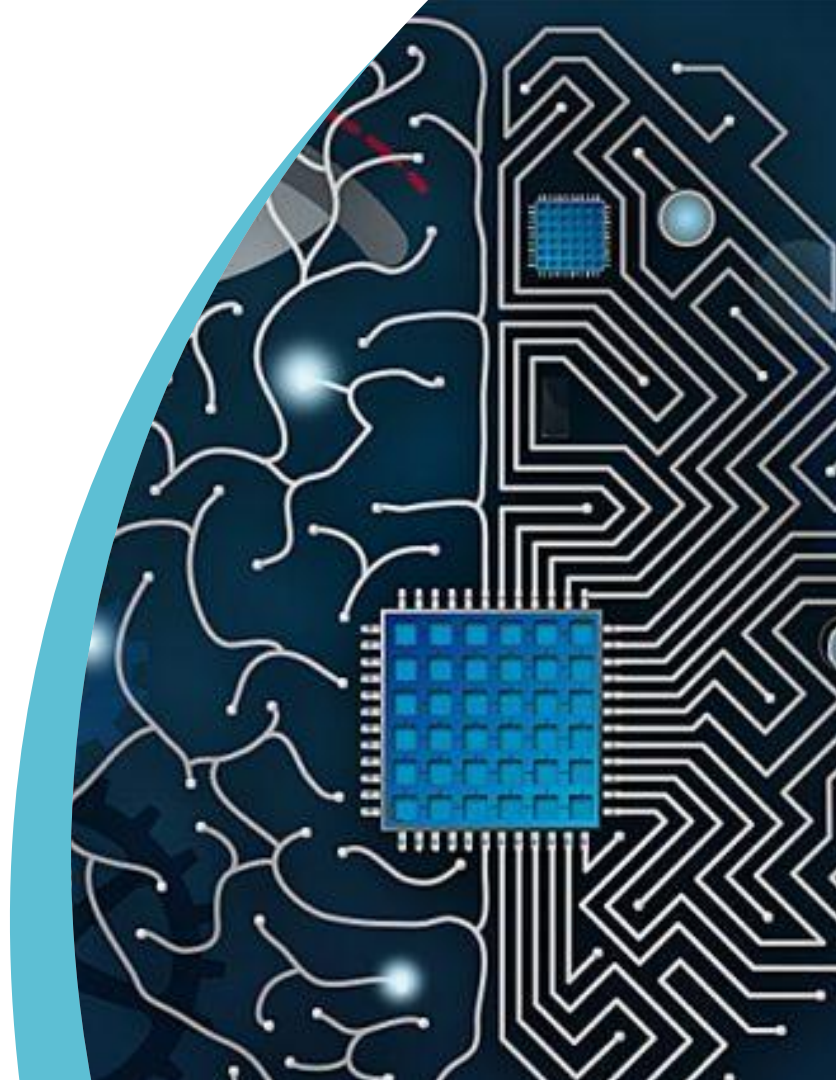
$$\begin{aligned} \left\langle \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} l \\ p \end{pmatrix}, \begin{pmatrix} j(\alpha) & \gamma \\ 0 & 0 \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} l \\ p \end{pmatrix}, \begin{pmatrix} A^T & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} j(\alpha) & \gamma \\ 0 & 0 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} l \\ p \end{pmatrix}, \begin{pmatrix} A^T j(\alpha) A & A^T \gamma \\ 0 & 0 \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} l \\ p \end{pmatrix}, \begin{pmatrix} j(A^T \alpha) & A^T \gamma \\ 0 & 0 \end{pmatrix} \right\rangle = \langle l, A^T \alpha \rangle + \langle p, A^T \gamma \rangle \\ &= \langle Al, \alpha \rangle + \langle Ap, \gamma \rangle \end{aligned}$$

➤ Adjoint operator:

$$(A, c) \cdot (l, p) = (Al + c \times Ap, Ap)$$



Main references



Main references

- Bargmann, V. : Irreducible unitary representations of the Lorentz group. Ann. Math. 48, pp.588-640, (1947).
- Souriau, J.-M. : Mécanique statistique, groupes de Lie et cosmologie, Colloques int. du CNRS numéro 237. In Proceedings of the Géométrie Symplectique et Physique Mathématique, Aix-en-Provence, France, 24–28, pp. 59–113, (1974)
- Souriau, J.-M. : Structure des systèmes dynamiques, Dunod, (1969).
- Kirillov, A.A. : Elements of the theory of representations, Springer-Verlag, Berlin, (1976).
- Marle, C.-M. : From Tools in Symplectic and Poisson Geometry to J.-M. Souriau's Theories of Statistical Mechanics and Thermodynamics. Entropy, 18, 370, (2016).
- Barbaresco, F. : Higher Order Geometric Theory of Information and Heat Based on Poly-Symplectic Geometry of Souriau Lie Groups Thermodynamics and Their Contextures: The Bedrock for Lie Group Machine Learning. Entropy, 20, 840, (2018).

Main references

- Cishahayo C., de Bièvre S. : On the contraction of the discrete series of $SU(1;1)$, Annales de l'institut Fourier, tome 43, no 2, p. 551-567, (1993).
- Cahen B. : Contraction de $SU(1,1)$ vers le groupe de Heisenberg, Travaux mathématiques, Fascicule XV, pp.19-43, (2004).
- Cahen, M., Gutt, S. and Rawnsley, J. : Quantization on Kähler manifolds I, Geometric in-terpretation of Berezin quantization, J. Geom. Phys. 7,45-62, (1990).
- Dai, J. : Conjugacy Classes, Coadjoint Orbits and Characters of $Diff+S^1$, PhD dissertation, The University of Arizona, Tucson, AZ, 85721, USA, (2000).
- Dai J., Pickrell D. : The orbit method and the Virasoro extension of $Diff+(S^1)$: I. Orbital integrals, Journal of Geometry and Physics, n°44, pp.623-653, (2003).
- Knapp, A. : Representation Theory of Semisimple Groups: An Overview based on Exam-ples, Princeton University press, (1986).
- Frenkel, I. : Orbital theory for affine Lie algebras, Invent. Math. 77, pp. 301–354, (1984).

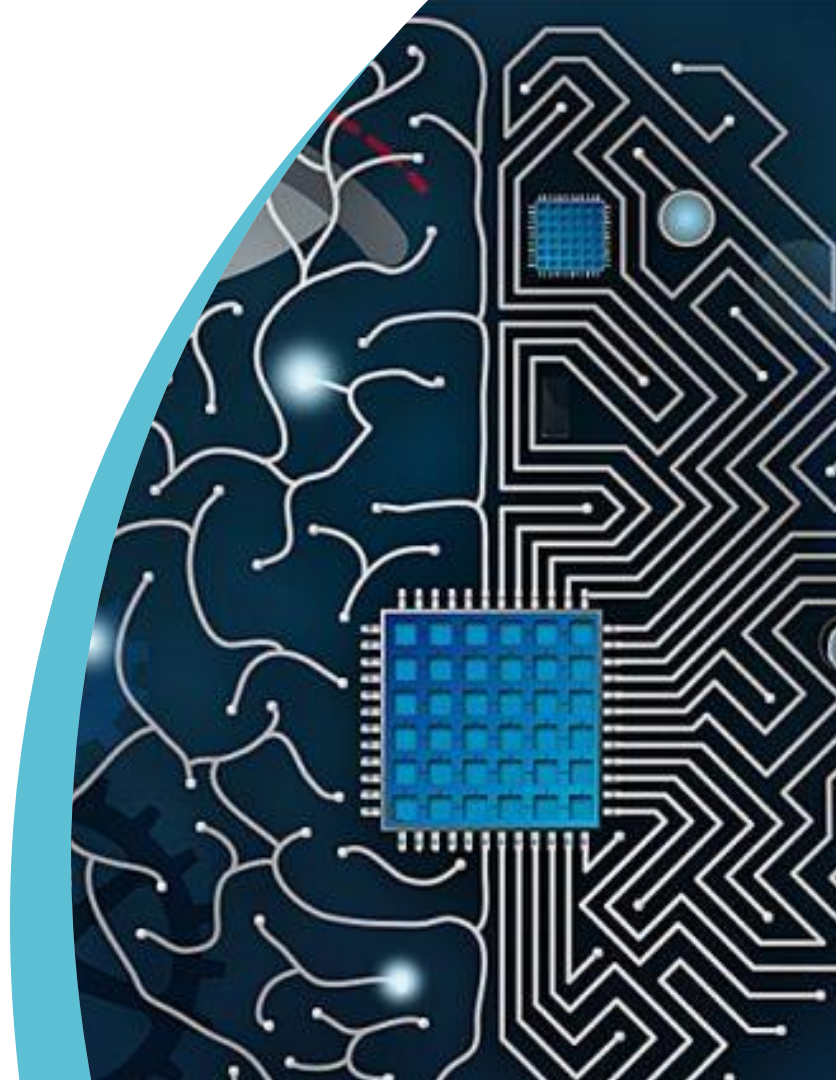
Main references

- Libine, M. : Introduction to Representations of Real Semisimple Lie Groups, arXiv:1212.2578v2, (2014).
- Guichardet, A. : La methode des orbites: historiques, principes, résultats. Leçons de ma-thématiques d'aujourd'hui, Vol.4, Cassini, pp. 33-59, (2010).
- Vergne, M. : Representations of Lie groups and the orbit method, Actes Coll. Bryn Mawr, p.59-101, Springer, (1983).
- Duflo, M. ; Heckman, G. ; Vergne, M.: Projection d'orbites, formule de Kirillov et formule de Blattner, Analyse harmonique sur les groupes de Lie et les espaces symétriques, Mé-moires de la Société Mathématique de France, Série 2, no. 15, p. 65-128, (1984).
- Witten, E: Coadjoint orbits of the Virasoro group, Com.. Math. Phys. 114, p. 1–53, (1988).
- Pukanszky, L. : The Plancherel formula for the universal covering group of $SL(2, \mathbb{R})$, Math. Ann. 156, pp.96-143, (1964).

Main references

- Clerc, J.L.; Orsted B.: The Maslov Index Revisited, Transformation Groups, vol. 6, n°4, pp.303-320, (2001).
- Foth, P.; Lamb M. : The Poisson Geometry of $SU(1,1)$, Journal of Mathematical Physics, Vol. 51, (2010).
- Perelomov, A.M. : Coherent States for Arbitrary Lie Group, Commun. math. Phys. 26, pp. 222-236, (1972).
- Ishi, H.: Kolodziejek, B: Characterization of the Riesz Exponential Family on Homogeneous Cones. arXiv:1605.03896, (2018).
- Tojo, K.; Yoshino, T. : A Method to Construct Exponential Families by Representation Theory. arXiv:1811.01394, (2018).

**Main Contribution of
Corps des Mines for
concept of Characteristic
Function:
Massieu, Poincaré, Levy &
Balian
(Le “bon sens” du Corps
des Mines)**



Elementary Structure of Information Geometry

Information Geometry Metric

$$g^* = d^2\Psi^* = d^2S$$

$$g = -d^2 \log \Phi = d^2\Psi$$

$$ds^2 = d^2 \text{ENTROPY}$$

$$ds^2 = -d^2 \text{LOG[LAPLACE]}$$

Legendre Transform

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

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

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

$$x^* = \frac{d\Psi(x)}{dx}, \quad x = \frac{d\Psi^*(x^*)}{dx^*}, \quad x^* = \int_{\Omega^*} \xi \cdot p_x(\xi) d\xi$$

Laplace/Fourier Transform

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

ENTROPY = LEGENDRE(- LOG[LAPLACE])

ENTROPY = FOURIER_(Min,*)(- LOG[FOURIER_(+,x)])

Seminal Work of François Jacques Dominique Massieu

Before introducing, Information Geometry with **Koszul model**, we have to explain the history of « **characteristic function** » that was initially introduced in Thermodynamics by (Corps des Mines Engineer) **François Jacques Dominique Massieu**. See:

- Roger Balian paper from French Academy of Sciences « **François Massieu et les potentiels thermodynamiques** »

http://www.academie-sciences.fr/pdf/hse/evol_Balian2.pdf

- Annale de l'Ecole des Mines: <http://www.annales.org/archives/x/massieu.html>



1st PhD on « sur les intégrales algébriques (algebraic integrals) » :

- Pour qu'il y ait une intégrale du premier degré dans le mouvement d'un point sur une surface, il faut et il suffit que cette surface soit développable sur une surface de révolution
- Pour qu'il y ait une intégrale du second degré dans le mouvement d'un point sur une surface, il faut et il suffit que cette surface ait son élément linéaire réductible à la forme de Liouville

Development of « Characteristic Function » Concept by Corps des Mines: Massieu, Poincaré, Levy & Balian

1869: François Massieu

- Introduction of Characteristic Function in Thermodynamics
- Use of Massieu Idea by Gibbs and Duhem to define Thermodynamics Potentials

$$S = \phi - \frac{1}{T} \cdot \frac{\partial \phi}{\partial \left(\frac{1}{T}\right)} = \phi - \beta \cdot U \quad (\text{Tr. Legendre})$$

S : Entropy, ϕ : Characteristic Function

« 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 »

F. Massieu

$$\phi = \log \psi$$

ϕ : characteristic function (from Massieu)

ψ : characteristic function (from Poincaré)

$$\left. \begin{aligned} S(\hat{D}) &= F(\hat{X}) - \langle \hat{D}, \hat{X} \rangle \\ F(\hat{X}) &= \log \text{Tr} \exp(\hat{X}) \end{aligned} \right\}$$

$$\Rightarrow ds^2 = -d^2 S = \text{Tr} \left[d\hat{D} d \log \hat{D} \right]$$

LES

whole or in
his reserved.



ished, trans
of Thales -



ent may not
losed to a



1908-1912: Henri Poincaré (+ Paul Levy)

- Poincaré introduces Characteristic Function in his 1908 Lecture on « Thermodynamics »
- Poincaré introduces Characteristic Function in his 1912 Lecture on « Probability »
- Paul Levy generalizes the Characteristic function in Probability

1986: Roger Balian

- Balian introduces the Fisher Quantum Metric as Hessian of von Neumann Entropy

■ Main papers of François Massieu

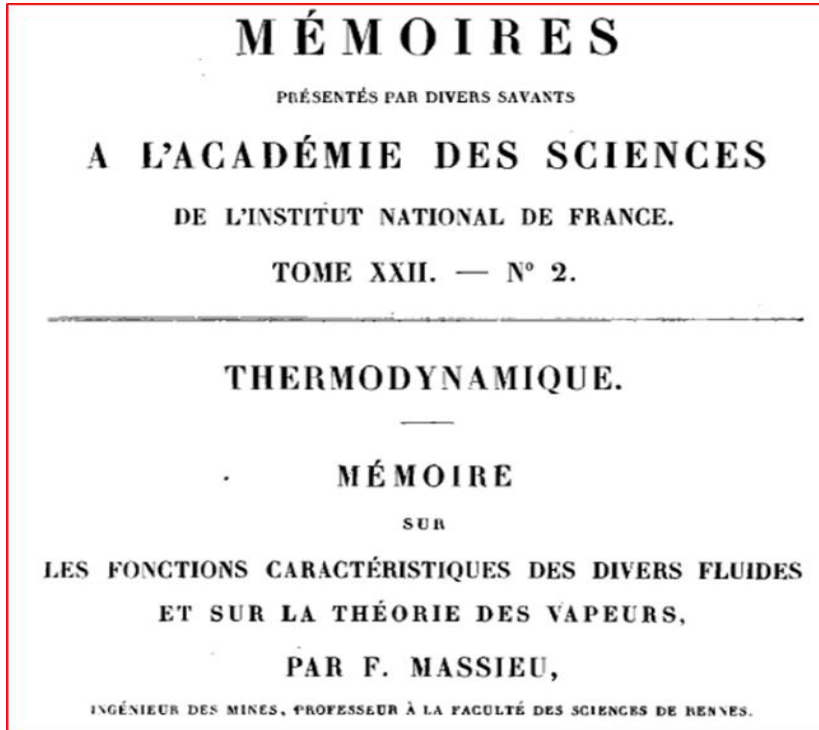
- Massieu, F. **Sur les Fonctions caractéristiques des divers fluides**. Comptes Rendus de l'Académie des Sciences 1869, 69, 858–862.
- Massieu, F. **Addition au précédent Mémoire sur les Fonctions caractéristiques**. Comptes Rendus de l'Académie des Sciences 1869, 69, 1057–1061.
- Massieu, F. **Exposé des principes fondamentaux de la théorie mécanique de la chaleur** (note destinée à servir d'introduction au Mémoire de l'auteur sur les fonctions caractéristiques des divers fluides et la théorie des vapeurs), 31 p., S.I. - s.n., 1873
- Massieu, F. **Thermodynamique: Mémoire sur les Fonctions Caractéristiques des Divers Fluides et sur la Théorie des Vapeurs**; Académie des Sciences: Paris, France, 1876; p. 92.

François Jacques Dominique Massieu : Initial paper on « Characteristic Function » in Rennes

Paper of François Massieu

$$S = \phi - \frac{1}{T} \cdot \frac{\partial \phi}{\partial \left(\frac{1}{T}\right)} = \phi - \beta \cdot U \quad (\text{Tr. Legendre})$$

S : Entropy, ϕ : Characteristic Function



MÉMOIRES PRÉSENTÉS.
THERMODYNAMIQUE. — *Addition au précédent Mémoire sur les fonctions caractéristiques.* Note de M. F. MASSIEU, présentée par M. Combes.

» Cette conclusion résultait *a posteriori* de la théorie même; mais j'ai reconnu qu'il était possible de l'établir de prime abord par un procédé qui a l'avantage de conduire plus simplement à la connaissance de la fonction caractéristique et de montrer la liaison de cette fonction avec d'autres fonctions déjà introduites dans la science, savoir : l'entropie S et l'énergie ou chaleur interne U . Je rappellerai d'ailleurs qu'une fois la fonction caractéristique d'un corps déterminée, la théorie thermodynamique de ce corps est faite.

$$\psi = S - \frac{U}{T}$$

Or, pour avoir S et U , et par suite ψ , il suffit de connaître quelles sont les quantités élémentaires de chaleur dQ qu'il faut fournir au corps suivant un cycle quelconque, pour le faire passer d'un état initial à un état déterminé, et en outre l'accroissement dU de sa chaleur interne pour les différents éléments de ce cycle, ou de tout autre cycle, reliant le même état initial au même état final.

Bad advice of Prof. Joseph Louis François Bertrand to Prof. Massieu

In following publications, François Massieu paper is reviewed by Joseph Louis François Bertrand, who give him a **bad advice** to replave variable $1/T$ by the variable T . If equations seem simpler, Structure support by Legendre transform is broken.



Joseph Louis
François Bertrand

⁽¹⁾ Dans le mémoire dont un extrait est inséré aux *Comptes rendus de l'Académie des sciences* du 18 octobre 1869, ainsi que dans la Note additionnelle insérée le 22 novembre suivant, j'avais adopté pour fonction caractéristique $\frac{H}{T}$, ou $S - \frac{U}{T}$; c'est d'après les bons conseils de M. Bertrand que j'y ai substitué la fonction H . dont l'emploi réalise quelques simplifications dans les formules.

Characteristic Fuction of Massieu and its good parameterization were discovered again by Max Planck (1897) and developed by Herbert Callen (1960) and Roger Balian.



M. Planck



H. Callen



R. Balian

Massieu Potential versus Gibbs-Duhem Potentials

GIBBS-DUHEM Potential: Free Energy

$$~~F = E - TS~~$$

Every mathematician knows it is impossible to understand any elementary course in thermodynamics.

Vladimir Arnold

MASSIEU Potential : characteristic function

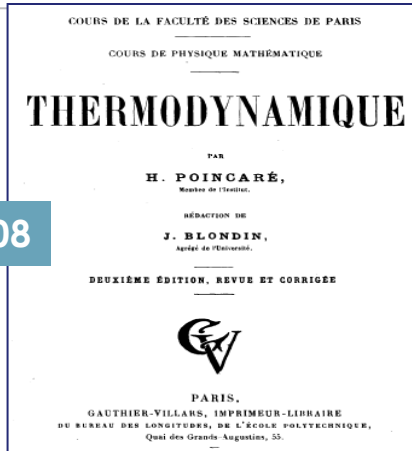
$$\frac{F}{T} = \frac{1}{T} E - S \quad \Rightarrow \quad \Psi = \left\langle \beta, E \right\rangle - S$$

$\beta = \frac{1}{T}$

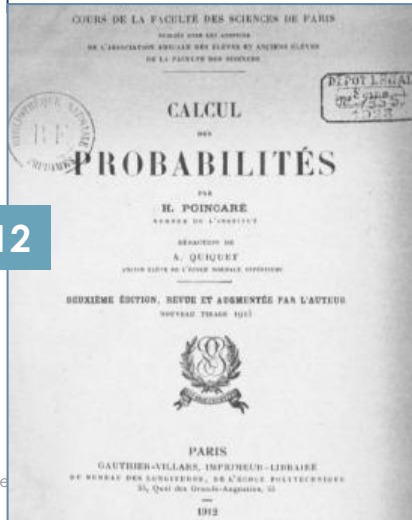
Preservation of Legendre Duality

Henri Poincaré Re-Use for Thermodynamics and Probability

1908



1912



125. Fonctions caractéristiques de M. Massieu. — Le théorème de Clausius nous a conduit à l'introduction d'une nouvelle fonction de l'état d'un système : son entropie S .

Si donc nous prenons comme variables indépendantes définissant l'état du système la pression p et le volume spécifique v , nous aurons à considérer, dans les applications, trois fonctions de ces variables : la température T , l'énergie interne U et l'entropie S .

M. Massieu a montré que, si l'on fait choix pour variables indépendantes de v et de T ou de p et de T , il existe une fonction, d'ailleurs inconnue, de laquelle les trois fonctions des variables, p , U et S dans le premier cas, v , U et S dans le second, peuvent se déduire facilement. M. Massieu a donné à cette fonction, dont la forme dépend du choix des variables, le nom de **fonction caractéristique.**

Puisque des fonctions de M. Massieu on peut déduire les autres fonctions des variables, toutes les équations de la Thermodynamique pourront s'écrire de manière à ne plus renfermer que ces fonctions et leurs dérivées; il en résultera donc, dans certains cas, une grande simplification. Nous verrons bientôt une application importante de ces fonctions.

Henri Poincaré Introduction of Characteristic Function in Probability

Characteristic Function in Probability

- Henri Poincaré introduced « characteristic function » in probability in his Lecture of 1912 (inspired by Massieu; both related by logarithm)
- It is **introduced with Laplace Transform**
- Characteristic function of a real random variable **defines completely its density of probability.**
- Moments of the random variable could be deduced from successive derivatives at zero of the characteristic function.
- The 2nd characteristic function is given by the logarithm,; **generating function of cumulants.**
- **Cumulants** have been introduced in 1889 by danish astronomer, mathematician and actuaire **Thorvald Nicolai Thiele** (1838 - 1910). Thiele called them **half-invariants** (demi-invariants).

Fonctions caractéristiques. — J'appelle fonction caractéristique $f(\alpha)$ la valeur probable de $e^{\alpha x}$; on aura donc

$$f(\alpha) = \sum p e^{\alpha x},$$

si la quantité x varie d'une manière discontinue et peut prendre seulement un nombre fini de valeurs, et

$$f(\alpha) = \int \varphi(x) e^{\alpha x} dx,$$

si x varie d'une manière continue et si $\varphi(x)$ représente la loi de probabilité. Il est clair que

$$f(\alpha) = 1 + \frac{\alpha}{1!} (x) + \frac{\alpha^2}{1.2} (x^2) + \frac{\alpha^3}{1.2.3} (x^3) + \dots,$$

(x^p) désignant la valeur probable de x^p . On voit que $f(0) = 1$.

La fonction caractéristique suffit pour définir la loi de probabilité. On a en effet par la formule de Fourier

$$f(i\alpha) = \int_{-\infty}^{+\infty} \varphi(x) e^{i\alpha x} dx,$$

$$2\pi \varphi(x) = \int_{-\infty}^{+\infty} f(i\alpha) e^{-i\alpha x} d\alpha.$$

Si deux quantités x et y sont indépendantes et si $f(\alpha)$, $f_1(\alpha)$ sont les fonctions caractéristiques correspondantes, la fonction relative à $x + y$ sera le produit $f(\alpha) f_1(\alpha)$. En effet, comme nous l'avons vu au paragraphe 130, la valeur probable du produit $e^{\alpha(x+y)}$ sera le produit des valeurs probables de $e^{\alpha x}$ et $e^{\alpha y}$.

Professor Roger Balian Contribution, 1986 paper

DISSIPATION IN MANY-BODY SYSTEMS: A GEOMETRIC APPROACH BASED ON INFORMATION THEORY

Roger BALIAN

*Service de Physique Théorique, Centre d'Etudes Nucléaires de Saclay, 91191 Gif-sur-Yvette Cedex,
France*

Yoram ALHASSID

A.W. Wright Nuclear Structure Lab., Yale University, New-Haven, Connecticut 06511, U.S.A.

and

Hugo REINHARDT

Zentralinstitut für Kernforschung, Rossendorf, 8051 Dresden, Pf 19, D.D.R.



NORTH-HOLLAND - AMSTERDAM

Roger Balian et al., A geometric approach to dissipation

Contents:

1. Incomplete description of a system	3	3. Geometry of the space of states and observables	24
1.1. Relevant and irrelevant variables	3	3.1. The Liouville representation of statistical mechanics	25
1.2. Relevant entropy relative to a subset of variables	5	3.2. Metric structure generated by the entropy	31
1.3. Geometric representation of complete, incomplete and reduced descriptions	8	3.3. Orthogonal projections	37
1.4. Outline	11	3.4. Geometry of macroscopic and of reduced states	41
2. Naive decoupling of the collective motion	15	4. Dynamical coupling between relevant and irrelevant variables	46
2.1. Derivation from a variational principle for states and observables	15	4.1. Projected equation of motion	46
2.2. Conservation laws; absence of dissipation	18	4.2. Memory kernel and generalized transport equations	49
2.3. Derivation of a time step procedure	19	4.3. Balance of the relevant entropy	53
2.4. Discarding the irrelevant information at finite time steps	21	4.4. Correlations in time: generalized Langevin equations	55
2.5. Dissipation as a memory loss	23	4.5. Projection referring to the full state	58
		4.6. Generalized fluctuation-dissipation theorems	64
		5. Approximation schemes	69
5.1. Decoupling and the mean-field approximation	70	7.1. The problem of selecting the relevant variables	122
5.2. Iteration procedure and linearization around mean field	71	7.2. Estimate of the coupling between relevant and irrelevant variables	124
5.3. Short memory approximation	73	7.3. Estimate of memory effects	125
5.4. Relaxation time approximation	76	7.4. Dissipation rate	127
5.5. Perturbation expansions	79	Appendix A: Properties of the relevant entropy	129
5.6. Projection near equilibrium	87	Appendix B: More on the Riemannian structure	131
6. Two illustrative examples	87	Appendix C: Various expressions for the dissipation rate	137
6.1. Brownian motion	87	Appendix D: Dissipation as correlation of random forces	141
6.2. Lipkin model	101	References	145
7. Conclusion: Possible criteria for improving the choice of collective variables	122		

Professor Roger Balian Contribution, 1986 paper

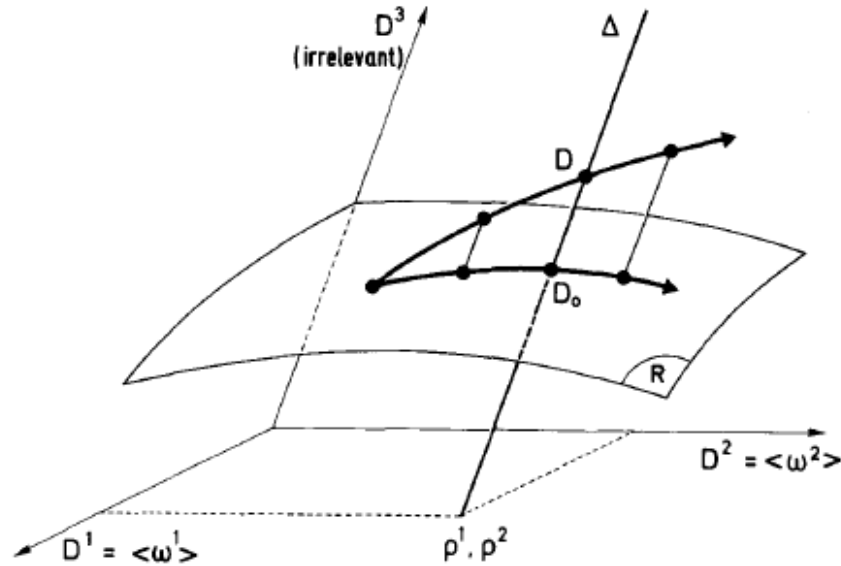


Fig. 1. Visualization of states in statistical mechanics. Three types of descriptions of a system should be distinguished. (1) A completely specified, *microscopic* state D is represented by a point, the coordinates D^μ of which are the expectation values of a complete set Ω^μ of linearly independent observables. The corresponding entropy $S(D)$ remains constant along a Hamiltonian trajectory. (2) A *macroscopic*, incomplete, description is characterized by the sole knowledge of the expectation values ρ^i of some set ω^i of relevant observables. This defines in the space of microscopic states a plane Δ (here a line), the dimensionality of which is the number of irrelevant variables. (3) The point D_0 of Δ which carries the minimum amount of information lies at the intersection of the plane Δ with the surface R of *reduced* states (having the same dimensionality as the relevant set). It provides for arbitrary observables the least biased predictions compatible with the macroscopic variables ρ . Its entropy $S(D_0)$ is identified with the macroscopic entropy $S(\rho)$ and may exhibit dissipation.

Entropy 2014, 16, 3878-3888; doi:10.3390/e16073878

OPEN ACCESS
entropy
ISSN 1099-4300
www.mdpi.com/journal/entropy

Article

The Entropy-Based Quantum Metric

Roger Balian

Institut de Physique Théorique, CEA/Saclay, F-91191 Gif-sur-Yvette Cedex, France;
E-Mail: roger@balian.fr

Received: 15 May 2014; in revised form: 25 June 2014 / Accepted: 11 July 2014 /

Published: 15 July 2014

Abstract: The von Neumann entropy $S(\hat{D})$ generates in the space of quantum density matrices \hat{D} the Riemannian metric $ds^2 = -d^2S(\hat{D})$, which is physically founded and which characterises the amount of quantum information lost by mixing \hat{D} and $\hat{D} + d\hat{D}$. A rich geometric structure is thereby implemented in quantum mechanics. It includes a canonical mapping between the spaces of states and of observables, which involves the Legendre transform of $S(\hat{D})$. The Kubo scalar product is recovered within the space of observables. Applications are given to equilibrium and non equilibrium quantum statistical mechanics. There the formalism is specialised to the relevant space of observables and to the associated reduced states issued from the maximum entropy criterion, which result from the exact states through an orthogonal projection. Von Neumann's entropy specialises into a relevant entropy. Comparison is made with other metrics. The Riemannian properties of the metric $ds^2 = -d^2S(\hat{D})$ are derived. The curvature arises from the non-Abelian nature of quantum mechanics; its general expression and its explicit form for q-bits are given, as well as geodesics.

Keywords: quantum entropy; metric; q-bit; information; geometry; geodesics; relevant entropy

Information, Entropy and Their Geometric Structures

Edited by

Frédéric Barbaresco and
Ali Mohammad-Djafari

Printed Edition of the Special Issue Published in *Entropy*

MDPI

www.mdpi.com/journal/entropy

- <https://www.mdpi.com/books/pdfview/book/127>
- <https://www.mdpi.com/books/pdfdownload/book/127>

Quantum Information Geometry of Roger Balian (1/3)

■ $Tr[\hat{D} \hat{O}]$ mean values through two dual spaces of observables \hat{O} and of the states \hat{D}

■ $S = -Tr[\hat{D} \log(\hat{D})]$ Entropy in space of states

■ Entropy S could be written as a scalar product $S = -\langle \hat{D}, \log(\hat{D}) \rangle$ where $\log(\hat{D})$ is an element of space of observables, allowing a physical geometrical structure in these spaces.

■ The 2nd differential $d^2 S$ is a non-negative quadratic form of coordinates of \hat{D} induced by the concavity of the Von Neumann Entropy S . Roger Balian has introduced distance ds between state \hat{D} and its neighborhood $\hat{D} + d\hat{D}$:

$$ds^2 = -d^2 S = Tr[d\hat{D}.d \log \hat{D}]$$

■ Where the Riemannian metric tensor is $-S(\hat{D})$ as function of a set of independant coordinates of \hat{D} .

Quantum Information Geometry of Roger Balian (2/3)

- It is possible to introduce the logarithm of a quantum characteristic function $F(\hat{X})$:

$$F(\hat{X}) = \log \text{Tr} \exp \hat{X}$$

- Von Neumann Entropy S appears as Legendre transform of $F(\hat{X})$:

$$S(\hat{D}) = F(\hat{X}) - \langle \hat{D}, \hat{X} \rangle$$

- with $S(\hat{D}) = -\text{Tr} \hat{D} \log \hat{D} = -\langle \hat{D}, \log \hat{D} \rangle$

- Where \hat{X} and \hat{D} are conjugate variable of the Legendre transform, making appear the algebraic/geometric duality between \hat{D} and $\log \hat{D}$.

- $F(\hat{X})$ characterizes canonical Thermodynamical equilibrium states with $\hat{X} = \beta \cdot \hat{H}$ and where hamiltonian is \hat{H} .

Quantum Information Geometry of Roger Balian (3/3)

■ $dF = Tr \hat{D} d\hat{X}$ with Maximum Entropy Gibbs Density:

$$\hat{D} = \frac{\exp \hat{X}}{Tr \exp \hat{X}}$$

■ dF are partial derivative of $F(\hat{X})$ with respect to coordinates of \hat{X} . \hat{D} is hermitian, normalised and positive and can be interpreted as a density matrix.

■ Legendre Transform appears with the following development:

$$S(\hat{D}) = -Tr \hat{D} \log \hat{D} = -Tr \left(\hat{D} (\hat{X} - \log Tr \exp \hat{X}) \right) = -Tr \hat{D} \hat{X} + Tr(\hat{D}) \log Tr \exp \hat{X}$$

$$Tr(\hat{D}) = 1 \Rightarrow S(\hat{D}) = F(\hat{X}) - \langle \hat{D}, \hat{X} \rangle$$

■ Roger Balian has defined the dual Riemannian metric from F , $ds^2 = d^2 F$ in conjugate space \hat{X} :

$$ds^2 = -dS^2 = Tr d\hat{D} d\hat{X} = d^2 F$$

■ Normalisation of \hat{D} implies $Tr d\hat{D} = 0$ and $Tr d^2 \hat{D} = 0$

Fundamental structure of Roger Balian Quantum Information Geometry

Legendre Transform

$$S(\hat{D}) = F(\hat{X}) - \langle \hat{D}, \hat{X} \rangle$$

Von Neumann Entropy

$$S = -\text{Tr}[\hat{D} \log(\hat{D})]$$

Characteristic Function

$$F(\hat{X}) = \log \text{Tr} \exp \hat{X}$$

Maximum Entropy Density

$$\hat{D} = \frac{\exp \hat{X}}{\text{Tr} \exp \hat{X}}$$

Balian Metric of Quantum Information Geometry (1986)

$$ds^2 = -d^2 S = \text{Tr} [d\hat{D} \cdot d \log \hat{D}]$$

Maurice Fréchet : IHP Lecture 1939, Paper 1943

in
ed.



- The Inverse of the Fisher/Information Matrix defines the lower bound of statistical estimators. Classically, this Lower bound is called Cramer-Rao Bound because it was described in the Rao's paper of 1945. Historically, this bound has been published first by Maurice Fréchet in 1939 in his winter "Mathematical Statistics" Lecture at the Institut Henri Poincaré during winter 1939–1940. Maurice Fréchet has published these elements in a paper as early as 1943. We can read at the bottom of the first page of his paper:

Fréchet, M. Sur l'extension de certaines évaluations statistiques au cas de petits échantillons. Revue de l'Institut International de Statistique 1943, 11, 182–205.

- At the bottom of 1st page of Fréchet's paper, we can read:
- The contents of this report formed a part of our lecture of mathematical statistics at the Henri Poincaré institute during winter 1939–1940. It constitutes one of the chapters of the second exercise book (in preparation) of our "Lessons of Mathematical Statistics", the first exercise book of which, "Introduction: preliminary Presentation of Probability theory" (119 pages quarto, typed) has just been published in the "Centre de Documentation Universitaire, Tournais et Constans. Paris".

This document may not be reproduced, modified, part or disclosed to a third party without the prior

Maurice Fréchet and Clairaut-Legendre Equation

■ Seminal work of Maurice Fréchet

- In Winter 1939, in his IHP Lecture, Maurice Fréchet introduced what has been called Cramer-Rao bound:

$$(\sigma_T)^2 \geq \frac{1}{n(\sigma_A)^2} \text{ avec } T = H(X_1, \dots, X_n), \quad A = \frac{1}{f(X, \theta)} \frac{\partial f(X, \theta)}{\partial \theta}$$

$\hat{\theta}$ estimateur de θ , borne de Fréchet : $R_{\theta} = E\left[(\theta - \hat{\theta})(\theta - \hat{\theta})^T\right] \geq I(\theta)^{-1}$

$$[I(\theta)]_{i,j} = -E\left[\frac{\partial^2 \log P(Z/\theta)}{\partial \theta_i \partial \theta_j}\right] = E\left[\frac{\partial \log P(Z/\theta)}{\partial \theta_i} \frac{\partial \log P(Z/\theta)}{\partial \theta_j}\right]$$

- In his 1943 paper, Maurice Fréchet studied “**densités distinguées**”, density with parameters that reach this bound. He proves that « distinguished density » should be exponential and given through **Alexis Clairaut equation**

(55)

$$\mu = \theta \mu' - \psi(\mu')$$

$$\Phi^*(x^*) = \left\langle x^*, \frac{d\Phi^*(x^*)}{dx^*} \right\rangle - \Phi\left(\frac{d\Phi^*(x^*)}{dx^*}\right) \text{ et } x^* = \frac{d\Phi(x)}{dx}$$

c'est-à-dire une équation de Clairaut. La solution $\mu' = \text{constante}$ réduirait $f(x, \theta)$, d'après (48) à une fonction indépendante de θ , cas où le problème n'aurait plus de sens. μ est donc donné par la solution singulière de (55), qui est unique et s'obtient en éliminant s entre $\mu = \theta s - \psi(s)$ et $\theta = \psi'(s)$ ou encore entre

Seminal Maurice Fréchet paper 1943

Fréchet, M. Sur l'extension de certaines évaluations statistiques au cas de petits échantillons. Revue de l'Institut International de Statistique 1943, 11, 182-205.

Etude des densités distinguées. Appelons (provisoirement, dans ce mémoire) *densité distinguée*, toute densité de probabilité $f(x, \theta)$ telle que la fonction

$$(46) \quad \theta + \frac{\frac{\partial L f(x, \theta)}{\partial \theta}}{\int_{-\infty}^{+\infty} \left[\frac{\partial}{\partial \theta} f(x, \theta) \right]^2 \frac{dx}{f(x, \theta)}}$$

soit indépendante de θ .

Pour ces densités distinguées, on va pouvoir déterminer la fonction minimale $H'(X_1, \dots, X_n)$ et étendre au cas des petits échantillons la comparaison des modes d'estimation faites par divers auteurs dans le cas des grands échantillons. Il vaut donc la peine de chercher la forme générale de $f(x, \theta)$ pour cette catégorie de variables.

de θ . En appelant $h(x)$ cette fonction, on voit qu'on a l'identité de la forme

$$(47) \quad \lambda(\theta) \frac{\partial}{\partial \theta} L f(x, \theta) = h(x) - \theta$$

où $\lambda(\theta) > 0$. On peut considérer $\frac{1}{\lambda(\theta)}$ comme la dérivée seconde d'une fonction $\mu(\theta)$; d'où $\frac{\partial}{\partial \theta} L f(x, \theta) = \mu_{\theta}''(\theta) [h(x) - \theta]$.

Par suite $L f(x, \theta) - \mu'_{\theta} [h(x) - \theta] - \mu(\theta)$ est une quantité indépendante de θ que nous pouvons représenter par $l(x)$.

Ainsi toute densité distinguée, $f(x, \theta)$, est de la forme

$$(48) \quad f(x, \theta) = e^{\mu'_{\theta} [h(x) - \theta] + \mu(\theta) + l(x)}$$

$$(52bis) \quad \lambda \mu'' = 1.$$

Incidentement, puisque, d'après (52), $\lambda(\theta)$ est positif, il en résulte aussi que $\mu'' \left(= \frac{1}{\lambda(\theta)} \right)$ est aussi positif. **Fisher metric**

On peut d'ailleurs préciser d'une manière plus directe que par (50), le choix des fonctions $\mu(\theta)$, $h(x)$, $l(x)$: on peut prendre arbitrairement $h(x)$ et $l(x)$ et alors $\mu(\theta)$ est déterminé par (50) ou même mieux par une formule explicite. En effet, (50) peut s'écrire

$$e^{\theta \mu' - \mu} = \int_{-\infty}^{+\infty} e^{\mu'_{\theta} h(x) + l(x)} dx.$$

Donnons-nous alors arbitrairement $h(x)$ et $l(x)$ et soit s une variable arbitraire: la fonction

$$\int_{-\infty}^{+\infty} e^{s h(x) + l(x)} dx \quad (1)$$

sera une fonction positive connue que nous pourrions représenter par $e^{\psi(s)}$. On voit alors que $\mu(\theta)$ sera défini par

$$\theta \mu' - \mu = \psi(\mu') \quad \text{ou}$$

$$(55) \quad \mu = \theta \mu' - \psi(\mu') \quad \text{Legendre-Clairaut}$$

c'est-à-dire une équation de Clairaut. La solution $\mu' = \text{constante}$ réduirait $f(x, \theta)$, d'après (48) à une fonction indépendante de θ , cas où le problème n'aurait plus de sens: μ est donc donné par la solution singulière de (55), qui est unique et s'obtient en éliminant s entre $\mu = \theta s - \psi'(s)$ et $\theta = \psi'(s)$ ou encore entre

$$(55bis) \quad \begin{aligned} e^{\theta \mu' - \mu} &= \int_{-\infty}^{+\infty} e^{s h(x) + l(x)} dx \text{ et} \\ \int_{-\infty}^{+\infty} e^{s h(x) + l(x)} [h(x) - \theta] dx &= 0. \end{aligned}$$

Si l'on veut, $\mu(\theta)$ est donné par la relation

$$e^{-\mu} = e^{-\theta s} \int_{-\infty}^{+\infty} e^{s h(x) + l(x)} dx$$

où s est donné en fonction de θ par la relation implicite (55bis).

Fréchet 1943 paper: Fréchet bound (Cramer Rao bound)

- We consider estimator of θ given by: $T = H(X_1, \dots, X_n)$
- And the random variable $A(X) = \frac{\partial \log p_\theta(X)}{\partial \theta}$ with associated value $U = \sum_i A(X_i)$
- The constraint $\int_{-\infty}^{+\infty} p_\theta(x) dx = 1$ induces:

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \prod_i p_\theta(x_i) dx_i = 1 \xRightarrow{\substack{\text{dérivée} \\ \text{par } \theta}} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \left[\sum_i A(x_i) \right] \prod_i p_\theta(x_i) dx_i = 0 \Rightarrow E_\theta[U] = 0$$

- If we assume $E_\theta[T] = \theta \Rightarrow \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} H(x_1, \dots, x_n) \prod_i p_\theta(x_i) dx_i = \theta \xRightarrow{\substack{\text{dérivée} \\ \text{par } \theta}} E[(T - \theta)U] = 1$

- Then as $E[T] = \theta$ and $E[U] = 0$, we have: $E[(T - E[T])(U - E[U])] = 1$

- According to Schwarz inequality: $[E(ZT)]^2 \leq E[Z^2]E[T^2]$

$$1 \leq E[(T - E[T])^2] E[(U - E[U])^2] = (\sigma_T \sigma_U)^2$$

- U as sum of independant variables, Bienaymé equality gives :

$$(\sigma_U)^2 = \sum_i [\sigma_{A(X_i)}]^2 = n(\sigma_A)^2 \Rightarrow (\sigma_T)^2 \geq \frac{1}{n(\sigma_A)^2}$$

Fréchet 1943 paper: Clairaut Equation

Fréchet studied « densités distinguées » (distinguished density): density with parameters that reach the Fréchet Bound

- Previous inequality becomes equality if there exist two numbers α et β (non random and non equal to zero) such that $\alpha(H' - \theta) + \beta U = 0$, with H' a function among admissible H function such that we have equality.

This equality could be written $H' = \theta + \lambda' U$ with λ' a deterministic value. Then if we use previous relation : $E[(T - E[T])(U - E[U])] = 1 \Rightarrow E[(H' - \theta)U] = \lambda' E_\theta[U^2] = 1$

- We deduce: $U = \sum A(X_i) \Rightarrow \lambda' n E_\theta[A^2] = 1$
- from wich we have $i \lambda'$ and the estimator H' that reach lower bound:

$$\lambda' = \frac{1}{nE[A^2]} \Rightarrow H' = \theta + \frac{1}{nE[A^2]} \sum_i \frac{\partial \log p_\theta(X_i)}{\partial \theta} = \theta + \frac{\sum_i \frac{\partial \log p_\theta(X_i)}{\partial \theta}}{n \int_{-\infty}^{+\infty} \left[\frac{\partial p_\theta(x)}{\partial \theta} \right]^2 \frac{dx}{p_\theta(x)}}$$

- with $E[H'] = \theta + \lambda' E[U] = \theta$

Fréchet 1943 paper: Clairaut Equation

- H' is then an admissible function, if H' is independent of θ . Indeed, if we consider $E_{\theta_0}[H'] = \theta_0$, $E[(H' - \theta_0)^2] \leq E_{\theta_0}[(H - \theta_0)^2] \forall H$ tq $E_{\theta_0}[H] = \theta_0$
- However $H = \theta_0$ verifies the equation and inequality proves that it is almost surely equal to θ_0 . So for estimating θ_0 , you should know first θ_0
- At this step, Fréchet studied « **densités distinguées** » (distinguished density), all density of probability $p_{\theta}(x)$ such that the following expression is independent of θ :

$$h(x) = \theta + \frac{\frac{\partial \log p_{\theta}(x)}{\partial \theta}}{\int_{-\infty}^{+\infty} \left[\frac{\partial p_{\theta}(x)}{\partial \theta} \right]^2 \frac{dx}{p_{\theta}(x)}}$$

- Fréchet objective is to determine the minimizing function $T = H'(X_1, \dots, X_n)$ that reach the bound. Previous equality could be written:

$$\lambda(\theta) \frac{\partial \log p_{\theta}(x)}{\partial \theta} = h(x) - \theta$$

$$\lambda(x) = \left[\int_{-\infty}^{+\infty} \left[\frac{\partial p_{\theta}(x)}{\partial \theta} \right]^2 \frac{dx}{p_{\theta}(x)} \right]^{-1} = \left[\int_{-\infty}^{+\infty} p_{\theta}(x) \left[\frac{\partial \ln p_{\theta}(x)}{\partial \theta} \right]^2 dx \right]^{-1}$$

Fréchet 1943 paper: Clairaut Equation

- However as $\lambda(\theta) > 0$, we can consider $\frac{1}{\lambda(\theta)}$ as second derivative of the function $\Phi(\theta)$ such that:

$$\frac{\partial \log p_\theta(x)}{\partial \theta} = \frac{\partial^2 \Phi(\theta)}{\partial \theta^2} [h(x) - \theta] \quad \frac{1}{\lambda(x)} = \int_{-\infty}^{+\infty} p_\theta(x) \left[\frac{\partial \ln p_\theta(x)}{\partial \theta} \right]^2 dx = E \left[\left[\frac{\partial \ln p_\theta(x)}{\partial \theta} \right]^2 \right] = \frac{\partial^2 \Phi(\theta)}{\partial \theta^2}$$

- we deduce that the following function is independant of θ :

$$\ell(x) = \log p_\theta(x) - \frac{\partial \Phi(\theta)}{\partial \theta} [h(x) - \theta] - \Phi(\theta)$$

- A **distinguished density** could be written:

$$p_\theta(x) = e^{\frac{\partial \Phi(\theta)}{\partial \theta} [h(x) - \theta] + \Phi(\theta) + \ell(x)} \quad \text{with} \quad \int_{-\infty}^{+\infty} p_\theta(x) dx = 1$$

- These 2 conditions are enough

Fréchet proved that « distinguished densities » are in the family of « exponential densities »
Fréchet proved that Fisher matrix is equal to the hessian of a function (Massieu Characteristic Function)

Fréchet 1943 paper: Clairaut Equation

- These 2 conditions are enough. Let 3 functions $\Phi(\theta)$, $h(x)$ and $\ell(x)$ such that we have for all θ :

$$\int_{-\infty}^{+\infty} e^{\frac{\partial \Phi(\theta)}{\partial \theta} [h(x) - \theta] + \Phi(\theta) + \ell(x)} dx = 1$$

- Then function is distinguished:

$$\theta + \frac{\frac{\partial \log p_{\theta}(x)}{\partial \theta}}{\int_{-\infty}^{+\infty} \left[\frac{\partial p_{\theta}(x)}{\partial \theta} \right]^2 \frac{dx}{p_{\theta}(x)}} = \theta + \lambda(x) \frac{\partial^2 \Phi(\theta)}{\partial \theta^2} [h(x) - \theta]$$

- if $\lambda(x) \frac{\partial^2 \Phi(\theta)}{\partial \theta^2} = 1$ when $\frac{1}{\lambda(x)} = \int_{-\infty}^{+\infty} \left[\frac{\partial \log p_{\theta}(x)}{\partial \theta} \right]^2 p_{\theta}(x) dx = (\sigma_A)^2$ the previous

function will reduce to $h(x)$ and then is not dependent to θ :

$$\theta + \lambda(x) \frac{\partial^2 \Phi(\theta)}{\partial \theta^2} [h(x) - \theta] \underset{\lambda(x) \frac{\partial^2 \Phi(\theta)}{\partial \theta^2} = 1}{\Rightarrow} h(x)$$

Fréchet 1943 paper: Clairaut Equation

- We the following relation:

$$\frac{1}{\lambda(x)} = \int_{-\infty}^{+\infty} \left(\frac{\partial^2 \Phi(\theta)}{\partial \theta^2} \right)^2 [h(x) - \theta]^2 e^{\frac{\partial \Phi(\theta)}{\partial \theta} (h(x) - \theta) + \Phi(\theta) + \ell(x)} dx$$

- The relation is true for all θ , we can derive previous expression with respect to θ

$$\int_{-\infty}^{+\infty} e^{\frac{\partial \Phi(\theta)}{\partial \theta} (h(x) - \theta) + \Phi(\theta) + \ell(x)} \left(\frac{\partial^2 \Phi(\theta)}{\partial \theta^2} \right) [h(x) - \theta] dx = 0$$

- We can divide by $\frac{\partial^2 \Phi(\theta)}{\partial \theta^2}$ that is not dependent to x . If we derive a 2nd time with respect to θ , we have:

$$\int_{-\infty}^{+\infty} e^{\frac{\partial \Phi(\theta)}{\partial \theta} (h(x) - \theta) + \Phi(\theta) + \ell(x)} \left(\frac{\partial^2 \Phi(\theta)}{\partial \theta^2} \right) [h(x) - \theta]^2 dx = \int_{-\infty}^{+\infty} e^{\frac{\partial \Phi(\theta)}{\partial \theta} (h(x) - \theta) + \Phi(\theta) + \ell(x)} dx = 1$$

- By combining this expression, with expression of $\frac{1}{\lambda(x)}$, we have : $\lambda(x) \cdot \frac{\partial^2 \Phi(\theta)}{\partial \theta^2} = 1$

- And as $\lambda(x) > 0$, then $\frac{\partial^2 \Phi(\theta)}{\partial \theta^2} > 0$

Fréchet 1943 paper: Clairaut Equation

- We can select arbitrary $h(x)$ and $l(x)$ and then $\Phi(\theta)$ is determined by:

$$\int_{-\infty}^{+\infty} e^{\frac{\partial\Phi(\theta)}{\partial\theta}[h(x)-\theta]+\Phi(\theta)+l(x)} dx = 1$$

- That we can write: $e^{\theta \cdot \frac{\partial\Phi(\theta)}{\partial\theta} - \Phi(\theta)} = \int_{-\infty}^{+\infty} e^{\frac{\partial\Phi(\theta)}{\partial\theta} h(x) + l(x)} dx$

- If we select arbitrary $h(x)$ and $l(x)$, and let s an arbitrary variable, the following function will be a positive function known, given by: $e^{\Psi(s)}$

$$\int_{-\infty}^{+\infty} e^{s \cdot h(x) + l(x)} dx = e^{\Psi(s)}$$

- We obtain $\Phi(\theta)$ by Clairaut Equation: $\Phi(\theta) = \theta \cdot \frac{\partial\Phi(\theta)}{\partial\theta} - \Psi\left(\frac{\partial\Phi(\theta)}{\partial\theta}\right)$

- The case $\frac{\partial\Phi(\theta)}{\partial\theta} = cste$ reduce the density to function independent to θ , then $\Phi(\theta)$ is given by singular solution of this Clairaut equation, that is unique and obtaining by eliminating s in the equation:

$$\Phi = \theta \cdot s - \Psi(s) \text{ et } \theta = \frac{\partial\Psi(s)}{\partial s}$$

Fréchet 1943 paper: Clairaut Equation

$$\Phi = \theta.s - \Psi(s) \text{ et } \theta = \frac{\partial \Psi(s)}{\partial s}$$

➤ The function $\Phi(\theta)$ is obtained by eliminating s between:

$$e^{\theta.s - \Phi(\theta)} = \int_{-\infty}^{+\infty} e^{s.h(x) + \ell(x)} dx \quad \int_{-\infty}^{+\infty} e^{s.h(x) + \ell(x)} [h(x) - \theta] dx = 0$$

➤ We obtain $\Phi(\theta) = -\log \int_{-\infty}^{+\infty} e^{s.h(x) + \ell(x)} dx + \theta.s$ where s is given by implicate equation:

$$\int_{-\infty}^{+\infty} e^{s.h(x) + \ell(x)} [h(x) - \theta] dx = 0$$

➤ Fréchet also observed that:

$$(\sigma_{T_n})^2 = \frac{1}{n(\sigma_A)^2} = \frac{1}{n \int_{-\infty}^{+\infty} \left[\frac{\partial p_\theta(x)}{\partial \theta} \right]^2 \frac{dx}{p_\theta(x)}} = \frac{1}{n \frac{\partial^2 \Phi(\theta)}{\partial \theta^2}}$$

➤ T_n follows:

$$p_\theta(t) = \sqrt{n} \frac{1}{\sigma_A \sqrt{2\pi}} e^{-\frac{n(t-\theta)^2}{2\sigma_A^2}} \text{ with } (\sigma_A)^2 = \frac{\partial^2 \Phi(\theta)}{\partial \theta^2}$$

Fréchet 1943 paper: Clairaut Equation

➤ We can write the estimator as :

$$H'(X_1, \dots, X_n) = \frac{1}{n} [h(X_1) + \dots + h(X_n)]$$

➤ Empirical value is given by :

$$t = H'(x_1, \dots, x_n) = \frac{1}{n} \sum_i h(x_i) = \theta + \lambda(\theta) \sum_i \frac{\partial \log p_\theta(x_i)}{\partial \theta}$$

➤ If we set $\theta = t$, we have as $\lambda(\theta) > 0$:

$$\sum_i \frac{\partial \log p_t(x_i)}{\partial t} = 0$$

➤ When $p_\theta(x)$ is a distinguished density, empirical value $\frac{\partial t}{\partial t}$ of θ corresponds at a sample x_1, \dots, x_n that is a root of previous equation in t . This equation has a root and only one root when X is a distinguished variable. With :

$$p_\theta(x) = e^{\frac{\partial \Phi(\theta)}{\partial \theta} [h(x) - \theta] + \Phi(\theta) + \ell(x)} \Rightarrow \sum_i \frac{\partial \log p_t(x_i)}{\partial t} = \frac{\partial^2 \Phi(t)}{\partial t^2} \left[\frac{\sum_i h(x_i)}{n} - t \right] \text{ avec } \frac{\partial^2 \Phi(t)}{\partial t^2} > 0$$

➤ We find the only root: $t = \frac{1}{n} \sum h(x_i)$

➤ $T \equiv H'(X_1, \dots, X_n) = \frac{1}{n} \sum_i h(X_i)$ cannot have an arbitrary expression, but it is given by a sum.

Basic Tool: Duality and Legendre Transform

Legendre Transform plays a central role related to duality & convexity

- Dual Potential Functions (entropy and characteristic function)
- Systems of dual coordinates

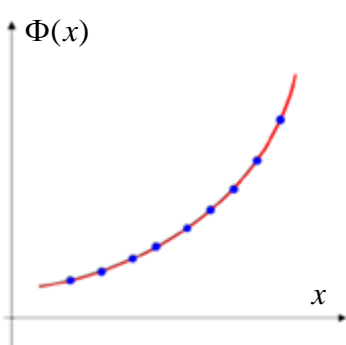
Roots of Legendre Transform

- Legendre Transform and Plücker Geometry
- Adrien-Marie Legendre and Gaspard Monge solve Minimal surface problem by use of Legendre Transform
- Chasles and Darboux interpreted Legendre Transform as reciprocal polar with respect to a paraboloid (re-use by Hadamard and Fréchet in calculus of variations)
- Alexis Clairaut introduced previously Clairaut Equation
- Maurice Fréchet introduced Clairaut equation associated to « distinguished densities » (densities with parameters achieving the Fréchet-Cramer-Rao Bound)

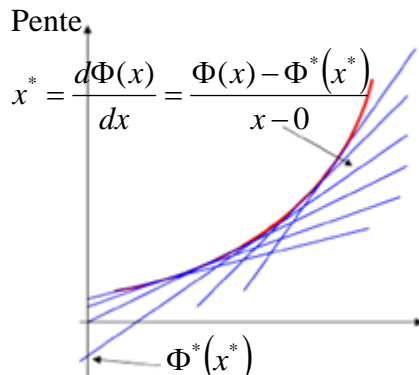
Legendre Transform interpretation

Legendre Transform

- Legendre Transform transforms one function defined by its value in one point in a function defined by its tangent.
- Used in thermodynamics and Lagrangian Mechanics



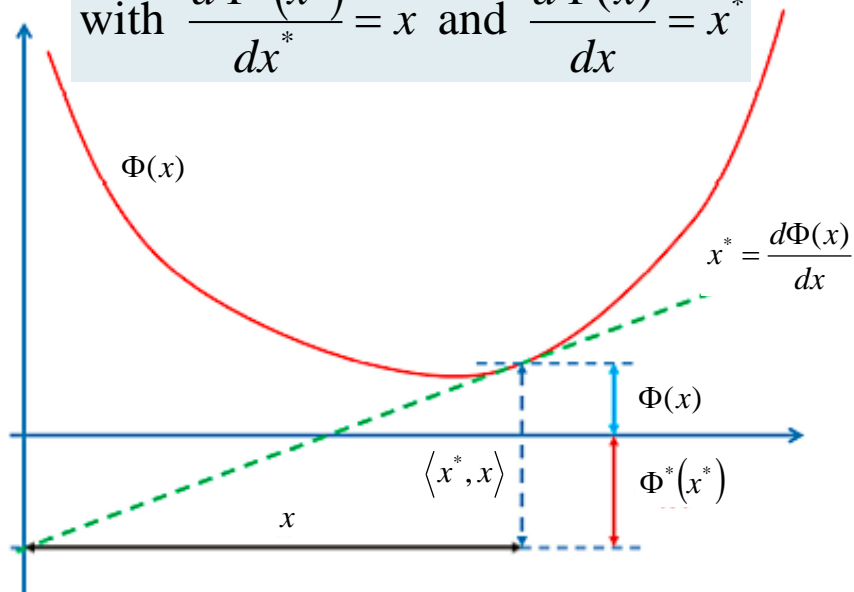
Classical Geometry
(curve is given by a continuum of points)



Plücker Geometry
(curve is given by the envelop of its tangents)

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

$$\text{with } \frac{d\Phi^*(x^*)}{dx^*} = x \text{ and } \frac{d\Phi(x)}{dx} = x^*$$



Legendre Transform is equivalent to Fourier Transform for convex function (duality)

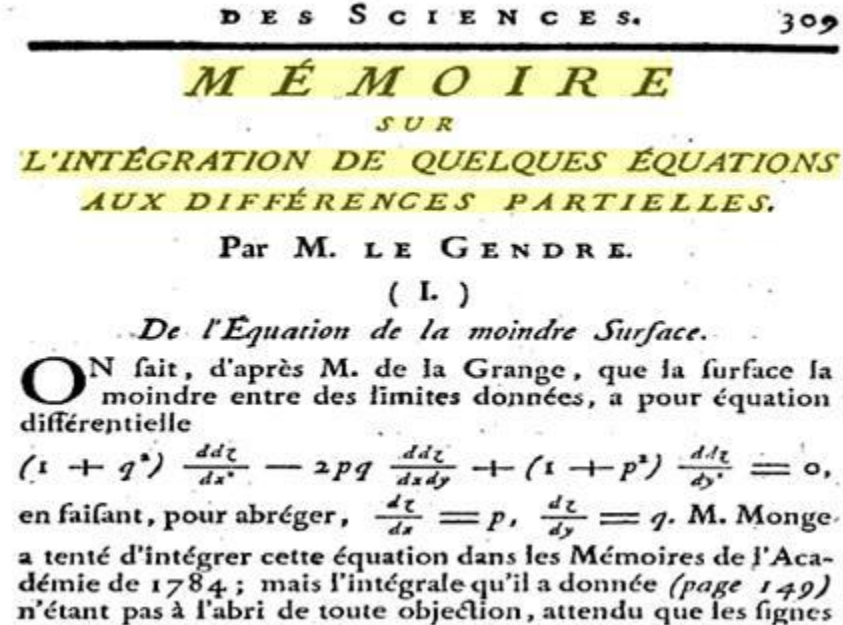
Brenier, Yann. Un algorithme rapide pour le calcul de transformées de Legendre-Fenchel discrètes, C. R. Acad. Sci. Paris Sér. I Math. 308 (1989), no. 20, 587–589.

Legendre Transform, 1787

1787, Adrien-Marie Legendre, "Mémoire sur l'intégration de quelques équations aux différences partielles".

- Adrien-Marie Legendre has introduced Legendre transform to solve a minimal surface problem given by Monge (Monge requested him to consolidate its proof).
- Legendre said "**J'y suis parvenu simplement par un changement de variables qui peut être utile dans d'autres occasions**".

Legendre, A.M. Mémoire Sur L'intégration de Quelques Equations aux Différences Partielles; Mémoires de l'Académie des Sciences: Paris, France, 1787; pp. 309–351.



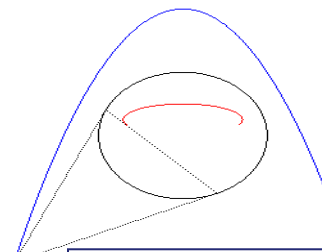
Lorsque la valeur de ω fera connue, il est clair qu'on aura celles de x, y, z , exprimées en p & q ; savoir,

$$x = \frac{d\omega}{dp}, \quad y = \frac{d\omega}{dq},$$
$$z = px + qy - \omega.$$

Reciprocal Polar with respect to a paraboloid

Legendre Transform & Reciprocal Polar

- **Darboux** gave in his book one interpretation of **Chasles** : « *Ce qui revient suivant une remarque de M. Chasles, à substituer à la surface sa polaire réciproque par rapport à un parabolôide* »
- In the lecture « Leçons sur le calcul des variations », **J. Hadamard**, followed by **M.E. Vessiot**, used reciprocal polar of figurative, and figuratrice.
- Note of **Paul Belgodère** presented by Elie Cartan « *Extrémale d'une surface* »



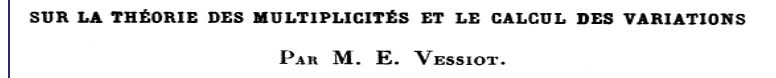
M. Chasles



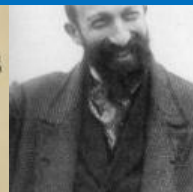
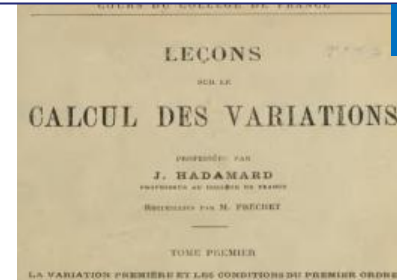
G. Darboux



M.E. Vessiot



J. Hadamard



Partons d'abord du problème de Lagrange, et posons comme précédemment

$$(200) \quad q_i = f'_i \quad (i = 1, \dots, n)$$

$$(201) \quad H = \sum_{i=1}^n y_i f'_i - f.$$

$$(202) \quad H = H(q_1, \dots, q_n, y_1, \dots, y_n, x)$$

La différentielle totale de H sera dès lors la même qu'à n° 140

$$(204) \quad dH = \sum_i y'_i dq_i - f_{y_i} dy_i.$$

La transformation de Legendre, définie par les équations (200), (201) reviendra encore à prendre la polaire réciproque de cette figurative par rapport au parabolôide

$$(205) \quad y'_1{}^2 + y'_2{}^2 - 2f_0 = 0.$$

Geometric Interpretation of Legendre Transform by Reciprocal Polar with respect to a paraboloid

- First, Let's consider the surface: $z = f(x, y)$ with $p = \frac{\partial z}{\partial x}$ and $q = \frac{\partial z}{\partial y}$
- We consider the equation of the paraboloid: $x^2 + y^2 = 2z$
- Reciprocal Polar with respect to paraboloid has coordinates: X, Y, Z
- the Polar plan with respect to paraboloid of this Reciprocal Polar $Xx + Yy - z - Z = 0$ should be equal to tangent plan of the surface at point (x_0, y_0, z_0) :

$$z - z_0 = p_0(x - x_0) + q_0(y - y_0) \Rightarrow p_0x + q_0y - z - (p_0x_0 + q_0y_0 - z_0) = 0$$

- This equality provides:

$$X = p_0, \quad Y = q_0, \quad Z = p_0x_0 + q_0y_0 - z_0$$

This is the **Legendre Transform**

(1) GOURSAT, *Traité d'Analyse*, tome I, p. 84; HENRI LEBESGUE, *Traité d'Analyse*, tome I, p. 105.

So in classical thermodynamics, Legendre transform $S(Q) = \langle \beta, Q \rangle - \Phi(\beta)$ is linked with polar reciprocal with respect to the paraboloid:

$$Q^2 = 2S(Q)$$

Geometric Interpretation of Legendre Transform by Reciprocal Polar with respect to a paraboloid

➤ We have $z = f(x, y)$ with $p = \frac{\partial z}{\partial x}$ and $q = \frac{\partial z}{\partial y}$ and $X = p$, $Y = q$, $Z = px + qy - z$ and the **Legendre Transform**

➤ We compute the first derivative of Z : $dZ = PdX + QdY$ with $P = \frac{\partial Z}{\partial X}$ and $Q = \frac{\partial Z}{\partial Y}$

$$Z = px + qy - z \Rightarrow dZ = pdx + qdy - dz + xdp + ydq \xRightarrow[\substack{dz=px+qdy \\ X=p, Y=q}]{\Rightarrow} dZ = x dX + y dQ \Rightarrow P = x, Q = y$$

➤ We compute 2nd derivative of Z :

$$R = \frac{\partial^2 Z}{\partial X^2} = \frac{\partial P}{\partial X} = \frac{\partial x}{\partial X}, \quad T = \frac{\partial^2 Z}{\partial X \partial Y} = \frac{\partial P}{\partial Y} = \frac{\partial Q}{\partial X} = \frac{\partial x}{\partial Y} = \frac{\partial y}{\partial X}, \quad S = \frac{\partial^2 Z}{\partial Y^2} = \frac{\partial Q}{\partial Y} = \frac{\partial y}{\partial Y}$$

$$r = \frac{\partial^2 z}{\partial x^2}, t = \frac{\partial^2 z}{\partial y^2}, s = \frac{\partial^2 z}{\partial x \partial y} \Rightarrow \begin{cases} dx = \frac{t}{rt - s^2} dX - \frac{s}{rt - s^2} dY \\ dy = \frac{-s}{rt - s^2} dX + \frac{r}{rt - s^2} dY \end{cases} \Rightarrow \begin{cases} R = \frac{\partial x}{\partial X} = \frac{t}{rt - s^2} \\ S = \frac{\partial x}{\partial Y} = \frac{-s}{rt - s^2} \\ T = \frac{\partial y}{\partial Y} = \frac{r}{rt - s^2} \end{cases} \Rightarrow \begin{cases} r = \frac{T}{RT - S^2} \\ s = \frac{-S}{RT - S^2} \\ t = \frac{R}{RT - S^2} \end{cases}$$

Geometric Interpretation of Legendre Transform by Reciprocal Polar with respect to a paraboloid

Links with Contact transformations

- Considering new variables X, Y, Z and P, Q the derivatives of Z with respect to X and Y , problem of finding in which case this five quantities could be express of x, y, z, p and q est the same problem where we look for five functions X, Y, Z, P and Q of five independant variables x, y, z, p and q satisfying the differential equation:

$$dZ - PdX - QdY = \rho(dz - pdx - qdy)$$

where ρ is a function of x, y, z, p and q

Proof:

$$\begin{cases} p = \frac{\partial z}{\partial x} \\ q = \frac{\partial z}{\partial y} \end{cases} \Rightarrow dz - pdx - qdy = 0 \Rightarrow dZ = PdX + QdY \Rightarrow \begin{cases} P = \frac{\partial Z}{\partial X} \\ Q = \frac{\partial Z}{\partial Y} \end{cases}$$

Reciprocal:

$$\rho = \frac{\partial Z}{\partial z} - P \frac{\partial X}{\partial z} - Q \frac{\partial Y}{\partial z}$$

Geometric Interpretation of Legendre Transform by Reciprocal Polar with respect to a paraboloid

Links with Ampere transformation

- Ampere transformation:

$$dz - pdx - qdy = d(z - qy) - pdx + ydq$$

$$\text{Set } \begin{cases} Z = z - qy, X = x, Y = q \\ P = p, Q = -y \end{cases} \Rightarrow dZ - p dX - Q dY = dz - pdx - qdy$$

- Then $\rho = 1$, and we have a contact transformation, also valid when Legendre transform is no longer valid (when $rt - s^2 = 0$, p and q are not independent)

Legendre transformation and Ampere transformation

- Legendre transform is obtained by same equality

$$dz - pdx - qdy = d(z - px - qy) - xdp - ydq$$

- We can set $X = p, Y = q, Z = z - px - qy$

$$P = x, Q = y$$

Clairaut Equation, 1734

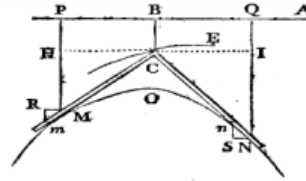
1724, Alexis Claude Clairaut

- Clairaut introduced before Legendre an equation related to Legendre transform.
- Legendre transform

$$\Phi^*(x^*) = \langle x^*, x \rangle - \Phi(x) \quad \text{avec} \quad \frac{d\Phi^*(x^*)}{dx^*} = x \quad \text{et} \quad \frac{d\Phi(x)}{dx} = x^*$$

➤ Clairaut Equation

$$\Phi^*(x^*) = \left\langle x^*, \frac{d\Phi^*(x^*)}{dx^*} \right\rangle - \Phi\left(\frac{d\Phi^*(x^*)}{dx^*}\right)$$



- Singular solution: envelop of straight lines solutions

$$\frac{d\Phi^*(x^*)}{dx^*} = \frac{d\Phi^*(x^*)}{dx^*} + \left\langle x^*, \frac{d^2\Phi^*(x^*)}{dx^{*2}} \right\rangle - \Phi' \left(\frac{d\Phi^*(x^*)}{dx^*} \right) \frac{d^2\Phi^*(x^*)}{dx^{*2}}$$

$$\Rightarrow 0 = \left\langle x^* - \Phi' \left(\frac{d\Phi^*(x^*)}{dx^*} \right), \frac{d^2\Phi^*(x^*)}{dx^{*2}} \right\rangle \Rightarrow x^* = \Phi' \left(\frac{d\Phi^*(x^*)}{dx^*} \right) \quad \text{et} \quad \Phi(x^*) = \langle C, x^* \rangle + \Phi(C)$$

196 MEMOIRES DE L'ACADEMIE ROYALE

SOLUTION DE PLUSIEURS PROBLEMES

Où il s'agit de trouver des Courbes dont la propriété consiste dans une certaine relation entre leurs branches, exprimée par une Equation donnée.

Par M. CLAIRAUT.

DANS les Courbes dont on parle dans ce Mémoire, il ne suffit pas, comme dans la plupart des autres, de considérer un de leurs points quelconques, ou une partie infiniment petite de la Courbe pour la déterminer toute entière. Les propriétés de celles-ci demandent nécessairement qu'on prenne à la fois plusieurs points à des distances finies les uns des autres, & dans des branches différentes.

Les Problemes que je vais donner, & ceux qui sont de la même espèce, seroient fort faciles, si, pour trouver les Courbes qui en font la solution, on se contentoit de prendre deux ou plusieurs branches de différentes Courbes, au lieu de trouver une seule Courbe qui les comprenne toutes. Prenant une branche d'une Courbe quelconque, on en trouveroit aisément d'autres par les méthodes ordinaires, qui auroient avec cette première la relation demandée. Mais pour faire en sorte que les différentes branches appartiennent toutes à la même Courbe, il faut nécessairement avoir recours à d'autres méthodes qui ajoutent de plus grandes difficultés à ces Problemes.

Il n'y a eu jusqu'ici, du moins que je sçache, que très-peu de Problemes de cette nature, on peut dire même qu'il n'y a d'expliqué que le fameux Probleme des Trajectoires réciproques, dont M.^{rs} Bernoulli, Pemberton & Euler ont donné des solutions dans les Actes de Leipzig, années 1718,

Foundations of Geometric Structure of Information: TRIBUTE TO J-L KOSZUL



Jean-Louis Koszul 1921-2018



**KOSZUL
- ENS ULM**

KOSZUL



KOSZUL - BOURBAKI GROUP



**KOSZUL
- STRASBOURG**



**KOSZUL
- GRENOBLE**



**KOSZUL
- GSI'13, Ecole des Mines de Paris**



KOSZUL - Last Interview

JEAN-LOUIS KOSZUL - DIRECTEUR DU LABORATOIRE, 1978-1981

Jean-Louis Koszul: 1921-2018



**KOSZUL
ENS ULM**

KOSZUL



**KOSZUL
BOURBAKI**



**KOSZUL
STRASBOURG**



**KOSZUL
GRENOBLE**



**KOSZUL
CIRM BIRTHDAY**

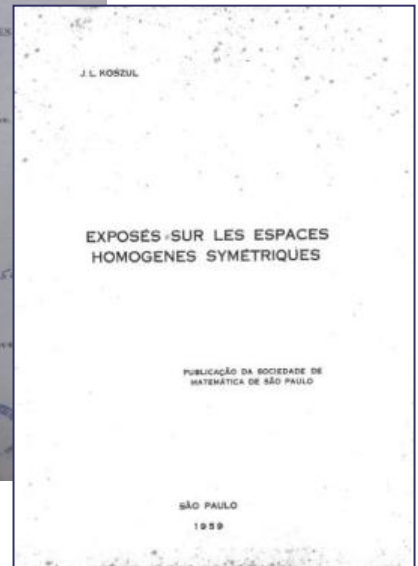
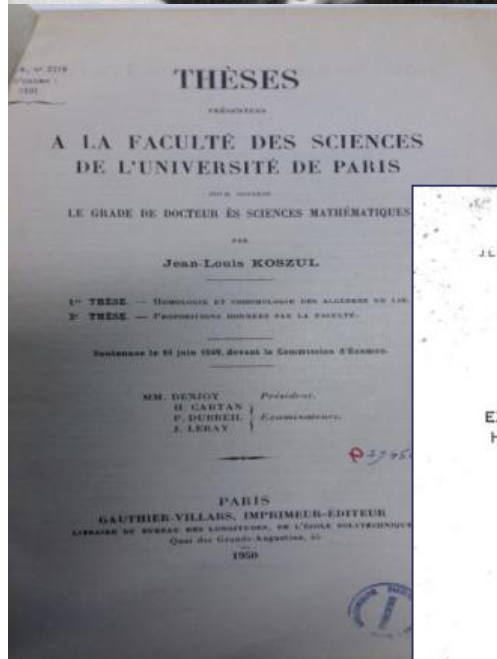


**KOSZUL
GSI '13 Mines**

Interview of Jean-Louis Koszul: <https://www.youtube.com/watch?v=AzK5K7Q05sw>

Jean-Louis Koszul Scientific Biography

- He entered ENS Ulm in the class of 1940 and defended his thesis with Henri Cartan.
- Koszul's thesis, defended in June 10th 1949 under the direction of Henri Cartan, dealt with the homology and cohomology of Lie algebras. The jury was composed of M. Denjoy (President), J. Leray, P. Dubreil and H. Cartan.
- He then taught in Strasbourg and was appointed Associate Professor at the University of Strasbourg in 1949, and had for colleagues R. Thom, M. Berger and B. Malgrange. He was promoted to professor status in 1956.
- He became a member of Bourbaki with the 2nd generation, J. Dixmier, R. Godement, S. Eilenberg, P. Samuel, J. P. Serre and L. Schwartz.
- Koszul was awarded by Jaffré Prize in 1975 and was elected correspondent at the Academy of Sciences on January 28th 1980. Koszul was one of the CIRM conference center founder at Luminy. The following year, he was elected to the Academy of São Paulo.



Cooperation with Japanese School of Differential Geometry: Hirohiko Shima Book on « the geometry of Hessian Structures »

- The elementary geometric structures discovered by Jean-Louis Koszul are the foundations of Information Geometry. These links were first established by Professor Hirohiko Shima.
- These links were particularly crystallized in Shima book 2007 "The Geometry of Hessian Structures", which is dedicated to Professor Koszul.
- The origin of this work followed the visit of Koszul in Japan in 1964, for a mission coordinated with the French government.
- Koszul taught lectures on the theory of flat manifolds at Osaka University. Hirohiko Shima was then a student and attended these lectures with the teachers Matsushima and Murakami.
- This lecture was at the origin of the notion of Hessian structures and the beginning of the works of Hirohiko Shima.



Henri Cartan noted concerning Koszul's ties with Japan, "Koszul has attracted eminent mathematicians from abroad to Strasbourg and Grenoble. I would like to mention in particular the links he has established with representatives of the Japanese School of Differential Geometry".

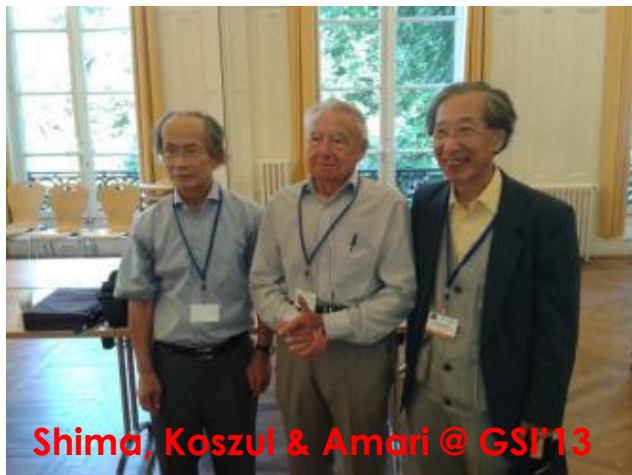
Jean-Louis Koszul (1921-2018) and Information Geometry



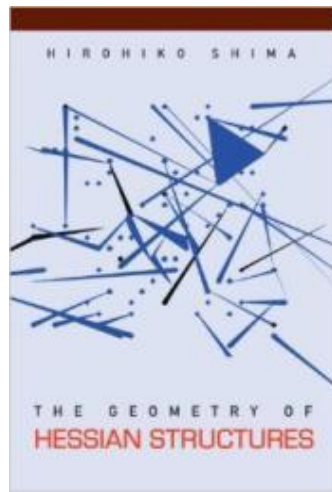
Koszul, Shima, Boyom @ GSI'13



H. Shima Talk « Koszul
Geometry of Hessian
Structure » @ GSI'13



Shima, Koszul & Amari @ GSI'13



HIRONOBU SHIMA

THE GEOMETRY OF
HESSIAN STRUCTURES



Koszul & Shima @ GSI'13

Jean-Louis Koszul and the elementary structures of Information Geometry

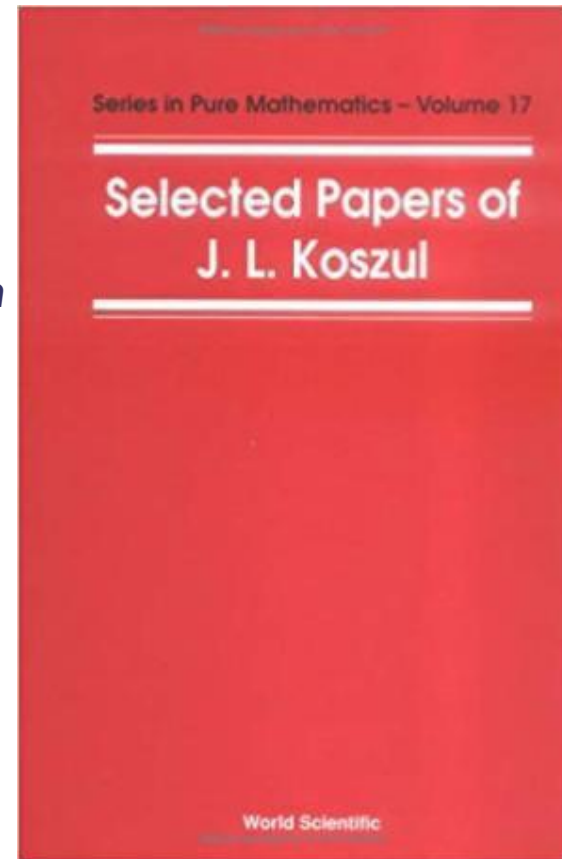
- This tribute is a scientific exegesis and admiration of Jean-Louis Koszul's works on homogeneous bounded domains that have appeared over time as elementary structures of Information Geometry.
- Koszul has introduced fundamental tools to characterize the geometry of sharp convex cones, as Koszul-Vinberg characteristic Function, Koszul Forms, and affine representation of Lie Algebra and Lie Group.
- The 2nd Koszul form is linked to an extension of classical Fisher metric.
- Koszul theory of hessian structures and Koszul forms could be considered as main foundation and pillars of Information Geometry.

Koszul « avant-garde »

- Inspired by the French mathematical tradition, and the teachings of his master Elie Cartan (Koszul was PhD student of Henri Cartan but was greatly influenced by Elie Cartan), Jean-Louis Koszul was a real “avant-garde”.
- We will explore only one part of his work which concerns **homogeneous bounded domains geometry**, from seminal Elie Cartan's earlier work on symmetric bounded domains. In a letter from André Weil to Henri Cartan, cited in the proceedings of the conference "Elie Cartan and today's mathematics" in 1984, Weil says "**As to the symmetrical spaces, and more particularly to the symmetric bounded domains at the birth of which you contributed, I have kept alive the memory of the satisfaction I felt in finding some incarnations in Siegel from his first works on quadratic forms, and later to convince Siegel of the value of your father's ideas on the subject**".
- At this 1984 conference "Elie Cartan and today's mathematics", two disciples of Elie Cartan gave a conference, Jean-Louis Koszul and Jean-Marie Souriau.

Koszul works on Homogeneous Bounded Domains

- In the book "**Selected papers of JL Koszul**", Koszul summarizes his work on homogeneous bounded domains: ***"It is with the problem of the determination of the homogeneous bounded domains posed by E. Cartan around 1935 that are related [my papers]. The idea of approaching the question through invariant Hermitian forms already appears explicitly in Cartan. This leads to an algebraic approach which constitutes the essence of Cartan's work and which, with the Lie J-algebras, was pushed much further by the Russian School. It is the work of Piatetski Shapiro on the Siegel domains, then those of E.B. Vinberg on the homogeneous cones that led me to the study of the affine transformation groups of the locally flat manifolds and in particular to the convexity criteria related to invariant forms"***.



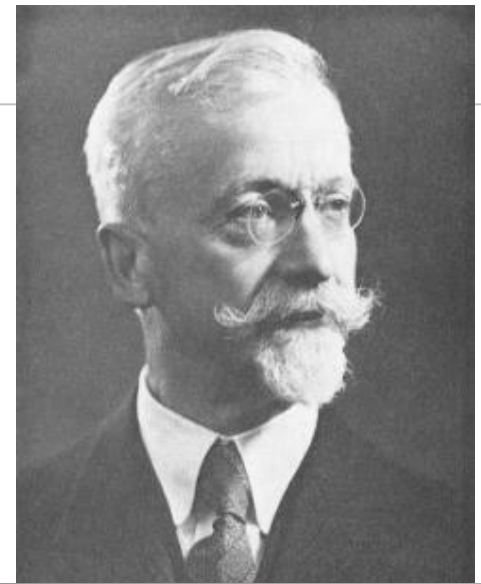
« La source » of Koszul inspiration: Elie Cartan

1935 Elie Cartan paper

➤ J.L. Koszul source of inspiration is given in this last sentence of Elie Cartan's 1935 paper where we can read in the last sentence:

➤ **“It is clear that if one could demonstrate that all homogeneous domains whose form is positive definite are symmetric, the whole theory of homogeneous bounded domains would be elucidated. This is a problem of Hermitian geometry certainly very interesting”** – Elie Cartan 1935

- Cartan E., Sur les invariants intégraux de certains espaces homogènes clos et les propriétés topologiques de ces espaces, Ann. Soc. Pol. De Math., n°8, pp.181-225, 1929
- Cartan E., Sur les domaines bornés de l'espace de n variables complexes, Abh. Math. Seminar Hamburg, vol. 1, p.116-162, 1935



$$\Phi \equiv \sum \frac{\partial^2 \log K}{\partial z_i \partial \bar{z}_j} dz_i d\bar{z}_j$$

THÉORÈME. Pour qu'un domaine homogène à $n \leq 3$ dimensions a groupe de stabilité clos soit équivalent à un domaine borné, il faut et il suffit que la forme différentielle Φ calculée, comme il a été dit plus haut, au moyen du groupe du domaine, soit définie positive.

Il serait intéressant de savoir si ce théorème s'étend aux valeurs de n supérieures à 3.

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

- **“[Détecter l'origine d'une notion ou la première apparition d'un résultat est souvent difficile. Je ne suis certainement pas le premier à avoir utilisé des représentations affines de groupes ou d'algèbres de Lie. On peut effectivement imaginer que cela se trouve chez Elie Cartan, mais je ne puis rien dire de précis. A propos d'Elie Cartan: je n'ai pas été son élève. C'est Henri Cartan qui a été mon maître pendant mes années de thèse. En 1941 ou 42 j'ai entendu une brève série de conférences données par Elie à l'Ecole Normale et ce sont des travaux d'Elie qui ont été le point de départ de mon travail de thèse.]** There are many things that I would like to understand (too much perhaps!), if only the relationship between what I did and the work of Souriau. Detecting the origin of a notion or the first appearance of a result is often difficult. I am certainly not the first to have used affine representations of groups or Lie algebras. We can imagine that it is at Elie Cartan, but I cannot say anything specific. About Elie Cartan: I was not his student. It was Henri Cartan who was my master during my years of thesis. In 1941 or 42, I heard a brief series of lectures given by Elie at the Ecole Normale and it was Elie's work that was the starting point of my thesis work”.

Koszul's paper at the foundation of the elementary structure of Information

- Koszul J.L., Sur la forme hermitienne canonique des espaces homogènes complexes. Can. J. Math., n°7, 562–576, 1955
- Koszul J.L., Exposés sur les Espaces Homogènes Symétriques; Publicação da Sociedade de Matematica de São Paulo: São Paulo, Brazil, 1959
- Koszul J.L., Domaines bornées homogènes et orbites de groupes de transformations affines, Bull. Soc. Math. France 89, pp. 515-533., 1961
- Koszul J.L., Ouverts convexes homogènes des espaces affines. Math. Z., n°79, 254–259, 1962
- Koszul J.L. Sous-groupes discrets des groupes de transformations affines admettant une trajectoire convexe, C.R. Acad. Sc. T.259, pp.3675-3677, 1964
- Koszul J.L., Variétés localement plates et convexité. Osaka. J. Math., n°2, 285–290, 1965
- Koszul J.L, Lectures on Groups of Transformations, Tata Institute of Fundamental Research, Bombay, 1965
- Koszul J.L., Déformations des variétés localement plates, .Ann Inst Fourier, n°18 , 103-114., 1968
- Koszul J.L., Trajectoires Convexes de Groupes Affines Unimodulaires. In Essays on Topology and Related Topics; Springer: Berlin, Germany, pp. 105–110, 1970
- Selected Papers of J L Koszul, Series in Pure Mathematics, Volume 17, World Scientific Publ, 1994

Koszul's main papers related to the elementary structures of information geometry (1/10)

1955 « Sur la forme hermitienne canonique des espaces homogènes complexes »

- Koszul considers the Hermitian structure of a homogeneous G/B manifold (G related Lie group and B a closed subgroup of G , associated, up to a constant factor, to the single invariant G , and to the invariant complex structure by the operations of G).
- Koszul says "**The interest of this form for the determination of homogeneous bounded domains has been emphasized by Elie Cartan: a necessary condition for G/B to be a bounded domain is indeed that this form is positive definite**". Koszul calculated this canonical form from infinitesimal data Lie algebra of G , the sub-algebra corresponding to B and an endomorphism algebra defining the invariant complex structure of G/B . The results obtained by Koszul proved that the homogeneous bounded domains whose group of automorphisms is semi-simple are bounded symmetric domains in the sense of Elie Cartan. Koszul also refers to André Lichnerowicz's work on Kählerian homogeneous spaces. In this seminal paper, Koszul also introduced a left invariant form of degree 1 on G :

$$\Psi(X) = \text{Tr}_{g/b} \left[\text{ad}(JX) - J.\text{ad}(X) \right] \quad \forall X \in g$$

- with J an endomorphism of the Lie algebra space and the trace $\text{Tr}_{g/b}[\cdot]$ corresponding to that of the endomorphism g/b . The Kähler form of the canonical Hermitian form is given by the differential of $-\frac{1}{4}\Psi(X)$ of this form of degree 1.

Koszul's main papers, related to the elementary structures of information geometry (2/10)

Koszul Forms for Homogeneous Bounded domains

- Koszul has developed his previously described theory for Homogeneous Siegel Domains SD. He has proved that there is a subgroup G in the group of the complex affine automorphisms of these domains (Iwasawa subgroup), such that G acts on SD simply transitively. The Lie algebra \mathfrak{g} of G has a structure that is an algebraic translation of the Kähler structure of SD.
- There is an integrable almost complex structure J on \mathfrak{g} and there exists $\eta \in \mathfrak{g}^*$ such that $\langle X, Y \rangle_\eta = \langle [JX, Y], \eta \rangle$ defines a J -invariant positive definite inner product on \mathfrak{g} . Koszul has proposed as admissible form $\eta \in \mathfrak{g}^*$, the form ξ :

$$\Psi(X) = \langle X, \xi \rangle = \text{Tr}[ad(JX) - J.ad(X)] \quad \forall X \in \mathfrak{g}$$

- Koszul has proved that $\langle X, Y \rangle_\xi$ coincides, up to a positive number multiple with the real part of the Hermitian inner product obtained by the Bergman metric of SD by identifying \mathfrak{g} with the tangent space of SD. The Koszul forms are then given by:

$$\alpha = -\frac{1}{4} d\Psi(X)$$

$$\beta = D\alpha$$

Koszul's main papers, related to the elementary structures of information geometry (3/10)

Koszul Forms

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

$$\Psi(X) = \text{Tr}_{\mathfrak{g}/\mathfrak{b}} [ad(JX) - Jad(X)] \quad \forall X \in \mathfrak{g}$$

➤ 2nd Koszul Form: $\beta = D\alpha$

Application for Poincaré Upper-Half Plane:

$$V = \{z = x + iy \mid y > 0\} \quad Y = y \frac{d}{dy} \Rightarrow \begin{cases} ad(Y).Z = [Y, Z] \\ [X, Y] = -Y \\ JX = Y \end{cases}$$

With $X = y \frac{d}{dx}$ and $\begin{cases} \text{Tr}[ad(JX) - Jad(X)] = 2 \\ \text{Tr}[ad(JY) - Jad(Y)] = 0 \end{cases}$

We can deduce that

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

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



Koszul's main papers, related to the elementary structures of information geometry (4/10)

■ Koszul form for Siegel Upper-Half Space: $V = \{Z = X + iY / Y > 0\}$

► Symplectic Group :

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

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

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

Koszul's main papers, related to the elementary structures of information geometry (5/10)

1959 « Exposés sur les espaces homogènes symétriques »

- It is a Lecture written as part of a seminar held in September and October 1958 at the University of Sao Paulo
- Koszul showed that any symmetric bounded domain is a direct product of irreducible symmetric bounded domains, determined by Elie Cartan (4 classes corresponding to classical groups and 2 exceptional domains).
- For the study of irreducible symmetric bounded domains, Koszul referred to Elie Cartan, Carl-Ludwig Siegel and Loo-Keng Hua.
- Koszul illustrated the subject with two particular cases, the half-plane of Poincaré and the half-space of Siegel, and showed that with its trace formula of endomorphism g/b , he found that the canonical Kähler hermitian form and the associated metrics are the same as those introduced by Henri Poincaré and Carl-Ludwig Siegel (who introduced them as invariant metric under action of the automorphisms of these spaces).

Koszul's main papers, related to the elementary structures of information geometry (6/10)

1961 « Domaines bornés homogènes et orbites de groupes de transformations affines »

- It is written by Koszul at the Institute for Advanced Study at Princeton during a stay funded by the National Science Foundation.
- On a complex homogeneous space, an invariant volume defines with the complex structure the canonical invariant Hermitian form. If the homogeneous space is holomorphically isomorphic to a bounded domain of a space C^n , this Hermitian form is positive definite because it coincides with the Bergmann metric of the domain.
- Koszul demonstrated in this article the reciprocal of this proposition for a class of complex homogeneous spaces. This class consists of some open orbits of complex affine transformation groups and contains all homogeneous bounded domains. Koszul addressed again the problem of knowing if a complex homogeneous space, whose canonical Hermitian form is positive definite is isomorphic to a bounded domain, but via the study of the invariant bilinear form defined on a real homogeneous space by an invariant volume and an invariant flat connection.

Koszul's main papers, related to the elementary structures of information geometry (7/10)

1961 « Domaines bornés homogènes et orbites de groupes de transformations affines »

- Koszul demonstrated that if this bilinear form is positive definite then the homogeneous space with its flat connection is isomorphic to a convex open domain containing no straight line in a real vector space and extended it to the initial problem for the complex homogeneous spaces obtained in defining a complex structure in the variety of vectors of a real homogeneous space provided with an invariant flat connection.
- It is in this article that Koszul used the affine representation of Lie groups and algebras. By studying the open orbits of the affine representations, he introduced an affine representation of G , written (\mathbf{f}, \mathbf{q}) , and the following equation setting f the linear representation of the Lie algebra \mathfrak{g} of G , defined by \mathbf{f} and q the restriction to \mathfrak{g} and the differential of \mathbf{q} (f and q are differential respectively of \mathbf{f} and \mathbf{q}):

$$f(X)q(Y) - f(Y)q(X) = q([X, Y]) \quad \forall X, Y \in \mathfrak{g}$$

$$\text{with } f : \mathfrak{g} \rightarrow gl(E) \quad \text{and} \quad q : \mathfrak{g} \mapsto E$$

Koszul's main papers, related to the elementary structures of information geometry (8/10)

1962 « Ouverts convexes homogènes des espaces affines »

- Koszul is interested in this paper by the structure of the convex open non-degenerate Ω (with no straight line) and homogeneous (the group of affine transformations of E leaving stable Ω operates transitively in Ω) in a real affine space of finite dimension.
- Koszul demonstrated that they can be all deduced from non-degenerate and homogeneous convex open cones built previously.
- He used for this the properties of the group of affine transformations leaving stable a non-degenerate convex open domain and an homogeneous domain.

Koszul's main papers, related to the elementary structures of information geometry (9/10)

1965 « Variétés localement plates et convexité »

➤ Koszul established the following **Koszul's theorem**:

- Let M be a locally related differentiable manifold. If the universal covering of M is isomorphic as a flat manifold with a convex open domain containing no straight line in a real affine space, then there exists on M a closed differential form α such that $D\alpha$ (D linear covariant derivative of zero torsion) is positive definite in all respects and which is invariant under every automorphism of M .
- If G is a group of automorphisms of M such that $G \backslash M$ is quasi-compact and if there exists on M a closed 1-differential form α invariant by G and such that $D\alpha$ is positive definite at any point, then the universal covering of M is isomorphic as a flat manifold with a convex open domain that does not contain a straight line in a real affine space.

1965 « Lectures on Groups of Transformations »

- This is lecture notes given by Koszul at Bombay "Tata Institute of Fundamental Research" on transformation groups.
- In particular in Chapter 6, Koszul studied discrete linear groups acting on convex open cones in vector spaces based on the work of C.L. Siegel (work on quadratic forms).
- Koszul used what we will call in the following Koszul-Vinberg characteristic function on convex sharp cone.

Koszul's main papers, related to the elementary structures of information geometry (10/10)

1968 « Déformations des variétés localement plates »

- Koszul provided other proofs of theorems introduced previously.
- Koszul considered related differentiable manifolds of dimension n and TM the fibered space of M . The linear connections on M constitute a subspace of the space of the differentiable applications of the $TM \times TM$ fiber product in the space $T(TM)$ of the TM vectors.
- Any locally flat connection D (the curvature and the torsion are zero) defines a locally flat connection on the covering of M , and is hyperbolic when universal covering of M , with this connection, is isomorphic to a sharp convex open domain (without straight lines) in \mathbb{R}^n .
- Koszul showed that, if M is a compact manifold, for a locally flat connection on M to be hyperbolic, it is necessary and sufficient that there exists a closed differential form of degree 1 on M whose covariant differential is positive definite.

1970 « Trajectoires Convexes de Groupes Affines Unimodulaires »

- Koszul demonstrated that a convex sharp open domain in \mathbb{R}^n that admits a unimodular transitive group of affine automorphisms is an auto-dual cone.
- This is a more geometric demonstration of the results shown by Ernest Vinberg on the automorphisms of convex cones.

The Geometry of Hessian Geometry and Koszul forms

The Elementary Structures:

- Codazzi Structure (D, g) , D is a connexion without torsion: $(D_X g)(Y, Z) = (D_Y g)(X, Z)$
- Hessian structure: (D, g) Codazzi with D is flat \Rightarrow dual structure (D', g) with : $D' = \nabla - D$ (with ∇ la Levi-Civita connexion)
- For a hessian structure (D, g) with $g = Dd\varphi$, $g = D'd\varphi'$ and the dual Codazzi structure (D', g) is also a hessian structure
- We have the property that φ' is the Legendre transform of φ : $\varphi' = \sum_i x^i \frac{\partial \varphi}{\partial x^i} - \varphi$
- A hessian structure (D, g) is a Koszul structure, if there is a closed 1-form ω such that $g = D\omega$
- Koszul has introduced a 2-form, that plays same role than Ricci tensor for a kahlerian metric: $\gamma = D\alpha$ with 1-form α , such that $D_X \nu = \alpha(X)\nu$ with volume element ν , and for (D', g) : $\alpha' = -\alpha$ and $\gamma' = \gamma - 2\nabla\alpha$
- For an homogeneous regular convex cone Ω , the hessian structure (D, g) is given by $g = Dd\psi$ with Koszul forms $\alpha = d \log \psi$ and $\gamma = g$
Volume element ν is invariant under the action of automorphisms of Ω .

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 dimension on R (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

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):

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

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

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

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

Koszul Entropy via Legendre Transform

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

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

with $x^* = D_x \Phi$ and $x = D_{x^*} \Phi^*$ where $\psi_\Omega(x) = \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \quad \forall x \in \Omega$

Demonstration: we set $x^* = \int_{\Omega^*} \xi \cdot e^{-\langle \xi, x \rangle} d\xi / \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$

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

and $-\langle x^*, x \rangle = \int_{\Omega^*} \log e^{-\langle \xi, x \rangle} \cdot e^{-\langle \xi, x \rangle} d\xi / \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$

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

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

Koszul-Vinberg Characteristic Function Legendre Transform

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

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

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

$$\Phi^*(x^*) = \left[\log \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \cdot \left(\int_{\Omega^*} \frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} d\xi \right) - \int_{\Omega^*} \log e^{-\langle \xi, x \rangle} \cdot \frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} d\xi \right] \text{ with } \int_{\Omega^*} \frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} d\xi = 1$$

$$\Phi^*(x^*) = \left[- \int_{\Omega^*} \frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} \cdot \log \left(\frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} \right) d\xi \right] \quad \text{SHANNON ENTROPY !}$$

To make the link with Fisher metric given by matrix $I(x)$, we can observe that the second derivative of $\log p_x(\xi)$ is given by:

$$\log p_x(\xi) = -\Phi^*(\xi) = \Phi(x) - \langle x, \xi \rangle$$

$$\frac{\partial^2 \log p_x(\xi)}{\partial x^2} = \frac{\partial^2 [\Phi(x) - \langle x, \xi \rangle]}{\partial x^2} = \frac{\partial^2 \Phi(x)}{\partial x^2}$$

$$\Rightarrow I(x) = -E_\xi \left[\frac{\partial^2 \log p_x(\xi)}{\partial x^2} \right] = -\frac{\partial^2 \Phi(x)}{\partial x^2} = \frac{\partial^2 \log \psi_\Omega(x)}{\partial x^2}$$

We could then deduce the close interrelation between Fisher metric and hessian of Koszul-Vinberg characteristic logarithm.

$$I(x) = -E_\xi \left[\frac{\partial^2 \log p_x(\xi)}{\partial x^2} \right] = \frac{\partial^2 \log \psi_\Omega(x)}{\partial x^2}$$

Koszul Metric and Fisher Metric as Variance

We can also observed that the Fisher metric or hessian of KVCF logarithm is related to the variance of ξ :

$$\log \Psi_{\Omega}(x) = \log \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \Rightarrow \frac{\partial \log \Psi_{\Omega}(x)}{\partial x} = - \frac{1}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} \int_{\Omega^*} \xi \cdot e^{-\langle \xi, x \rangle} d\xi$$

$$\frac{\partial^2 \log \Psi_{\Omega}(x)}{\partial x^2} = - \frac{1}{\left(\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \right)^2} \left[- \int_{\Omega^*} \xi^2 \cdot e^{-\langle \xi, x \rangle} d\xi \cdot \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi + \left(\int_{\Omega^*} \xi \cdot e^{-\langle \xi, x \rangle} d\xi \right)^2 \right]$$

$$\frac{\partial^2 \log \Psi_{\Omega}(x)}{\partial x^2} = \int_{\Omega^*} \xi^2 \cdot \frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} d\xi - \left(\int_{\Omega^*} \xi \cdot \frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} d\xi \right)^2 = \int_{\Omega^*} \xi^2 \cdot p_x(\xi) d\xi - \left(\int_{\Omega^*} \xi \cdot p_x(\xi) d\xi \right)^2$$

$$I(x) = -E_{\xi} \left[\frac{\partial^2 \log p_x(\xi)}{\partial x^2} \right] = \frac{\partial^2 \log \psi_{\Omega}(x)}{\partial x^2} = E_{\xi} [\xi^2] - E_{\xi} [\xi]^2 = \text{Var}(\xi)$$

Definition of Maximum Entropy Density

How to replace x by mean value of ξ , $\bar{\xi} (= x^*)$ in :

$$p_x(\xi) = \frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} \quad \text{with} \quad \bar{\xi} = \int_{\Omega^*} \xi \cdot p_x(\xi) d\xi$$

Legendre Transform will do this inversion by inverting $\bar{\xi} = \frac{d\Phi(x)}{dx}$

We then observe that Koszul Entropy provides density of Maximum Entropy with this general definition of density:

$$p_{\bar{\xi}}(\xi) = \frac{e^{-\langle \xi, \Theta^{-1}(\bar{\xi}) \rangle}}{\int_{\Omega^*} e^{-\langle \xi, \Theta^{-1}(\bar{\xi}) \rangle} d\xi}$$

$$x = \Theta^{-1}(\bar{\xi})$$

$$\bar{\xi} = \Theta(x) = \frac{d\Phi(x)}{dx}$$

where $\bar{\xi} = \int_{\Omega^*} \xi \cdot p_{\bar{\xi}}(\xi) d\xi$ and $\Phi(x) = -\log \int_{\Omega^*} e^{-\langle x, \xi \rangle} d\xi$

From Cartan-Killing Form to Koszul Information Metric

$$B(x, y) = \text{Tr}(ad_x ad_y)$$

Cartan – Killing Form

$$\langle x, y \rangle = -B(x, \theta(y))$$

with $\theta \in \mathfrak{g}$, Cartan Involution



Koszul Characteristic Function

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



Koszul Entropy

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

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

$$\text{with } x^* = \int_{\Omega^*} \xi \cdot p_x(\xi) d\xi$$

Koszul Density

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



Koszul Metric

$$I(x) = -E_{\xi} \left[\frac{\partial^2 \log p_x(\xi)}{\partial x^2} \right]$$

$$I(x) = -\frac{\partial^2 \Phi(x)}{\partial x^2} = \frac{\partial^2 \log \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi}{\partial x^2}$$

Relation of Koszul density with Maximum Entropy Principle

■ The density from Maximum Entropy Principle is given by:

$$\text{Max}_{p_x(\cdot)} \left[- \int_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi \right] \text{ such } \begin{cases} \int_{\Omega^*} p_x(\xi) d\xi = 1 \\ \int_{\Omega^*} \xi \cdot p_x(\xi) d\xi = x^* \end{cases}$$

■ If we take $q_x(\xi) = e^{-\langle \xi, x \rangle} / \int_{\Omega^*} e^{-\langle \zeta, x \rangle} d\zeta = e^{-\langle x, \xi \rangle - \log \int_{\Omega^*} e^{-\langle \zeta, x \rangle} d\zeta}$ such that

$$\begin{cases} \int_{\Omega^*} q_x(\xi) \cdot d\xi = \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi / \int_{\Omega^*} e^{-\langle \zeta, x \rangle} d\zeta = 1 \\ \log q_x(\xi) = \log e^{-\langle x, \xi \rangle - \log \int_{\Omega^*} e^{-\langle \zeta, x \rangle} d\zeta} = -\langle x, \xi \rangle - \log \int_{\Omega^*} e^{-\langle \zeta, x \rangle} d\zeta \end{cases}$$

■ Then by using the fact that $\log x \geq (1 - x^{-1})$ with equality if and only if $x = 1$, we find the following:

$$- \int_{\Omega^*} p_x(\xi) \log \frac{p_x(\xi)}{q_x(\xi)} d\xi \leq - \int_{\Omega^*} p_x(\xi) \left(1 - \frac{q_x(\xi)}{p_x(\xi)} \right) d\xi$$

Relation of Koszul density with Maximum Entropy Principle

■ We can then observe that:

$$\int_{\Omega^*} p_x(\xi) \left(1 - \frac{q_x(\xi)}{p_x(\xi)} \right) d\xi = \int_{\Omega^*} p_x(\xi) d\xi - \int_{\Omega^*} q_x(\xi) d\xi = 0$$

because $\int_{\Omega^*} p_x(\xi) d\xi = \int_{\Omega^*} q_x(\xi) d\xi = 1$

■ We can then deduce that:

$$-\int_{\Omega^*} p_x(\xi) \log \frac{p_x(\xi)}{q_x(\xi)} d\xi \leq 0 \Rightarrow -\int_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi \leq -\int_{\Omega^*} p_x(\xi) \log q_x(\xi) d\xi$$

■ If we develop the last inequality, using expression of $q_x(\xi)$:

$$-\int_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi \leq -\int_{\Omega^*} p_x(\xi) \left[-\langle x, \xi \rangle - \log \int_{\Omega^*} e^{-\langle x, \xi \rangle} d\xi \right] d\xi$$

$$-\int_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi \leq \left\langle x, \int_{\Omega^*} \xi \cdot p_x(\xi) d\xi \right\rangle + \log \int_{\Omega^*} e^{-\langle x, \xi \rangle} d\xi$$

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

$$-\int_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi \leq \Phi^*(x^*)$$

General Theory: 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

Filiation Poincaré/Cartan/Koszul

« 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



Henri Poincaré
(half-plane) $n=1$



Elie Cartan
(classification in 6 classes of symmetric homogeneous bounded domains)
 $n \leq 3$



Carl Ludwig Siegel
(Siegel space of 1st kind and Symplectic Geometry)



Lookeng Hua
(Bergman Kernel, Cauchy and Poisson of Siegel Domains)



Ernest Vinberg
(Siegel Domains of 2nd kind)

Structure of Information
Geometry
(Koszul Hessian
Geometry)

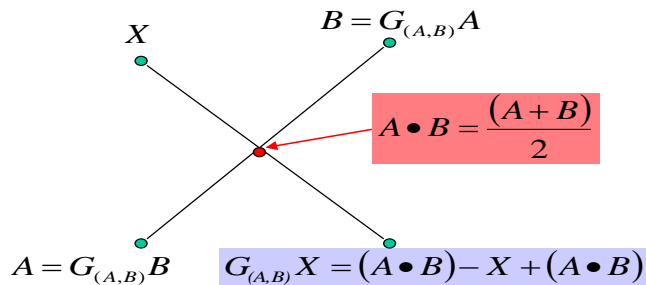


Jean-Louis Koszul
(hermitian canonical forms of complex homogeneous spaces, a complex homogeneous space with positive definite hermitian canonical form is isomorphic to a bounded domain, study of affine transformation groups of locally flat manifolds)

Etudes des espaces symétriques et leur classification

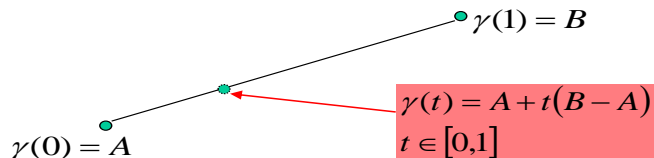
$$G_{(A,B)}X = (A \circ B)X^{-1}(A \circ B) \quad \text{avec} \quad A \circ B = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}$$

Espace Euclidien: isométrie

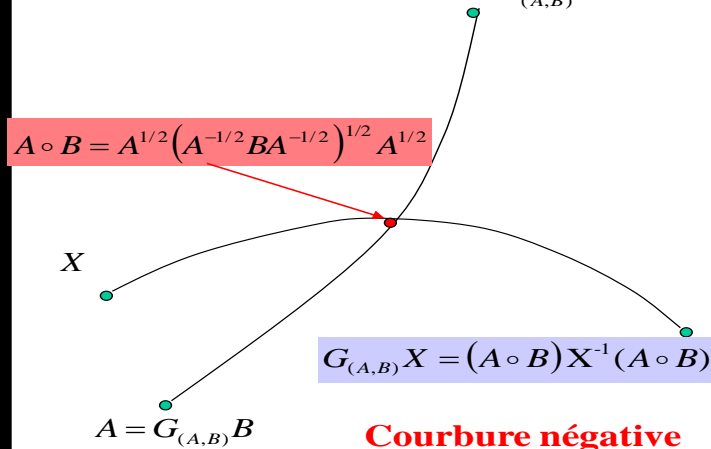


Courbure nulle

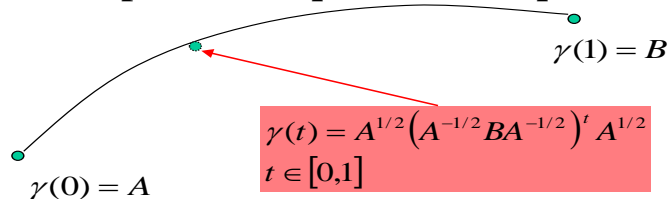
Espace Euclidien: Géodésique



**Espace métrique : isométrie
(espace matrices symétriques définies positives)**



Espace métrique : Géodésique



E. J. Cartan



M. Berger

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

Premier Miracle :

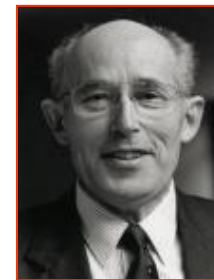
- La théorie des espaces symétriques peut être considérée comme le premier miracle de la géométrie riemannienne, en fait comme un nœud de forte densité dans l'arbre de toutes les mathématiques On doit à Elie Cartan dans les années 1926 d'avoir découvert que ces géométries sont , dans une dimension donnée,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

- 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 hyperkähleriennes**



M. Berger

Elementary Structure of Information Geometry

Information Geometry Metric

$$g^* = d^2\Psi^* = d^2S$$

$$g = -d^2 \log \Phi = d^2\Psi$$

$$ds^2 = d^2 \text{ENTROPY}$$

$$ds^2 = -d^2 \text{LOG[LAPLACE]}$$

Legendre Transform

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

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

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

$$x^* = \frac{d\Psi(x)}{dx}, \quad x = \frac{d\Psi^*(x^*)}{dx^*}, \quad x^* = \int_{\Omega^*} \xi \cdot p_x(\xi) d\xi$$

Laplace/Fourier Transform

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

$$\text{ENTROPY} = \text{LEGENDRE}(-\text{LOG[LAPLACE]})$$

$$\text{ENTROPY} = \text{FOURIER}_{(\text{Min}, *)}(-\text{LOG[FOURIER}_{(+, X)])}$$

Global Legendre Structure for Information Geometry (J.L. Koszul view)

$\langle \cdot, \cdot \rangle$ inner product from Cartan-Killing Form :

$$\langle \hat{\xi}, \beta \rangle = -B(\hat{\xi}, \theta(\beta)) \quad \text{with} \quad B(\hat{\xi}, \theta(\beta)) = \text{Tr}(ad_{\hat{\xi}} ad_{\theta(\beta)})$$



$$S(\hat{\xi}) = \langle \hat{\xi}, \beta \rangle - \Phi(\beta)$$

Legendre Transform

$$\Phi(\beta) = -\log \psi_{\Omega}(\beta)$$

$$S(\hat{\xi}) = -\int_{\Omega^*} p_{\hat{\xi}}(\xi) \log p_{\hat{\xi}}(\xi) . d\xi$$



$$\text{with} \quad \psi_{\Omega}(\beta) = \int_{\Omega^*} e^{-\langle \beta, \xi \rangle} d\xi$$

$$p_{\hat{\xi}}(\xi) = \frac{e^{-\langle \Theta^{-1}(\hat{\xi}), \xi \rangle}}{\int_{\Omega^*} e^{-\langle \Theta^{-1}(\hat{\xi}), \xi \rangle} . d\xi} \quad \hat{\xi} = \Theta(\beta) = \frac{\partial \Phi(\beta)}{\partial \beta}$$

$$\beta = \frac{\partial S(\hat{\xi})}{\partial \hat{\xi}}$$

$$I(\beta) = -E \left[\frac{\partial^2 \log p_{\beta}(\xi)}{\partial \beta^2} \right]$$

$$ds_g^2 = \sum_{ij} g_{ij} d\beta_i d\beta_j$$

$$ds_h^2 = \sum_{ij} h_{ij} d\hat{\xi}_i d\hat{\xi}_j$$

$$I(\beta) = -\frac{\partial^2 \Phi(\beta)}{\partial \beta^2}$$

$$\text{with} \quad g_{ij} = \left[\frac{\partial^2 \Phi(\beta)}{\partial \beta^2} \right]_{ij}$$

$$ds_g^2 = ds_h^2$$

$$\text{with} \quad h_{ij} = \left[\frac{\partial^2 S(\hat{\xi})}{\partial \hat{\xi}^2} \right]_{ij}$$

Application for Density of Symmetric Positive Definite Matrices

■ If we apply previous equation for Symmetric Positive Definite Matrices:

$$p_{\hat{\xi}}(\xi) = \frac{e^{-\langle \Theta^{-1}(\hat{\xi}), \xi \rangle}}{\int_{\Omega^*} e^{-\langle \Theta^{-1}(\hat{\xi}), \xi \rangle} . d\xi} \quad \hat{\xi} = \Theta(\beta) = \frac{\partial \Phi(\beta)}{\partial \beta} \quad \Phi(\beta) = -\log \psi_{\Omega}(\beta)$$

with $\psi_{\Omega}(\beta) = \int_{\Omega^*} e^{-\langle \beta, \xi \rangle} d\xi$

$$\langle \eta, \xi \rangle = \text{Tr}(\eta^T \xi), \quad \forall \eta, \xi \in \text{Sym}(n)$$

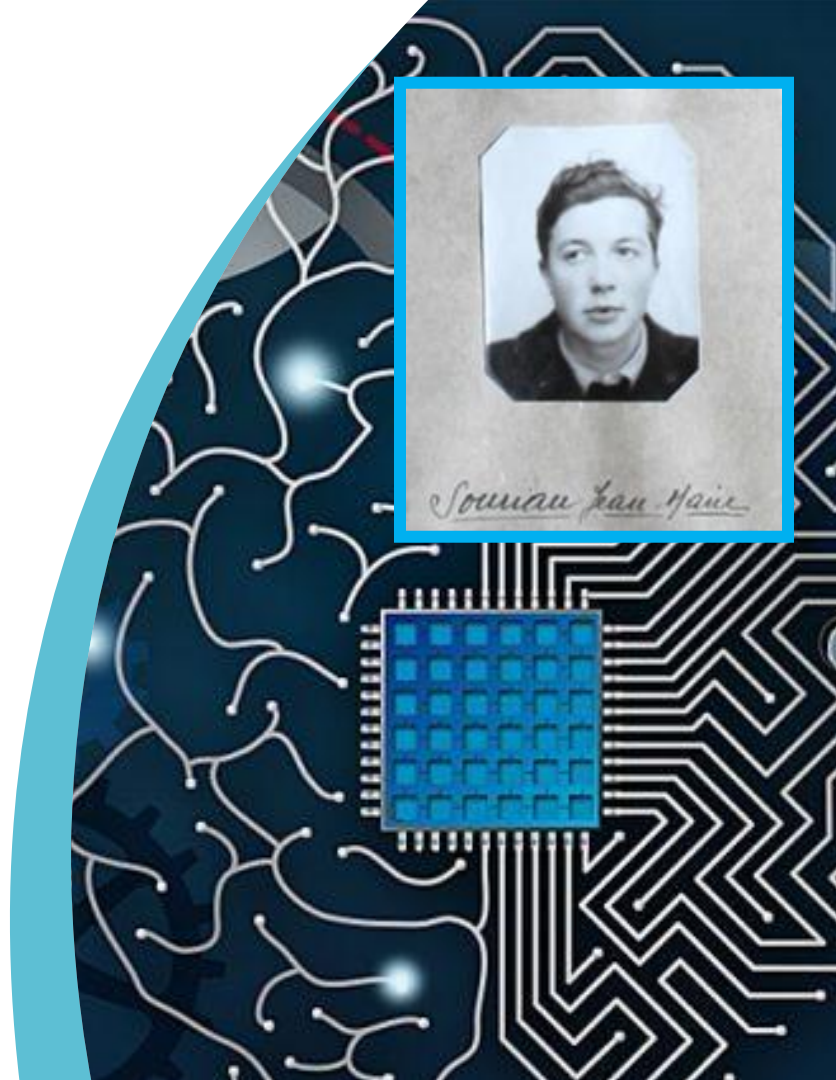
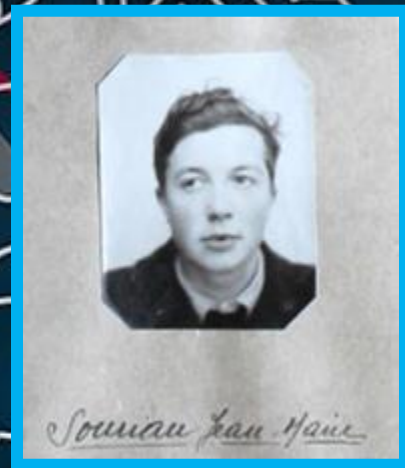
Application: $\psi_{\Omega}(\beta) = \int_{\Omega^*} e^{-\langle \beta, \xi \rangle} d\xi = \det(\beta)^{-\frac{n+1}{2}} \psi_{\Omega}(I_d)$

$$\hat{\xi} = \frac{\partial \Phi(\beta)}{\partial \beta} = \frac{\partial(-\log \psi_{\Omega}(\beta))}{\partial \beta} = \frac{n+1}{2} \beta^{-1}$$

$$p_{\hat{\xi}}(\xi) = e^{-\langle \Theta^{-1}(\hat{\xi}), \xi \rangle + \Phi(\Theta^{-1}(\hat{\xi}))} = \psi_{\Omega}(I_d) \cdot [\det(\alpha \hat{\xi}^{-1})] e^{-\text{Tr}(\alpha \hat{\xi}^{-1} \xi)} \quad \text{with } \alpha = \frac{n+1}{2}$$



Conclusion & perspectives



SOURIAU 2019

- Web site: <http://souriau2019.fr>
- In 1969, the groundbreaking book of Jean-Marie Souriau appeared "**Structure des Systèmes Dynamiques**". We will celebrate, in 2019, the jubilee of its publication, with a conference in honour of the work of this great scientist.
- Topics: Symplectic Mechanics, Geometric Quantization, Relativity & General Covariance, Thermodynamics, Cosmology, Diffeology, Philosophy
- Panel on Thermodynamics (including "Lie Groups Thermodynamics, Souriau-Fisher Metric")

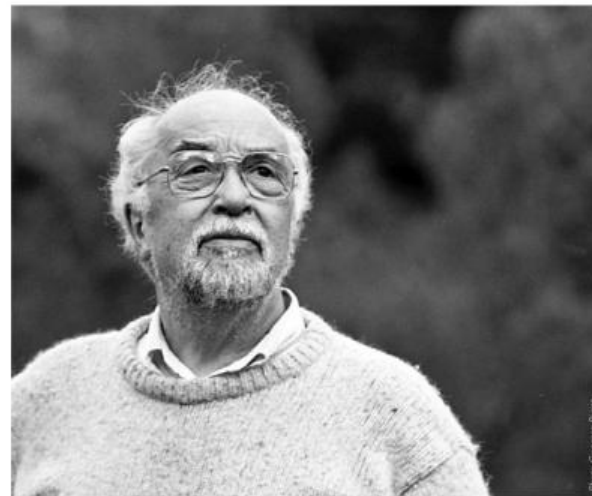


Photo: Georges Egger

- Fredéric Barbaresco
- Daniel Bernequand
- Jean-Pierre Bouguignon
- Pierre Cartier
- Dan Christensen
- Maurice Courbage
- Thibault Damour
- Paul Donato
- Paolo Giordano
- Serg Güler
- Patrick Iglesias-Zemmour
- Yael Karibon
- Jean-Pierre Magnot
- Yvette Koeman-Schwartzbach
- Marc Lichnerowicz
- Martin Pincussault
- Elsa Prato
- Urs Schweizer
- Jean-Jacques Szczeciniarz
- Roland Traay
- Jordan Watts
- Ervin Wu
- San-Wo Ngogo
- Alan Weinstein

JEAN-MARIE SOURIAU

In 1969, the groundbreaking book of Jean-Marie Souriau appeared "*Structure des Systèmes Dynamiques*". We will celebrate, in 2019, the jubilee of its publication, with a conference in honour of the work of this great scientist.

Symplectic Mechanics, Geometric Quantization, Relativity, Thermodynamics, Cosmology, Diffeology & Philosophy

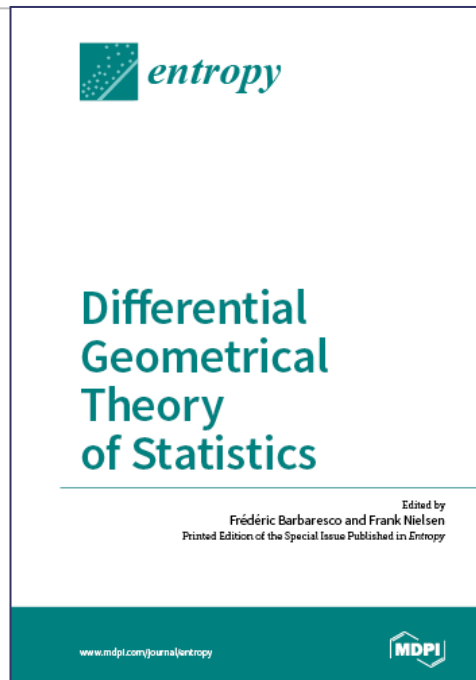
80|Prime



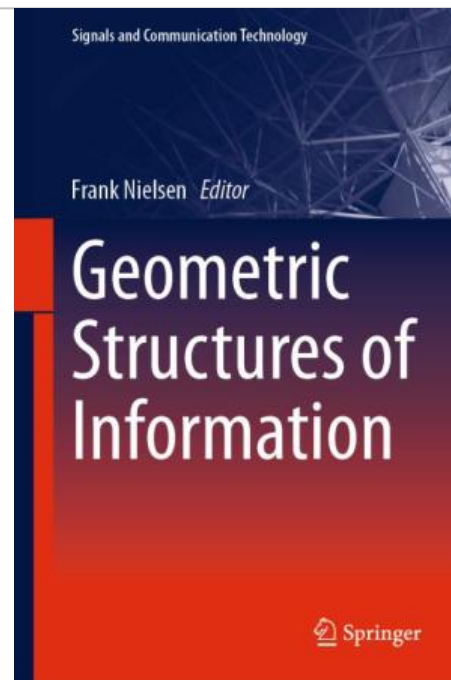
Books



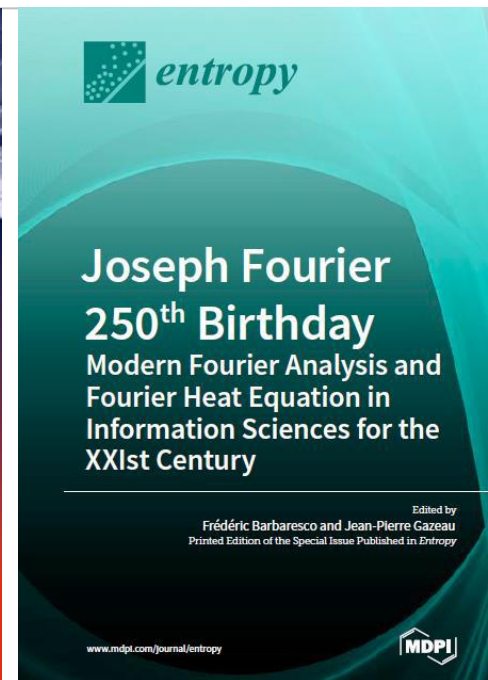
<https://www.mdpi.com/books/pdfview/book/127>



<https://www.mdpi.com/books/pdfview/book/313>

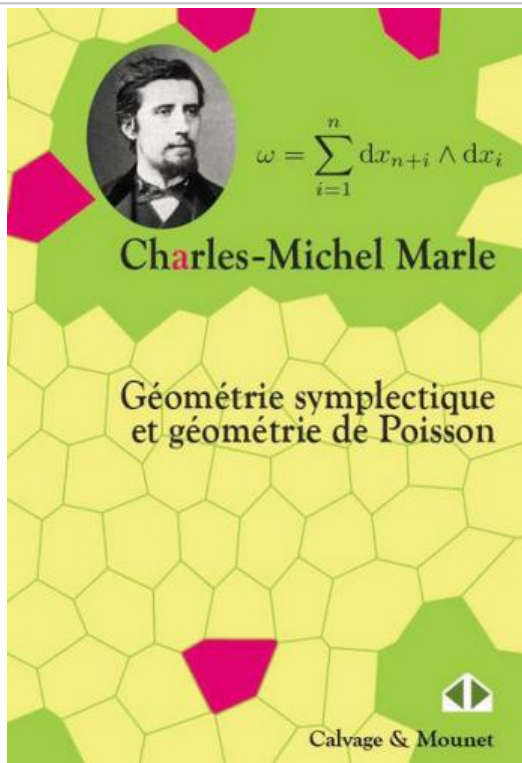


<https://www.springer.com/us/book/9783030025199>

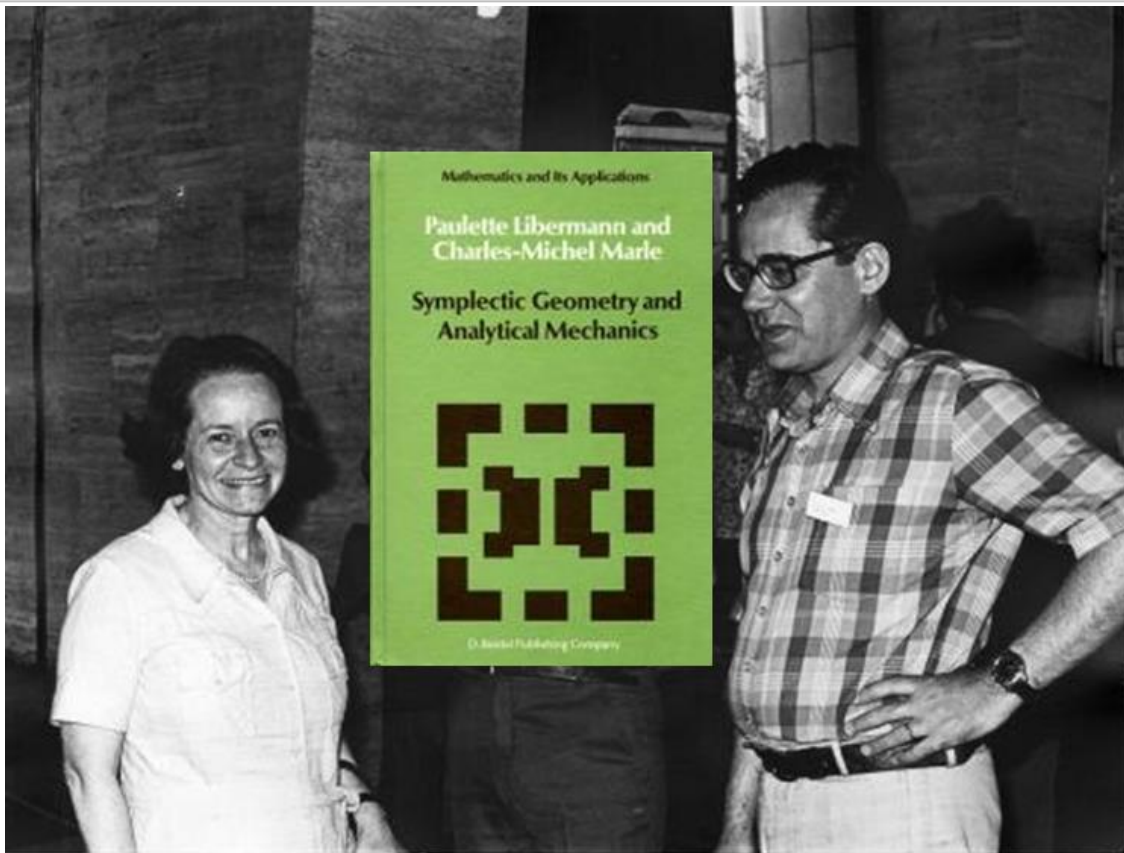


<https://www.mdpi.com/journal/entropy/special-issues/fourier>

Charles-Michel Marle Books

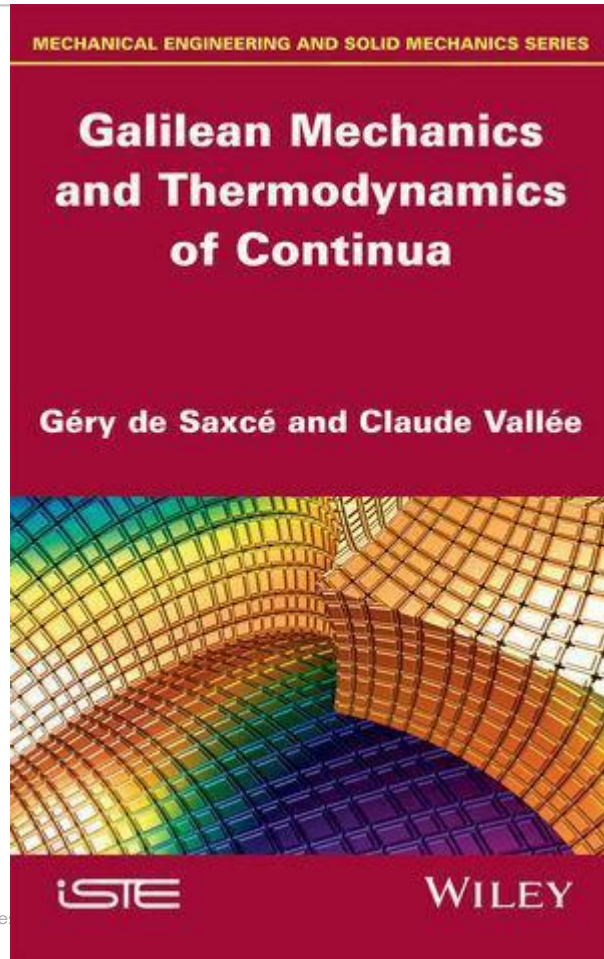


https://www.amazon.fr/gp/product/2916352708/ref=dbs_a_d_ef_rwt_bibl_vppl_i0



OPEN

Référence Book: Gery de Saxcé & Claude Vallée



This title proposes a unified approach to continuum mechanics which is consistent with Galilean relativity. Based on the notion of affine tensors, a simple generalization of the classical tensors, this approach allows gathering the usual mechanical entities — mass, energy, force, moment, stresses, linear and angular momentum — in a single tensor.

Starting with the basic subjects, and continuing through to the most advanced topics, the authors' presentation is progressive, inductive and bottom-up. They begin with the concept of an affine tensor, a natural extension of the classical tensors. The simplest types of affine tensors are the points of an affine space and the affine functions on this space, but there are more complex ones which are relevant for mechanics – torsors and momenta. The essential point is to derive the balance equations of a continuum from a unique principle which claims that these tensors are affine-divergence free.

<https://www.wiley.com/en-us/Galilean+Mechanics+and+Thermodynamics+of+Continua-p-9781848216426>

OPEN

CIRM Seminar, August 2017

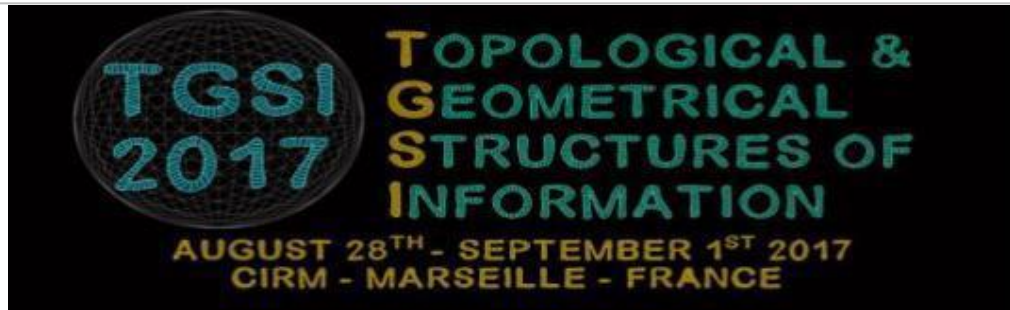
TGSI'17 « Topological & Geometrical Structures of Information »

TGSI'17 videos & slides

<http://forum.cs-dc.org/category/94/tgsi2017>

Special Issue "Topological and Geometrical Structure of Information", Selected Papers from CIRM conferences 2017"

http://www.mdpi.com/journal/entropy/special_issues/topological_geometrical_info



Talk on Koszul-Souriau Characteristic Function:

309 <https://www.youtube.com/watch?v=VXxiMCn-tsE&feature=youtu.be>

THALES

FGSI'19 Cartan-Koszul-Souriau Foundations of Geometric Structures of Information



<https://fgsi2019.sciencesconf.org/>
<https://fgsi2019.sciencesconf.org/resource/page/id/5>

FOUNDATIONS OF GEOMETRIC STRUCTURE OF INFORMATION

Cartan - Koszul - Souriau



Anton ALEKSEEV (Geneva Univ.)

Dmitri ALEKSEEVSKY (Moscow IITP)

John BAEZ (Riverside UC)

Michel BRION (Grenoble Univ.)

Misha GROMOV (Paris IHES)

Patrick IGLESIAS-ZEMMOUR
(Aix-Marseille Univ.)

Yann OLLIVIER (Paris Facebook)

Vasily PESTUN (Paris IHES)

Aissa WADE (Penn State Univ.)

Panel sessions: SYMPLECTIC GEOMETRY IN PHYSICS
TRIBUTE TO J-L KOSZUL & J-M SOURIAU

Registration/ Informations : <https://fgsi2019.sciencesconf.org/>

Contact : fgsi2019@sciencesconf.org

Scientific Committee: Frédéric Barbaresco (Thales), Damien Calaque (Montpellier Univ.), Michel N'Guillo-Boyon (Montpellier Univ.), Alan Weinstein (Berkeley Univ.)

Organizations: Jérémie Brieussel (Montpellier Univ.), Damien Calaque (Montpellier Univ.), Daniel Massart (Montpellier Univ.), Michel N'Guillo-Boyon (Montpellier Univ.)

4 - 6 FEBRUARY 2019

IMAG - Université de Montpellier

Trimester 2019 Labex CIMI, Toulouse « Statistics with Geometry & Topology »

Trimester « Statistics with Geometry & Topology », Toulouse, Aout-Sept. 2019

- Opening Event: **Geometric Science of Information** (GSI 19), 27-29 August 2019, ENAC
- **Geometric Statistics**, 30 Aout au 6 Septembre
- **Information Geometry**, 14 au 19 Octobre 2019, IMT
- **Topology for Learning and Data Analysis**, 29 Septembre-4 Octobre 2019, IMT
- **Computational Aspects of Geometry**, 6-8 Novembre 2019, IMT

<https://perso.math.univ-toulouse.fr/statistics-geometry-and-topology/>

Topics

- Information geometry
- Topology for learning and data
- Computational algebraic geometry
- Optimization and statistical applications

Mini courses

- M. BOYOM
- F. CHAZAL
- A. CUEVAS
- R. ELKAN
- P. MASSART
- B. MICHEL
- E. MILLER
- G. PISTONE
- X. PENNEC
- D. STEURER
- A. TROUVÉ
- C. ULLER

Thematic weeks

Aug 26	Aug 29	Geometric Science of Information (GSI-19)
Aug 30	Sep 6	Information geometry
Sep 29	Sep 4	Topology for Learning and Data Analysis
Nov 6	Nov 8	Computational Aspects of Geometry

Local Organisers:
F. Costantino, F. Gamboa, D. Henrion, T. Klein, A. Le Brigan, F. Nicol, E. Pauwels

Scientific Committee:
M. Arnaudon, F. Barbaresco, J. Bigot, A. Garivier, J.-B. Lasserre, X. Pennec, S. Puechmorel

WebPage:
<http://www.cimi.univ-toulouse.fr/>

Title: Structures jointes et socle commun de la physique statistique, de la géométrie de l'information et de l'inférence pour l'apprentissage

➤ Structures géométriques de la physique statistique et de l'Information

- Mécanique Statistique et Mécanique Géométrique
- Thermodynamique, Géométrie Symplectique et de Contact
- Thermodynamique des groupes de Lie
- Thermodynamique relativiste et des milieux continus
- Intégrateurs Symplectiques

➤ Structures physiques de l'inférence et l'apprentissage

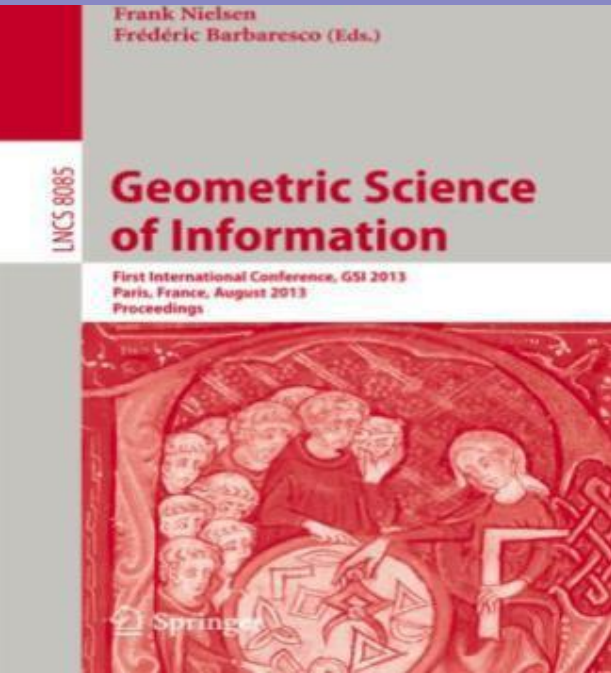
- gradient stochastique de la dynamique de Langevin
- géométrie de l'information, métrique de Fisher et gradient naturel
- méthodes de Monté-Carlo Hamiltonniennes
- inférence variationnelle et contrôles hamiltonien
- Machine de Boltzman



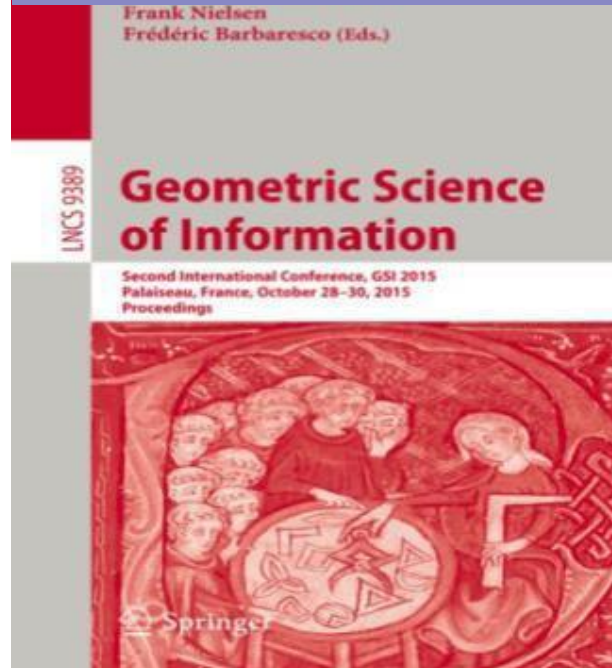
GSI SPRINGER PROCEEDINGS

LECTURE NOTES IN COMPUTER SCIENCE

GSI'13 Springer Proceedings:
<http://www.springer.com/us/book/9783642400193>

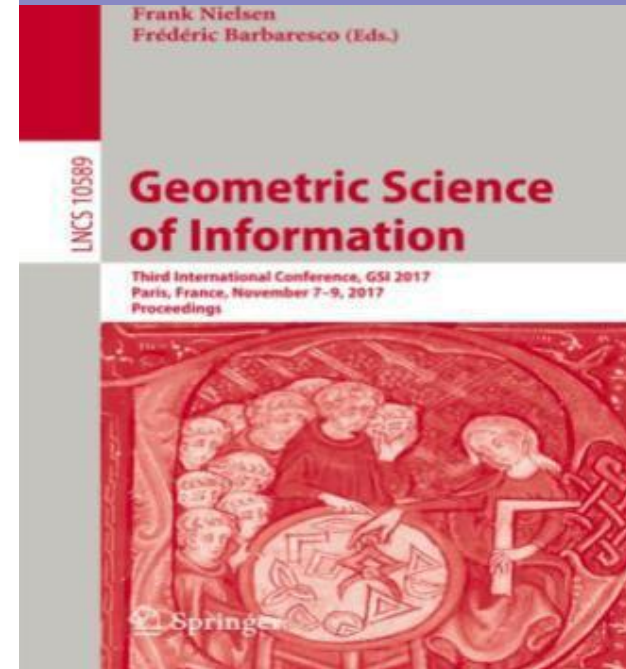


GSI'15 Springer Proceedings:
<http://www.springer.com/la/book/9783319250397>



OPEN

GSI'17 Springer Proceedings:
<http://www.springer.com/cn/book/9783319684444>



THALES

GSI'13

Mines ParisTech

Slides:

<https://www.see.asso.fr/gsi2013>



GSI'15

Ecole Polytechnique

Videos:

<https://www.youtube.com/channel/UC5HHo1jbQXusNQzU1iekaGA>

UNITWIN website (slides & videos):

<http://forum.cs-dc.org/category/90/gsi2015>

GSI'17: Mines ParisTech

150 attendees from 38 countries

Videos:

<https://www.youtube.com/channel/UCnE9-LbfFRqtaes49cN2DVg/videos>

Website: www.gsi2017.org



**Société
Mathématique
de France**



Koszul Book on Souriau Work

Jean-Louis Koszul · Yiming Zou

Introduction to Symplectic Geometry

Forewords by Michel Nguiffo Boyom, Frédéric Barbaresco and Charles-Michel Marle

This introductory book offers a unique and unified overview of symplectic geometry, highlighting the differential properties of symplectic manifolds. It consists of six chapters: Some Algebra Basics, Symplectic Manifolds, Cotangent Bundles, Symplectic G-spaces, Poisson Manifolds, and A Graded Case, concluding with a discussion of the differential properties of graded symplectic manifolds of dimensions (o, n) . It is a useful reference resource for students and researchers interested in geometry, group theory, analysis and differential equations.

$$\mu : M \longrightarrow \mathfrak{g}^*$$

$$\mu(sx) = s\mu(x) = \text{Ad}^*(s)\mu(x) + \varphi_\mu(s), \quad \forall s \in G, x \in M$$

$$c_\mu(a, b) = \{ \langle \mu, a \rangle, \langle \mu, b \rangle \} - \langle \mu, [a, b] \rangle = \langle d\varphi_\mu(a), b \rangle, \quad \forall a, b \in \mathfrak{g}$$

Jean-Louis Koszul
Yiming Zou

Introduction to Symplectic Geometry

$$\mu : M \longrightarrow \mathfrak{g}^*$$

$$\mu(sx) = s\mu(x) = \text{Ad}^*(s)\mu(x) + \varphi_\mu(s), \quad \forall s \in G, x \in M$$

$$c_\mu(a, b) = \{ \langle \mu, a \rangle, \langle \mu, b \rangle \} - \langle \mu, [a, b] \rangle = \langle d\varphi_\mu(a), b \rangle, \quad \forall a, b \in \mathfrak{g}$$

Science Press
Beijing

Springer

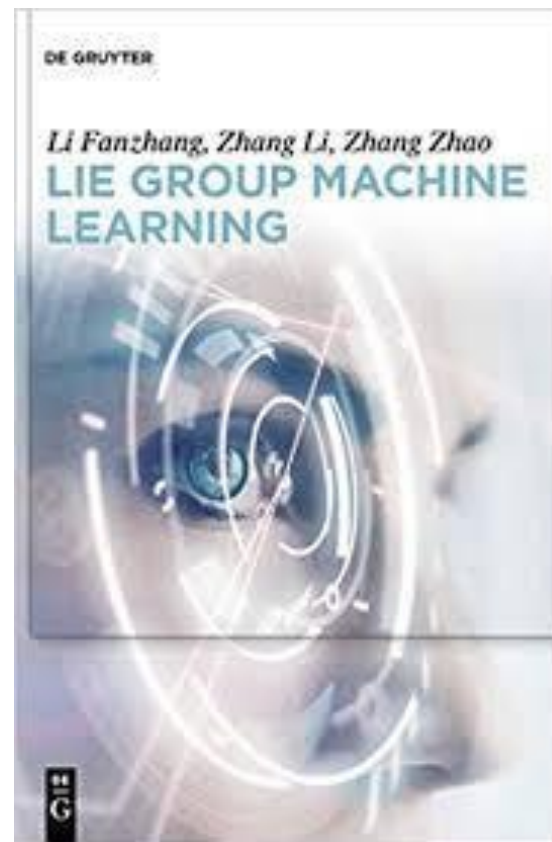
Lie Group Machine Learning (Chinese Book, De Gruyter, 2019)

Fanzhang Li, Li Zhang, Zhao Zhang

LIE GROUP MACHINE LEARNING

<https://www.degruyter.com/view/product/477343>

This book explains deep learning concepts and derives semi-supervised learning and nuclear learning frameworks based on cognition mechanism and Lie group theory. Lie group machine learning is a theoretical basis for brain intelligence, Neuromorphic learning (NL), advanced machine learning, and advanced artificial intelligence. The book further discusses algorithms and applications in tensor learning, spectrum estimation learning, Finsler geometry learning, Homology boundary learning, and prototype theory. With abundant case studies, this book can be used as a reference book for senior college students and graduate students as well as college teachers and scientific and technical personnel involved in computer science, artificial intelligence, machine learning, automation, mathematics, management science, cognitive science, financial management, and data analysis. In addition, this text can be used as the basis for teaching the principles of machine learning.





entropy



Joseph Fourier 250th Birthday

Modern Fourier Analysis and
Fourier Heat Equation in
Information Sciences for the
XXIst Century

Edited by
Frédéric Barbaresco and Jean-Pierre Gazeau
Printed Edition of the Special Issue Published in *Entropy*

www.mdpi.com/journal/entropy



Jean-Marie Souriau Geometric Theory of Heat, 250 years after Joseph Fourier

MDPI Entropy Book for Joseph Fourier 250th Birthday

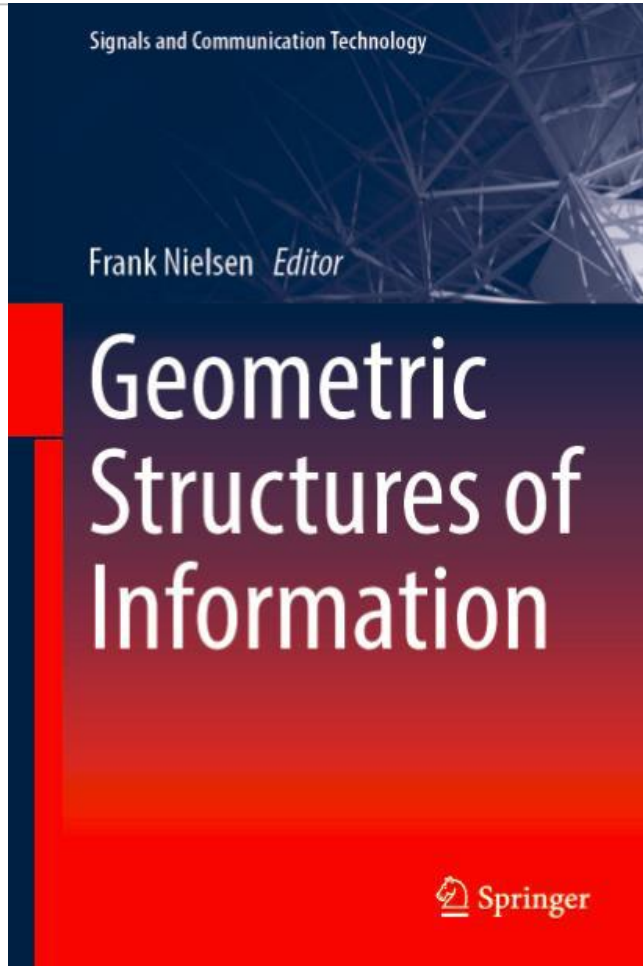
► https://www.mdpi.com/journal/entropy/special_issues/fourier

Jean-Marie Souriau Geometric Theory of Heat: Bedrock for Lie Group Machine Learning

► Barbaresco, F. : Higher Order Geometric Theory of Information and Heat Based on Poly-Symplectic Geometry of Souriau Lie Groups Thermodynamics and Their Contextures: The Bedrock for Lie Group Machine Learning. *Entropy*, 20, 840, (2018).

FOURIER 250.

THALES



Geometric Structures of Information

➤ <https://www.springer.com/us/book/9783030025199>

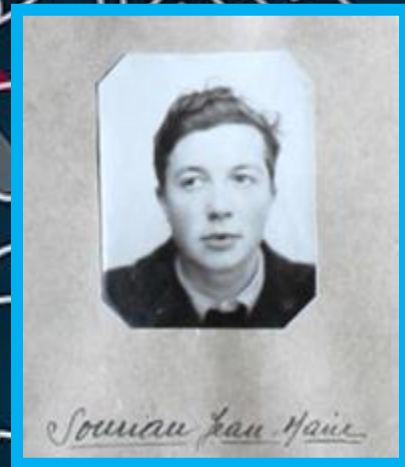
Paper on Jean-Louis Koszul

➤ Barbaresco, F. , Jean-Louis Koszul and the Elementary Structures of Information Geometry, Geometric Structures of Information, pp 333-392, SPRINGER, 2018

➤ https://link.springer.com/chapter/10.1007%2F978-3-030-02520-5_12



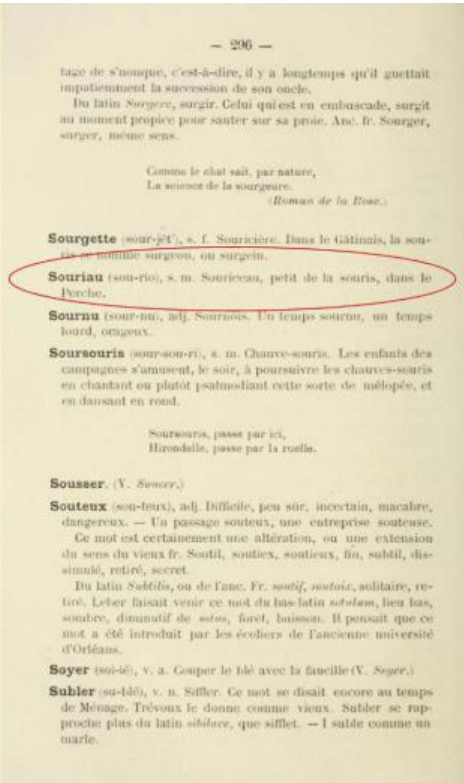
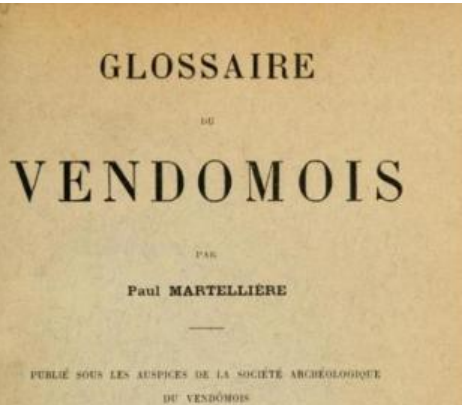
Souriau Genealogy



Origine du nom et de la famille Souriau

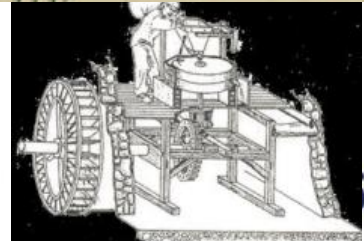
Origine du mot Souriau

- Petite souris (souriceau) dans le Perche



Origine de la famille Souriau

- Dans le Vendomois, les SOURIAU de 1490 à 1819 ont tous été "Maîtres" Laboureurs ou "Maîtres" Meuniers (moulins à eau)



Souriau Esthetism on « structure of motion » by 3 generation of ENS Ulm graduated students

Esthetism of motion

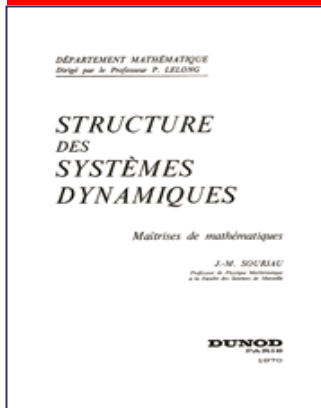


Paul Souriau
ENS Ulm 1873



Souriau Esthetism on « Structure of motion »

Structure of Motion

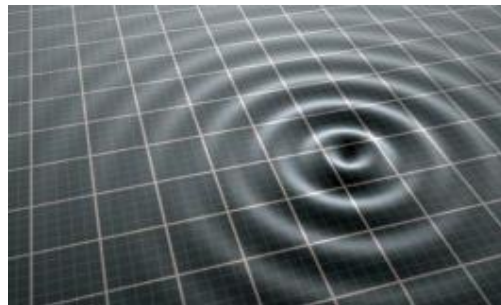


Jean-Marie Souriau
ENS Ulm 1942

Structure of Esthetism



Etienne Souriau
ENS Ulm 1912



Motivation :
caractérisation statistique de
la mesure digitale des
fluctuations de l'onde
électromagnétique radar
(première idée de Henri
Poincaré)

ALEAE GEOMETRIA
Géométrie du hasard

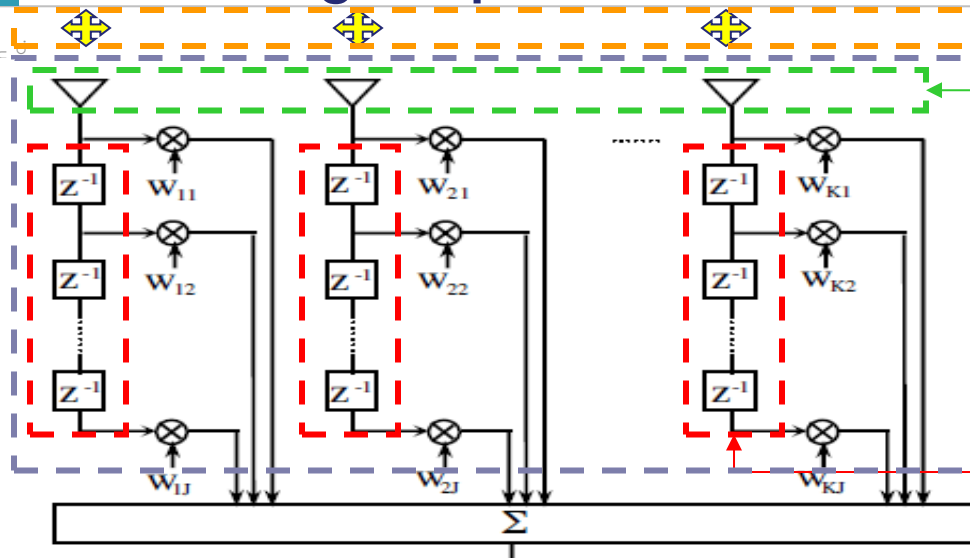


Blaise Pascal

Mesure digitale des fluctuations de l'onde électromagnétique radar

- L'idée est de « coder » l'état stationnaire d'une mesure digitale de l'onde électromagnétique (Direction, Doppler et Polarimétrie) par un point dans un espace métrique.
- Comme il s'agit d'une mesure statistique, la métrique naturelle est donnée en géométrie de l'information par le hessien de l'Entropie du processus stationnaire (métrique, potentiels et système de coordonnées duaux de la métrique de Fisher via la transformée de Legendre).
- En utilisant, les structures Toeplitz (respectivement Bloc-Toeplitz) des matrices de covariance du signal temporel [Doppler] ou spatial [direction] (respectivement spatio-temporelle), on démontre que la métrique naturelle est Kählérienne dans l'espace produit du polydisque de Poincaré (info. Doppler) et du polydisque de Siegel (info. spatio-temporel)

Mesure Digitale Spatio-Temporelle (et Polarimétrique) de l'Onde Electromagnétique Radar



Mesure digitale polarimétrique
(Matrice de covariance polarimétrique)

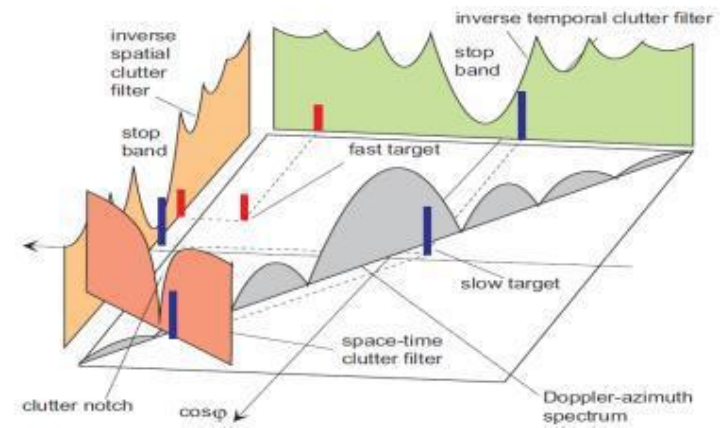
Mesures digitale spatio-temporelle
(Matrice de covariance spatio-temporelle)

Mesure digitale spatiale
(Matrice de covariance spatiale)

Mesure digitale temporelle
(Matrice de covariance Doppler)

Mesure digitale de l'onde EM

- Mesure polarimétrique
- Mesure temporelle => information Doppler
- Mesure spatiale => Direction du front d'onde
- Mesure spatio-temporelle



This document may not be repr

Approche historique de Henri Poincaré de 1891 pour la polarisation

Géométrie des états de polarisation

- Henri Poincaré montre qu'il est possible de représenter une vibration lumineuse polarisée elliptiquement, par l'affixe d'un nombre complexe: chaque état de polarisation de l'onde est représentée par un point sur le plan complexe
- La projection stéréographique du plan complexe sur une sphère de diamètre unité permet de représenter tous les états de polarisation d'une onde par un point et un seul sur une sphère
- ½ siècle plus tard, Jones et Muller proposèrent des méthodes de calcul matriciel, mais le modèle de Poincaré reste le plus élégant et efficace.



CHAPITRE XII POLARISATION ROTATOIRE.—THÉORIE DE M. MALLARD

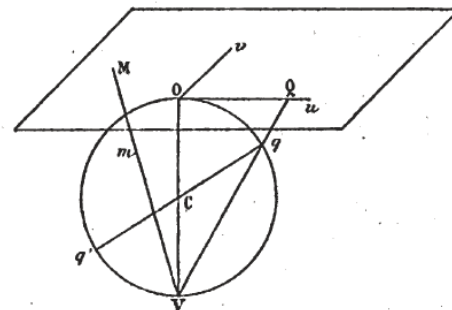


Fig. 42.

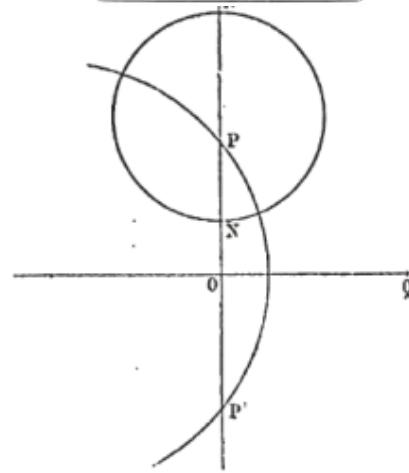


Fig. 41.

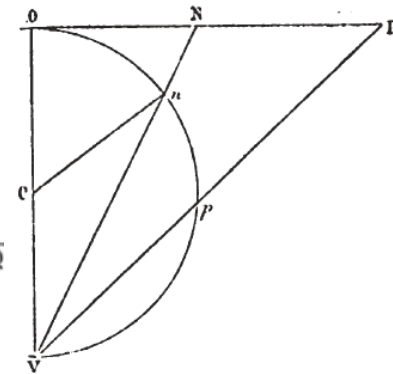


Fig. 43

Modélisation géométrique de Poincaré des états de polarisation

Représentation de l'état de polarisation par un point du plan complexe

- Champ Electromagnétique E est décrit par:

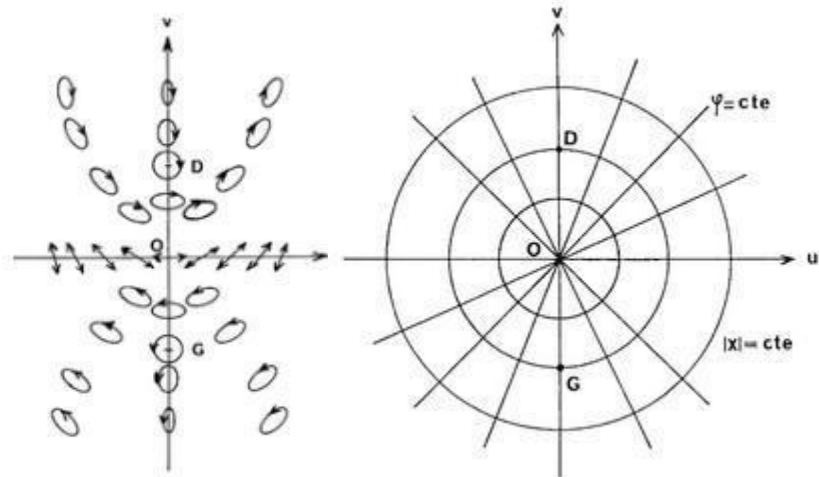
$$E(z, t) = \text{Re} \left[A e^{i(kz - \omega t)} \right] = E_x \vec{i} + E_y \vec{j}$$

$$\text{avec } A = A_x e^{i\varphi_x} \vec{i} + A_y e^{i\varphi_y} \vec{j}$$

$$\chi = \frac{A_y e^{i\varphi_y}}{A_x e^{i\varphi_x}} = \frac{A_y}{A_x} e^{i(\varphi_y - \varphi_x)} = u + iv$$

- La valeur de ce rapport nous fait connaître la forme de l'ellipse τ et son orientation ϕ décrite par le vecteur électromagnétique dans le plan perpendiculaire à l'axe de propagation et décrivant l'état polarimétrique de l'onde:

$$|\chi| = \frac{A_y}{A_x} = \tan \tau \quad \text{et} \quad \arg \chi = \phi = \varphi_y - \varphi_x \quad \chi = \frac{\tan \phi + i \tan \tau}{1 - i \tan \phi \tan \tau}$$



Modélisation géométrique de Poincaré des états de polarisation

Représentation de l'état de polarisation par un point du plan complexe

➤ Ellipse de polarisation

$$\mathbf{E}_M = \begin{bmatrix} E_{M,x}(t) \\ E_{M,y}(t) \end{bmatrix} = \begin{bmatrix} A_x(t) \cdot e^{i\varphi_x(t)} \\ A_y(t) \cdot e^{i\varphi_y(t)} \end{bmatrix} = p e^{i\varphi} \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} \cos \tau \\ i \sin \tau \end{bmatrix}$$

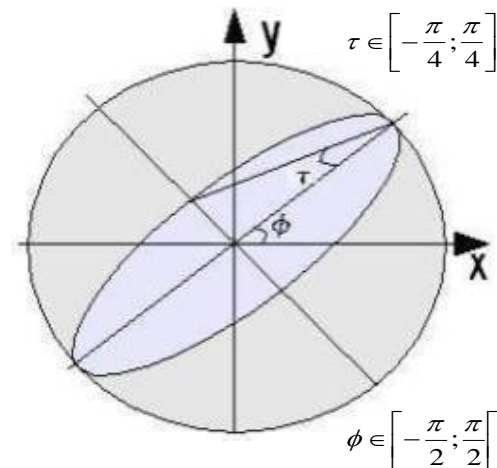
- τ : angle d'éllipticité caractérisant l'élongation (la forme)
- ϕ : caractérise l'inclinaison de l'ellipse

➤ Matrice de cohérence de Jones:

$$\mathbf{J} = \mathbf{E}(\mathbf{E}_M \cdot \mathbf{E}_M^+) = \begin{bmatrix} E(|E_{M,x}|^2) & E(E_{M,y} E_{M,x}^*) \\ E(E_{M,x} E_{M,y}^*) & E(|E_{M,y}|^2) \end{bmatrix} = \begin{bmatrix} J_{11} & J_{21} \\ J_{12} & J_{22} \end{bmatrix}$$

➤ Vecteurs de Stokes:

$$\vec{\mathbf{S}} = \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} J_{11} + J_{22} \\ J_{11} - J_{22} \\ 2 \operatorname{Re}(J_{21}) \\ -2 \operatorname{Im}(J_{12}) \end{bmatrix} = \begin{bmatrix} s_0 \\ s_0 \cos 2\tau \cos 2\phi \\ s_0 \cos 2\tau \sin 2\phi \\ s_0 \sin 2\tau \end{bmatrix}$$



$$\phi = \frac{1}{2} \arctan \left(\frac{s_2}{s_1} \right)$$

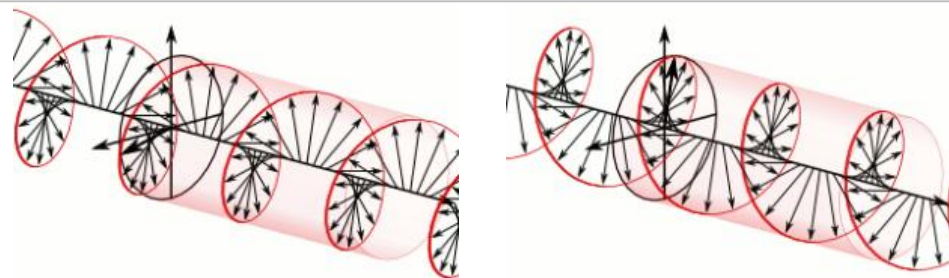
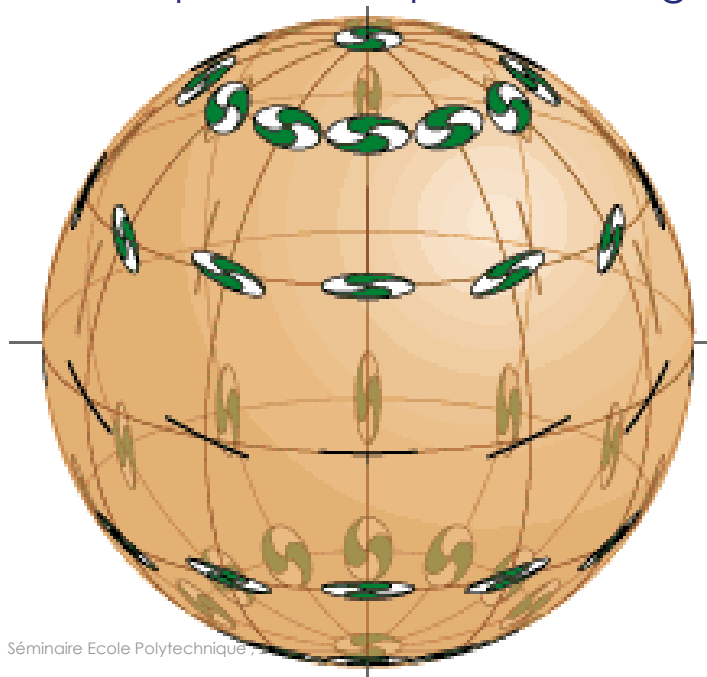
$$\tau = \frac{1}{2} \arctan \left(\frac{s_3}{\sqrt{s_1^2 + s_2^2}} \right)$$

$$s_0^2 \geq s_1^2 + s_2^2 + s_3^2$$

Information polarimétrique: représentée par un point sur la sphère

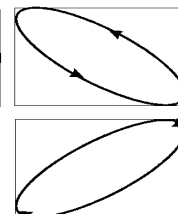
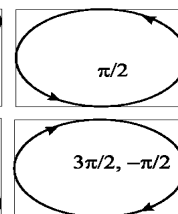
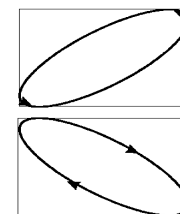
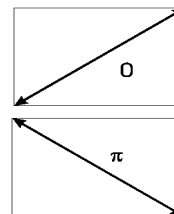
Coordonnées sphériques

- Sur la sphère: $(2\phi, 2\tau)$
- Hémisphère Nord: polarisation droite
- Hémisphère Sud: polarisation gauche

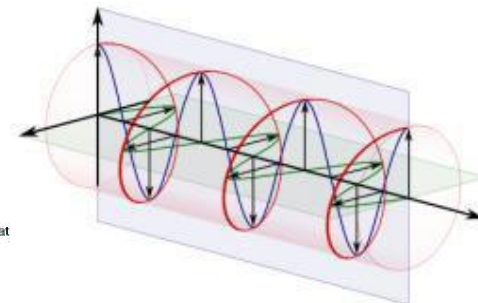
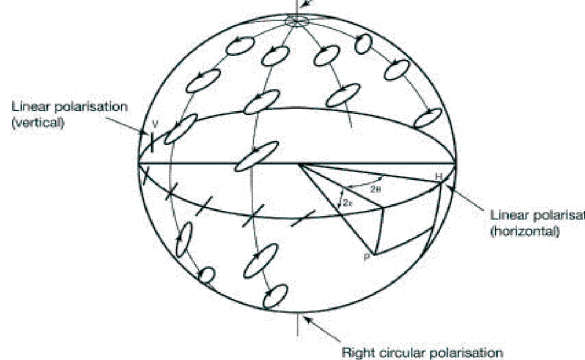


Polarisation circulaire gauche

Polarisation circulaire droite

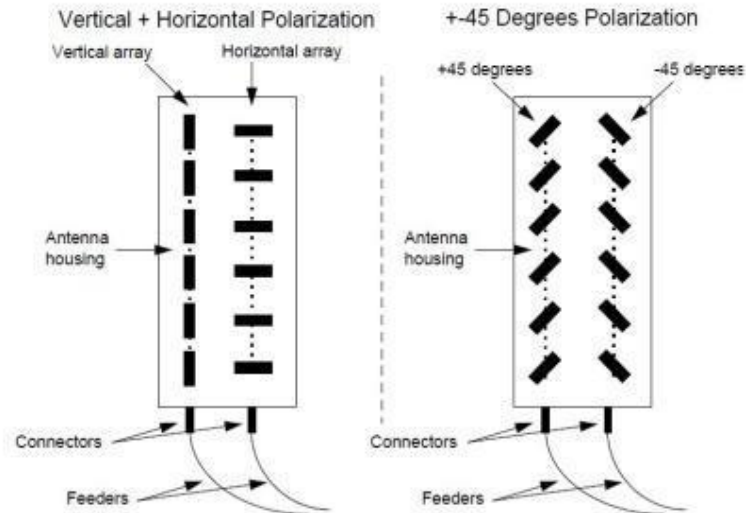
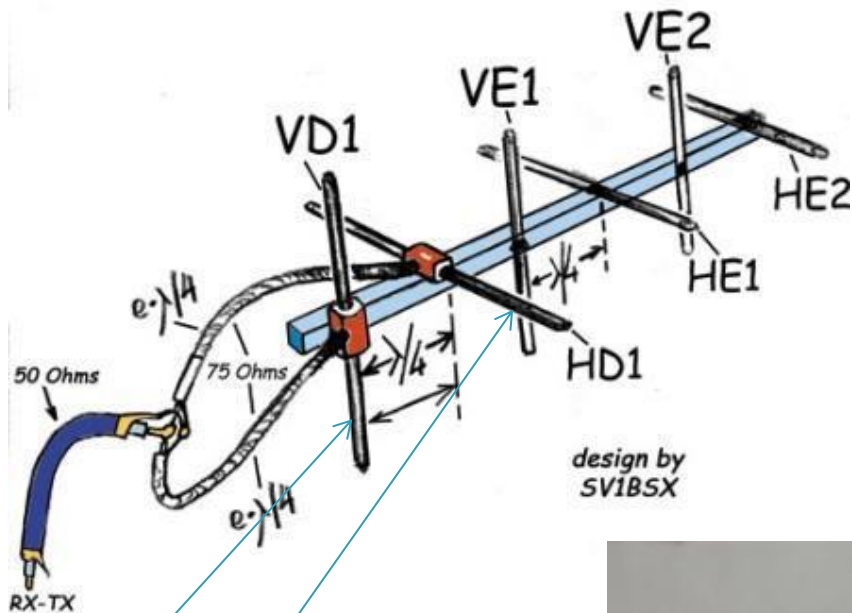


Left circular polarisation

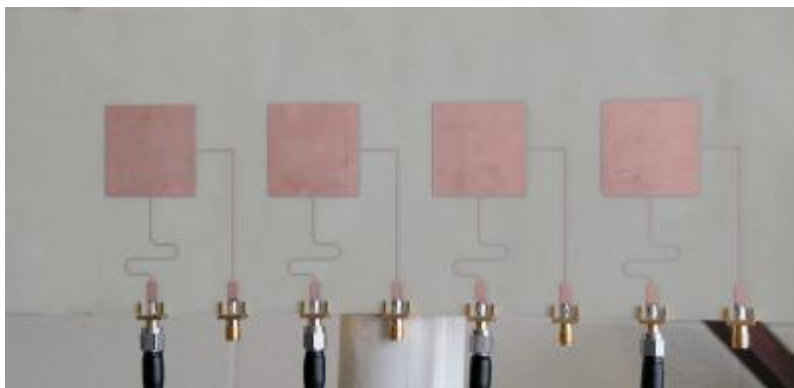


Exemples d'Antennes Polarimétriques

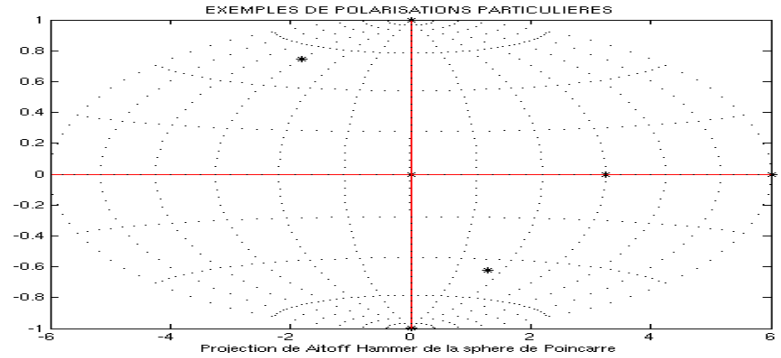
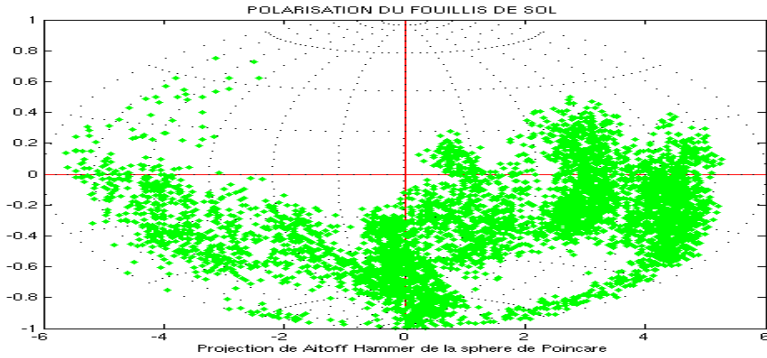
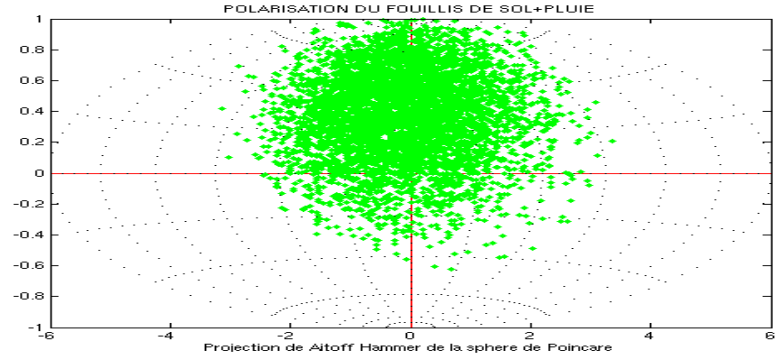
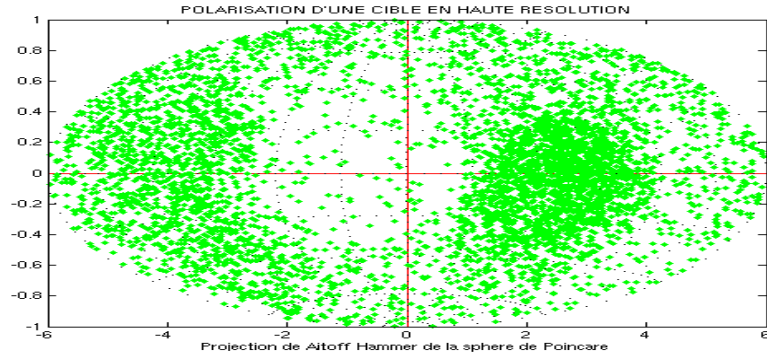
This document may not be reproduced, modified, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales. - © Thales 2015 All rights reserved.



$$E_M = \begin{bmatrix} E_{M,x}(t) \\ E_{M,y}(t) \end{bmatrix} = \begin{bmatrix} A_x(t) \cdot e^{i\varphi_x(t)} \\ A_y(t) \cdot e^{i\varphi_y(t)} \end{bmatrix}$$

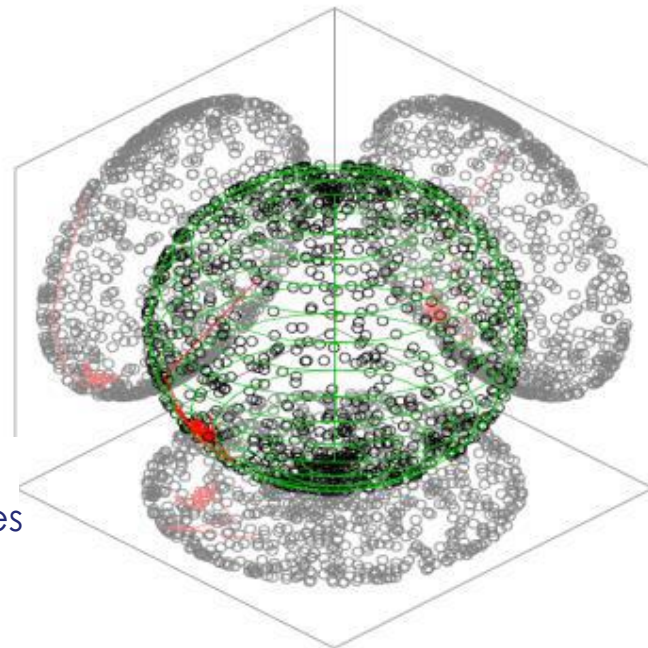
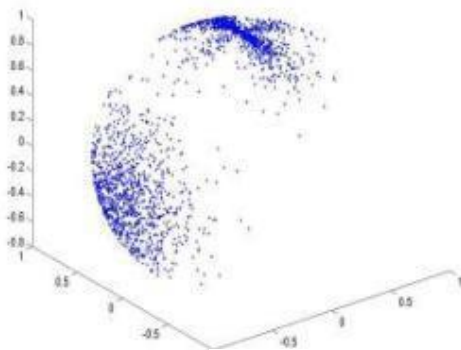


Etat polarimétrique: Projection de Mercator de la sphère de Poincaré

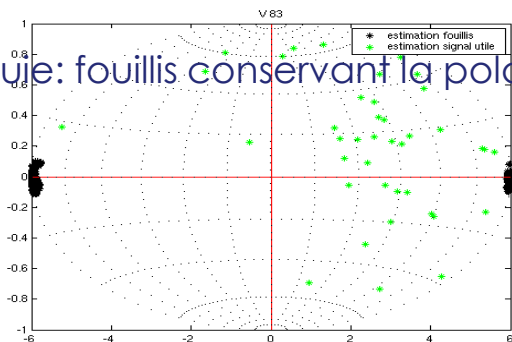


This document may not be reproduced, modified, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales. - © Thales. 2015 All rights reserved.

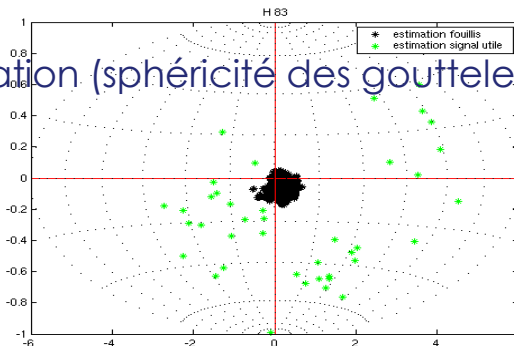
Comment décrire la statistique de fluctuation de l'état polarimétrique ?



Pluie: fouillis conservant la polarisation (sphéricité des gouttelettes)



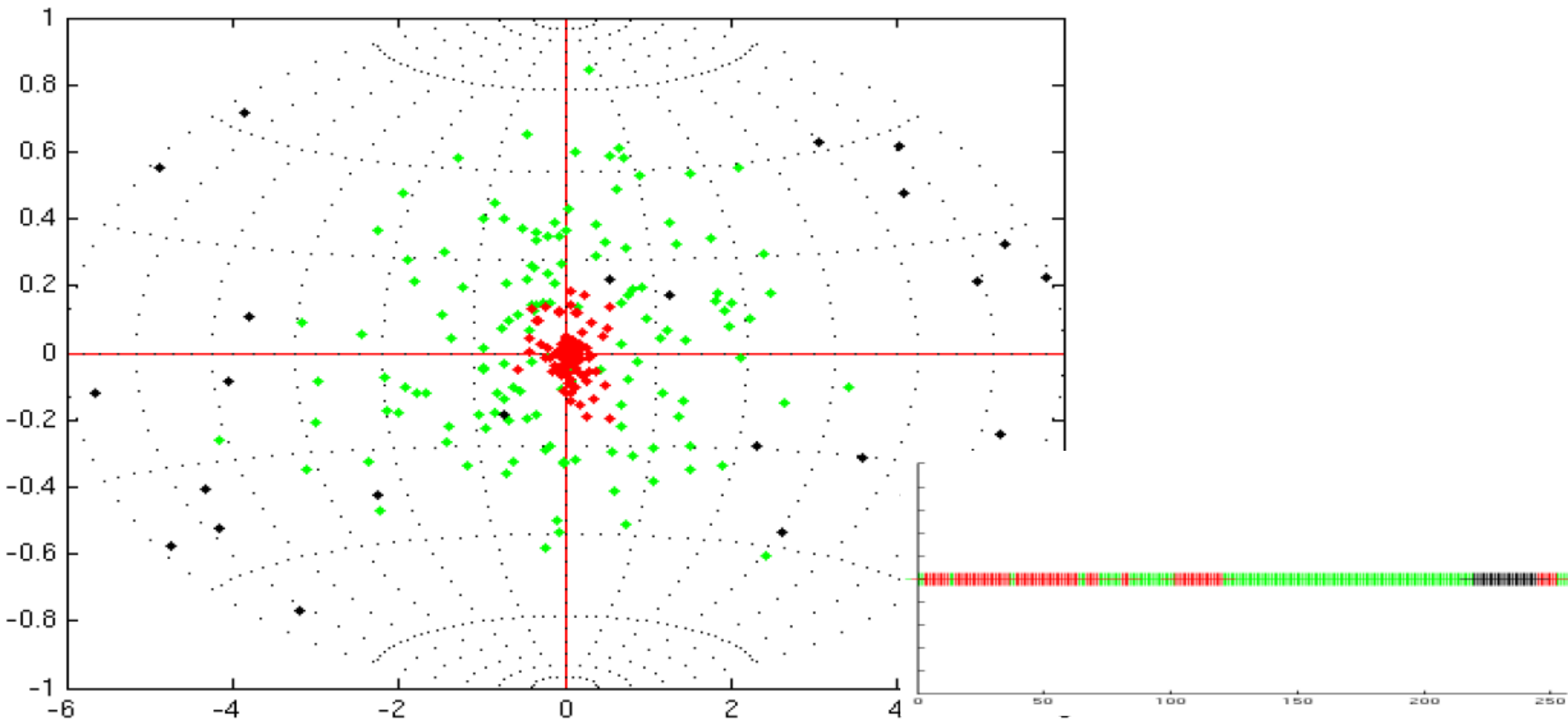
Pluie: Émission verticale



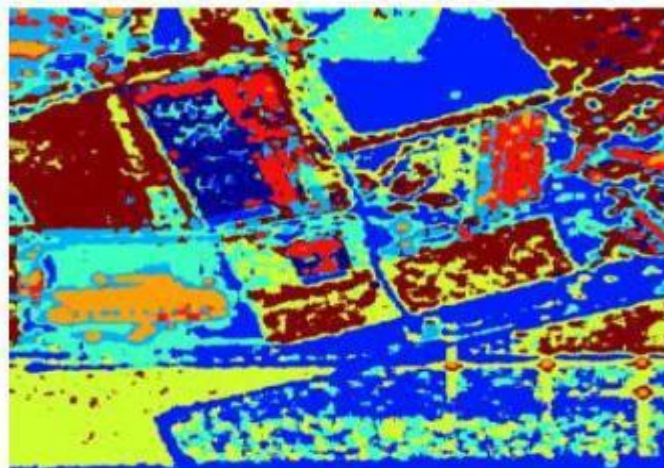
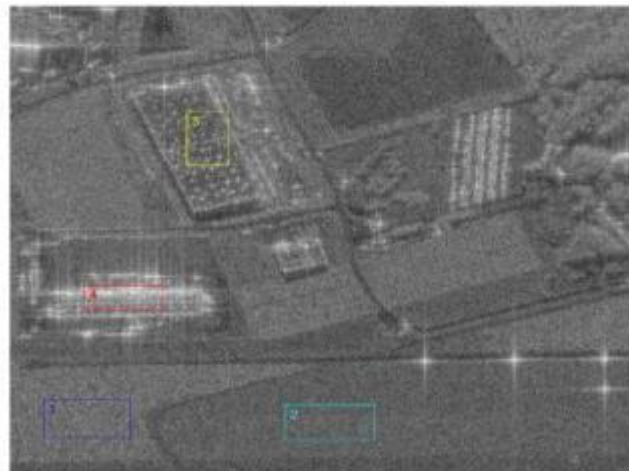
Pluie: Émission horizontale

Segmentation/classification de données polarimétriques

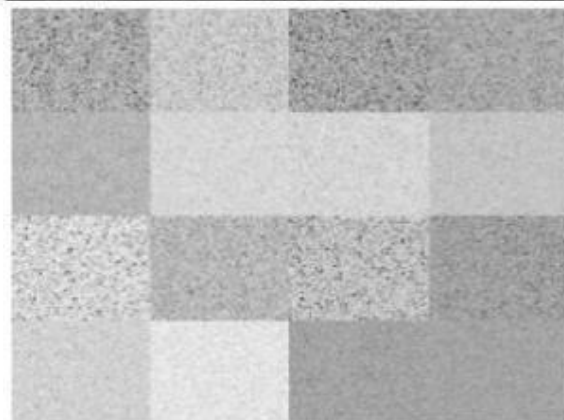
This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - © Thales 2015 All rights reserved.



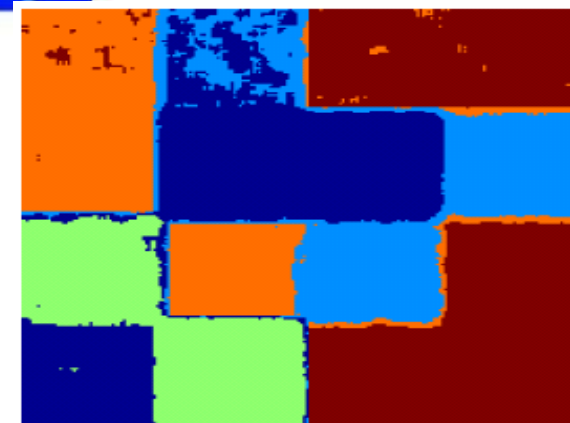
Pierre Formont (Thèse ONERA): Segmentation des données polarimétriques



Segmentation sur la distance au sens de la géométrie de l'information



$\sim \mathcal{K}(M_1, \lambda_1)$	$\sim \mathcal{K}(M_1, \lambda_2)$	$\sim \mathcal{K}(M_2, \lambda_1)$	$\sim \mathcal{K}(M_2, \lambda_2)$
$\sim \mathcal{K}(M_1, \lambda_3)$	$\sim \mathcal{K}(M_1, \lambda_4)$	$\sim \mathcal{K}(M_2, \lambda_3)$	$\sim \mathcal{K}(M_2, \lambda_4)$
$\sim \mathcal{K}(M_3, \lambda_1)$	$\sim \mathcal{K}(M_3, \lambda_2)$	$\sim \mathcal{K}(M_4, \lambda_1)$	$\sim \mathcal{K}(M_4, \lambda_2)$
$\sim \mathcal{K}(M_3, \lambda_3)$	$\sim \mathcal{K}(M_3, \lambda_4)$	$\sim \mathcal{K}(M_4, \lambda_3)$	$\sim \mathcal{K}(M_4, \lambda_4)$

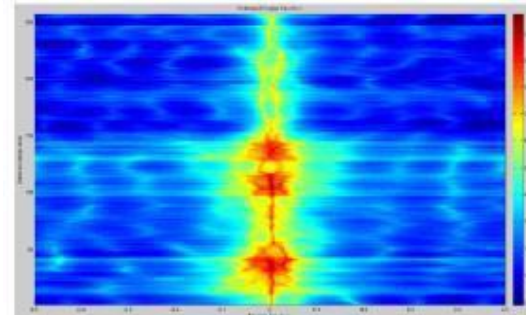


SCM - GM

Les éléments aléatoires de nature quelconque dans un espace distancié pour caractériser la fluctuation du signal électromagnétique

Dans la mesure digitale de l'onde électromagnétique, il s'agit de caractériser les fluctuations :

- **En amplitude:** variation du signal en puissance
- **Sur la polarimétrie:** variation de la polarisation de l'onde
- **Spatialement:** variation de la direction d'arrivée du front d'onde
- **Temporellement:** Variation du spectre Doppler



Spectre Doppler/Distance

Exemple du signal radar Doppler relatif à la série temporelle de mesures

Mesure de la série temporelle du signal digitalisé (pour une direction donnée)

$$Z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \text{ avec } z_i \in \mathbb{C}$$

$$R_n = E[ZZ^+] = \begin{bmatrix} r_0 & r_1^* & \cdots & r_{n-1}^* \\ r_1 & r_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_1^* \\ r_{n-1} & \cdots & r_1 & r_0 \end{bmatrix}$$

Stationnaire \rightarrow R Toeplitz

$$\text{avec } r_k = E[z_m z_{m-k}^*]$$

avec r_k : coefficient de corrélation dont on déduit le spectre Doppler

$$|S_Z(f)|^2 = \left| \sum_{k=-(n-1)}^{n-1} r_k e^{-j2\pi f k} \right|^2$$

Etude du cas le plus simple: matrice THDP de taille 2x2

Exemple sur matrice 2x2 Toeplitz Hermitienne définie positive:

➤ 1 matrice de covariance d'un signal stationnaire 2x2 peut être représentée par « 1 point » sur une variété dans $R_+^* \times D$ (avec D le disque de Poincaré unité)

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

avec $a, b \in R$ et $h \in R_+^*$

$$\Omega^+ = \Omega$$

$$\Omega = h \cdot \begin{bmatrix} 1 & \mu^* \\ \mu & 1 \end{bmatrix} \text{ avec } \mu = \frac{a + ib}{h}$$

$$|\mu|^2 = \frac{a^2 + b^2}{h^2} < 1 \text{ car } h^2 > a^2 + b^2$$

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

$$\det(\Omega) = h^2 - (a + ib)(a - ib)$$

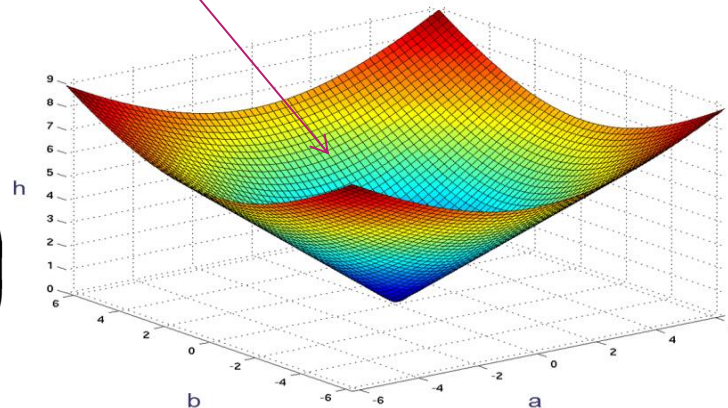
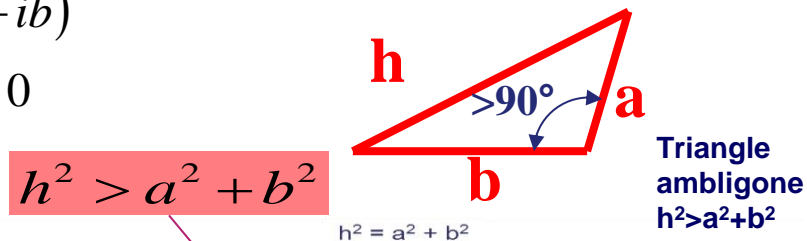
$$\det(\Omega) = h^2 - (a^2 + b^2) > 0$$

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

Doppler moyen

φ_{Doppler} : phase de μ

$$\varphi_{\text{Doppler}} = \arctan\left(\frac{b}{a}\right)$$



Quelles sont les contraintes entre les paramètres de la matrice de covariance d'un signal complexe (circulaire) stationnaire

Les paramètres d'une matrice de covariance d'une série temporelle d'un signal complexe (circulaire) stationnaire sont contraints par :

➤ Structure Toeplitz (éléments égaux sur la diagonale) : $\forall n, E[z_n z_{n-k}^*] = r_k$

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



Otto Toeplitz

➤ Structure Hermitienne :

$$R_n^+ = R_n \text{ avec } + : \text{transposée et conjuguée}$$

➤ Structure définie positive (valeurs propres positives) :

$$\forall Z \in C^n, Z^+ R_n Z > 0 \text{ et } \lambda_i > 0 \quad i=1, \dots, n \text{ avec } \det(R_n - \lambda I) > 0$$

Théorème de Trench/Verblunsky (décomposition d'Iwasawa partielle) pour les matrices Toeplitz Hermitiennes Définies Positives

Structure: matrice de covariance Toeplitz Hermitienne Définie Positive THDP

➤ Toute matrice THDP est diffeomorphe à l'espace produit $(P_0, \mu_1, \dots, \mu_n) \in \mathbb{R}^+ \times D_n$:

- P_0 est un paramètre d'échelle (puissance du signal)
- μ_k sont des paramètres de forme (forme du spectre) appelés coefficients de réflexion/Verblunsky

$$\begin{aligned} \varphi : THDP(n) &\rightarrow \mathbb{R}_+^* \times D^{n-1} \\ R_n &\mapsto (P_0, \mu_1, \dots, \mu_n) \\ \text{avec } D &= \{z \in \mathbb{C} / |z| < 1\} \end{aligned}$$

➤ Cette décomposition par bloc est liée à la décomposition d'Iwasawa partielle:

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

$$A_n = \begin{bmatrix} A_{n-1} \\ 0 \end{bmatrix} + \mu_n \begin{bmatrix} A_{n-1}^{(-)} \\ 1 \end{bmatrix}$$

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

$$V^{(-)} = J.V^*$$

Théorème de Trench/Verblunsky et décomposition de Cholesky

■ L'itération de la structure bloc de la matrice:

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

Génère la décomposition de (André-Louis) Cholesky:

$$\Omega_n = (\alpha_n R_n)^{-1} = \Omega_n^{1/2} \cdot \Omega_n^{1/2+}$$

$$\Omega_n^{1/2} = \sqrt{1 - |\mu_n|^2} \begin{bmatrix} 1 & 0_{n-1}^+ \\ A_{n-1} & \Omega_{n-1}^{1/2} \end{bmatrix}$$

A. Cholesky, Sur la résolution numérique des systèmes d'équations linéaires, Manuscrit, Fonds A. Cholesky, Archives de l'École Polytechnique, Palaiseau.



A.L. Cholesky
(Commandant d'Artillerie Cholesky
tué pendant la grande guerre)



Machine à
calculer
Dactyle

THALES

Métrique de Fisher de la géométrie de l'Information

■ Nous considérons comme métrique, le **hessien de l'Entropie**:

➤ Entropie: $S(\bar{R}_n) = \log(\det \bar{R}_n^{-1}) - \log(\pi.e)$

➤ Développement de l'entropie en utilisant sa matrice blocs et

$$\det \begin{bmatrix} P & q^+ \\ q & Q \end{bmatrix} = p \cdot \det(Q - qq^+)$$

$$\bar{R}_n^{-1} = \begin{bmatrix} \bar{\alpha}_{n-1} & \bar{\alpha}_{n-1} \bar{A}_{n-1}^+ \\ \bar{\alpha}_{n-1} \bar{A}_{n-1} & \bar{R}_{n-1}^{-1} + \bar{\alpha}_{n-1} \bar{A}_{n-1} \bar{A}_{n-1}^+ \end{bmatrix} \quad \bar{\alpha}_n^{-1} = \left[1 - |\bar{\mu}_n|^2 \right] \bar{\alpha}_{n-1}^{-1}$$

$$\bar{P}_0 = \bar{\alpha}_0^{-1}$$



$$S(\bar{R}_n) = - \sum_{k=1}^{n-1} (n-k) \log \left[1 - |\bar{\mu}_k|^2 \right] - n \log [\pi.e.\bar{P}_0]$$

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

Exemple de l'article de 1932 de E. Kähler et appelé cas hyper-abélien



E. Kähler

■ La paramétrisation est donnée par:

$$\theta^{(n)} = \left[\bar{P}_0 \quad \bar{\mu}_1 \quad \dots \quad \bar{\mu}_{n-1} \right]^T = E \left[\left[P_0 \quad \mu_1 \quad \dots \quad \mu_{n-1} \right]^T \right]$$

■ La métrique au sens de la géométrie de l'information est donné par le hessien de l'Entropie:

$$g_{ij} = \frac{\partial^2 S}{\partial \theta_i \partial \theta_j^*}$$

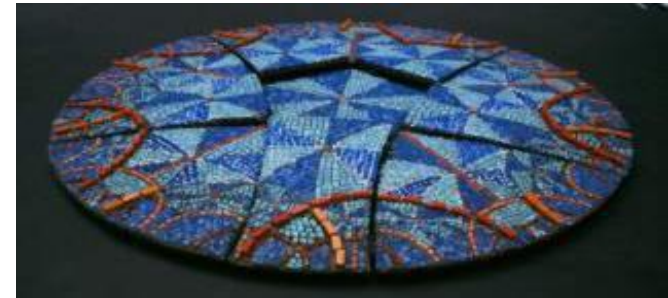
$$ds_{dual}^2 = \sum_{i,j} g_{ij}^{dual} d\bar{\theta}_i^{(n)} d\bar{\theta}_j^{(n)*} = n \cdot \left(\frac{d\bar{P}_0}{\bar{P}_0} \right)^2 + \sum_{i=1}^{n-1} (n-i) \frac{|d\bar{\mu}_i|^2}{\left(1 - |\bar{\mu}_i|^2 \right)^2}$$

Le disque de Poincaré: géométrie naturelle des gaussiennes

This document may not be reproduced, modified, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales. - © Thales 2015 All rights reserved.



HENRI POINCARÉ DANS SON CABINET DE TRAVAIL. — 1904. —



Introduction du disque de Poincaré comme modèle de la géométrie hyperbolique (courbure négative constante)

sd, in any way, in whole or in
 parts. 2015 All rights reserved.



$$ds^2 = \frac{dx^2}{y^2} + \frac{dy^2}{y^2} = \frac{|dz|^2}{\text{Im}(z)} \text{ avec } z = x + iy$$

Invariant par les automorphismes du demi-plan $M(z)$

$$M(z) = \frac{a.z + b}{c.z + d} \text{ avec } ad - bc = 1$$

THÉORIE DES GROUPES FUCHSIENS
 PAR H. POINCARÉ
 À PARIS.

Dans une série de mémoires présentés à l'Académie des Sciences j'ai défini certaines fonctions nouvelles que j'ai appelées fuchsienues, kleinéennes, thétafuchsienues et zétafuchsienues. De même que les fonctions elliptiques et abéliennes permettent d'intégrer les différentielles algébriques, de même les nouvelles transcendentes permettent d'intégrer les équations différentielles linéaires à coefficients algébriques. J'ai résumé succinctement les résultats obtenus dans une note insérée aux *Mathematische Annalen*. Ayant l'intention de les exposer en détail, je commencerai, dans le présent travail, par étudier les propriétés des groupes fuchsienus, me réservant de revenir plus tard sur leurs conséquences au point de vue de la théorie des fonctions.

§ 1. Substitutions réelles.

Soit z une variable imaginaire définie par la position d'un point dans un plan; t une fonction imaginaire de cette variable définie par la relation:

$$(1) \quad t = \frac{az + b}{cz + d}$$

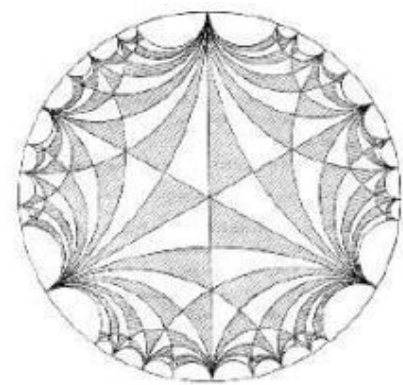
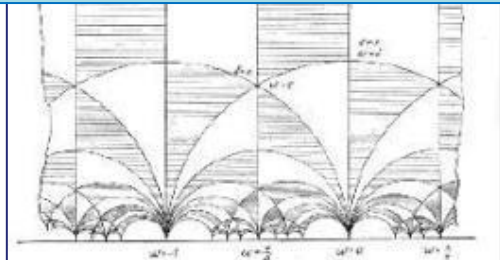
Je supposerai, ce qui ne restreint pas la généralité, que l'on a:

$$ad - bc = 1.$$

Si le point z décrit deux arcs de courbe se coupant sous un certain angle α , le point t décrira de son côté deux arcs de courbe se coupant

Acta Mathematica, 1.

- « Acta Mathematica » fondé par Gösta Mittag-Leffler en 1882
- Henri Poincaré publie dans le n°1, Vol.1 page 1: **Henri Poincaré (1882) "Théorie des Groupes Fuchsienus", Acta Mathematica**



Moyenner les matrices de covariance structurées

- La distance entre 2 matrices de covariance est donné par la distance dans l'espace produit $R^+ \times D_n$ (via la distance de Poincaré dans le disque):

$$d^2 \left[\left(\bar{P}_{0,1}, \{\bar{\mu}_{i,1}\}_{i=1}^{N-1} \right), \left(\bar{P}_{0,2}, \{\bar{\mu}_{i,2}\}_{i=1}^{N-1} \right) \right] = N \log^2 \left(\frac{\bar{P}_{0,2}}{\bar{P}_{0,1}} \right) + \sum_{i=1}^{N-1} (N-i) \left(\frac{1}{2} \log \left(\frac{1+\delta_i}{1-\delta_i} \right) \right)^2$$

$$\text{avec } \delta_i = \left| \frac{\bar{\mu}_{i,1} - \bar{\mu}_{i,2}}{1 - \bar{\mu}_{i,1} \bar{\mu}_{i,2}^*} \right|$$

- La moyenne est alors définie comme le barycentre géodésique de Fréchet qui minise la somme des carrés des distances géodésiques:

$$\left(P_{0,barycenter}, \{\mu_{i,barycenter}\}_{i=1}^{N-1} \right) =$$

$$P_{0,median}, \{\mu_{i,median}\}_{i=1}^{N-1} = \underset{P_{0,barycenter}, \{\mu_{i,barycenter}\}_{i=1}^{N-1}}{\text{ArgMin}} \sum_{k=1}^M d^p \left[\left(P_{0,barycenter}, \{\mu_{i,barycenter}\}_{i=1}^{N-1} \right), \left(\bar{P}_{0,k}, \{\bar{\mu}_{i,k}\}_{i=1}^{N-1} \right) \right]$$



- Le calcul du barycentre de Fréchet est calculé par le flot de Hermann Karcher

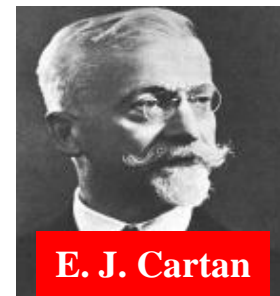
Unicité du centre de masse: de Elie Cartan à Hermann Karcher

Centre de Masse de Cartan

- Elie Cartan a prouvé que la fonctionnelle suivante:

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

est strictement convexe et possède un unique minimum (centre de masse de A pour la distribution da) pour les variétés à courbure négative.



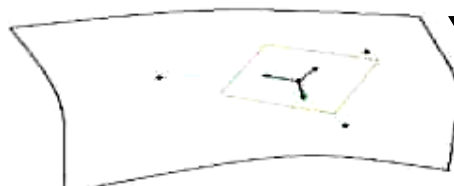
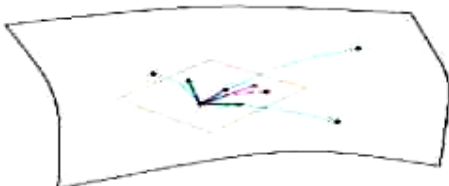
Flot de Karcher

- Hermann Karcher a prouvé la convergence du flot suivant vers le centre de masse:

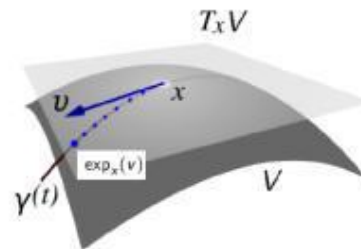
$$m_{n+1} = \gamma_n(t_n) = \exp_{m_n}(-t_n \cdot \nabla f(m_n)) \text{ avec } \dot{\gamma}_n(0) = -\nabla f(m_n)$$



H. Karcher



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



Flot de Karcher pour calculer la moyenne ou la médiane géodésique

Flot de Karcher

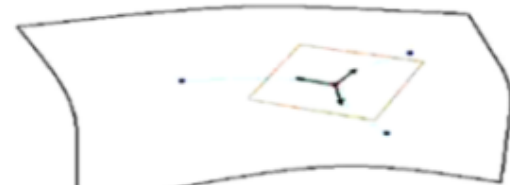
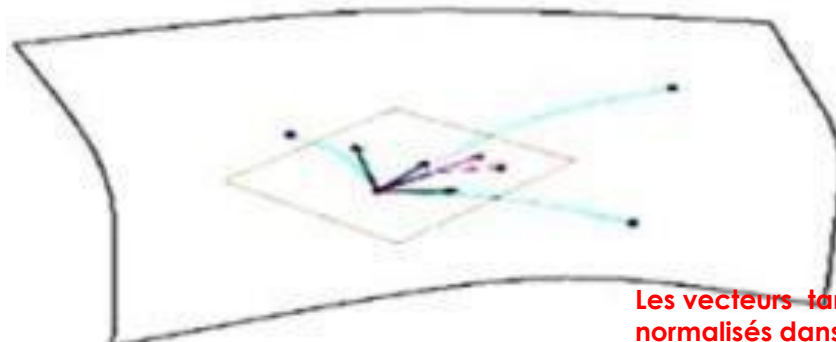
➤ Passage de l'Euclidien à la variété via la carte exponentielle pour le médian

$$m_{n+1} = t \sum_k \frac{x_k - m_n}{\|x_k - m_n\|}$$

Euclidien

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

Riemannien



Les vecteurs tangents aux géodésiques sont normalisés dans le cas du médian

Riemannian Center of Mass and Mollifier Smoothing*

H. KARCHER
Bonn University



Hermann KARCHER

- Partir d'un point arbitraire sur la variété
- Calculer les géodésiques de ce point aux N points et les vecteurs tangents (normalisés)
- Sommer les vecteurs tangents (normalisés)
- Déplacer le point sur la variété dans la direction de ce vecteur somme (via la carte exponentielle)

Riemannian Center of Mass and so called karcher mean

JULY 2014

Hermann Karcher, Bonn

Abstract. The *Riemannian center of mass* was constructed in [GrKa] (1973). In [GKR1, GKR2, Gr, Ka, BuKa] (1974-1981) it was successfully applied with more refined estimates. Probably in 1990 someone renamed it without justification into *karcher mean* and references to the older papers were omitted by those using the new name. As a consequence newcomers started to reprove results from the above papers. – Here I explain the older history.

Wikipedia talks about a generalization of centroids to metric spaces **if** the minimizer of the function $f(x) := \sum m_i \cdot d(x, p_i)^2$ **exists and is unique** and continues:

“It is named after Maurice Fréchet. Karcher means are a closely related construction named after Hermann Karcher.”

Who believes that Fréchet would have been proud of this naming?



Centre de masse : Moyenne arithmétique et médian sur R^n

Dans R^n , le centre de masse est défini pour M points $\{x_i\}_{i=1, \dots, M}$

- Moyenne arithmétique: $x_{center} = \frac{1}{M} \sum_{i=1}^M x_i$
- Ce point minimise la fonction des distances: $x_{center} = \arg \text{Min}_x \sum_{i=1}^M d^2(x, x_i)$
- Le médian (Point Fermat-Weber) minimise : $x_{median} = \arg \text{Min}_x \sum_{i=1}^M d(x, x_i)$

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

$m_{mean} = \text{Min}_m E[|x - m|^2]$

$\sum_{i=1}^n x_{moyen} x_i = 0$

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

$m_{median} = \text{Min}_m E[|x - m|]$

$\sum_{i=1}^n \frac{x_{median} x_i}{\|x_{median} x_i\|} = 0$

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



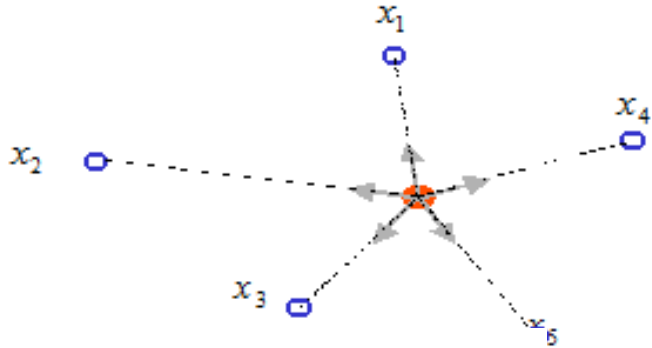
Laplace et la valeur probable ou milieu des probabilité (1774)

hed, translated, in any way, in whole or in part. © Thales - 2015 All rights reserved.

Médian (terme introduit par Cournot)

Médian géométrique (barycentre de Fréchet)

Ne nécessite pas de relation d'ordre totale

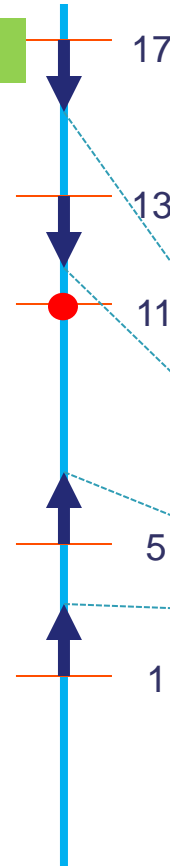


$$\sum_{i=1}^n \frac{x_{median} x_i}{\|x_{median} x_i\|} = 0$$

Médian classique

Nécessite une relation d'ordre totale

Changer 17 en 1000000, ne change pas la valeur médiane



$$1 < 5 < \underline{11} < 13 < 17$$

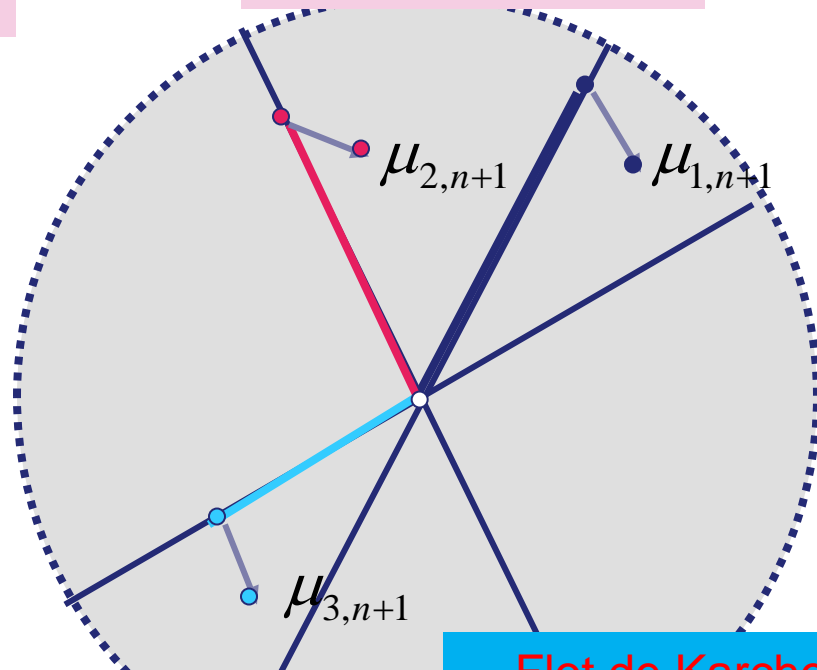
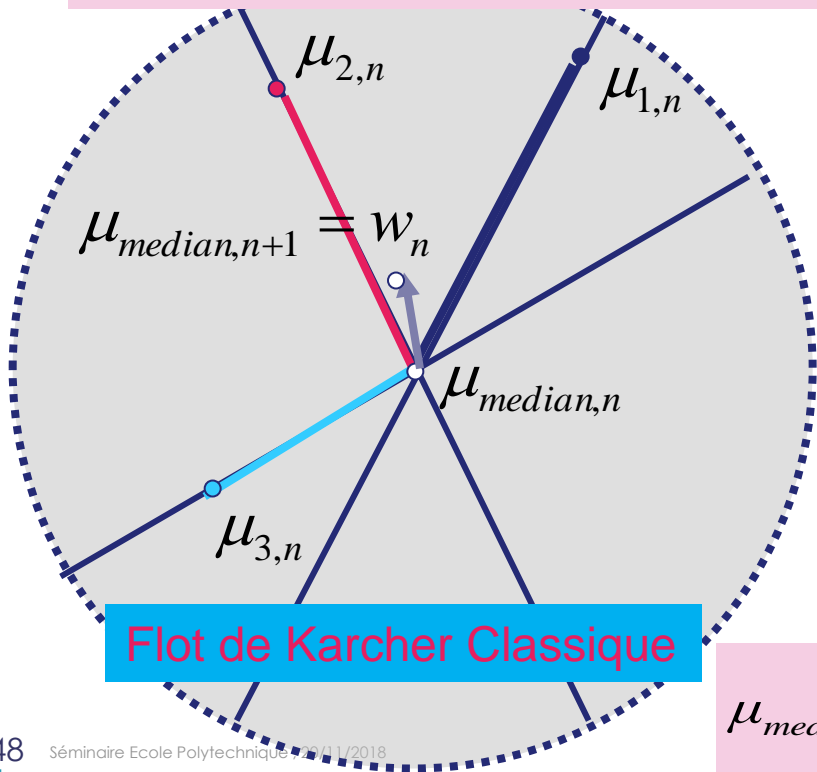
Valeur médiane (on ordonne les valeurs et on choisit la valeur pour avoir autant de valeurs au dessus et en dessous)

$$\sum_{i=1}^n \frac{x_{median} x_i}{\|x_{median} x_i\|} = 0$$

Flot de Karcher pour le barycentre de Fréchet dans le disque unité de Poincaré: calcul des coefficients de réflexions/Verblunsky

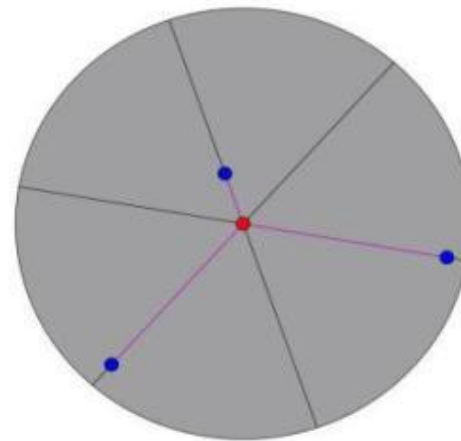
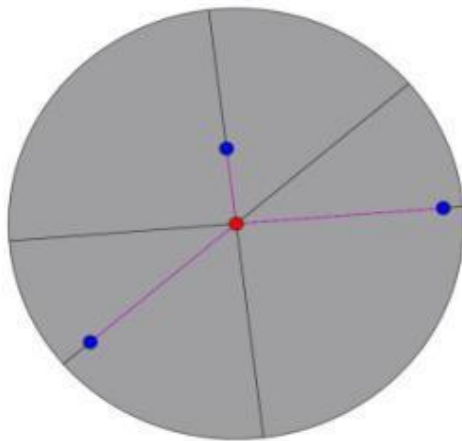
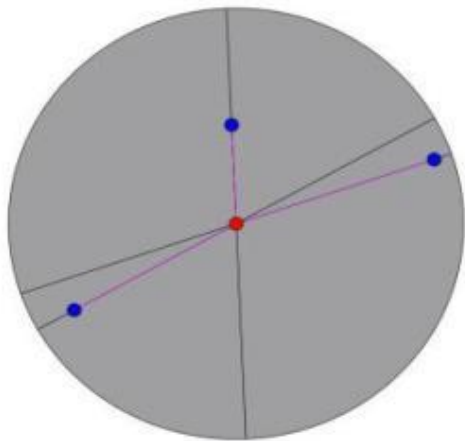
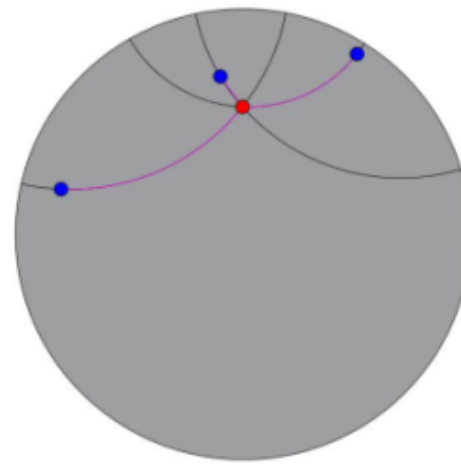
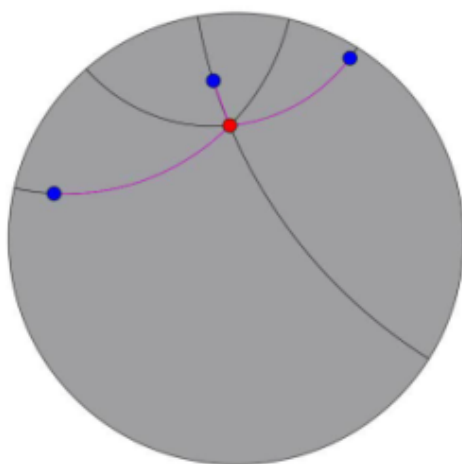
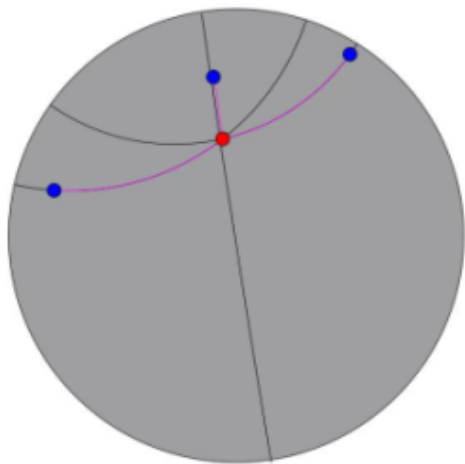
$$w_n = \gamma_n \sum_{\substack{k=1 \\ k \neq l}}^m \frac{\mu_{k,n}}{|\mu_{k,n}|} \quad \text{avec} \quad \left\{ \mathcal{V} \mid |\mu_{l,n}| < \varepsilon \right\}$$

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

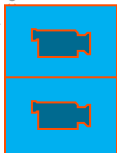


$$\mu_{median,n+1} = \frac{\mu_{median,n} + w_n}{1 + \mu_{median,n} w_n^*}$$

Flot de Karcher pour le barycentre de Fréchet dans le disque unité de Poincaré: calcul des coefficients de réflexions/Verblunsky

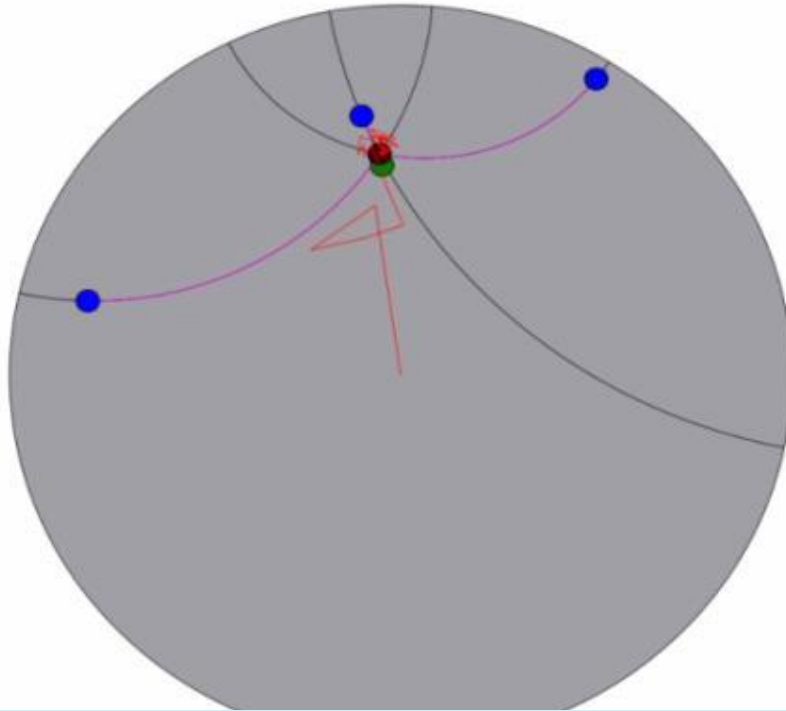


This document may not be reproduced, modified, adapted, published, in any way, in whole or in part or disclosed to a third party without the consent of Thales - © Thales 2015 All rights reserved.

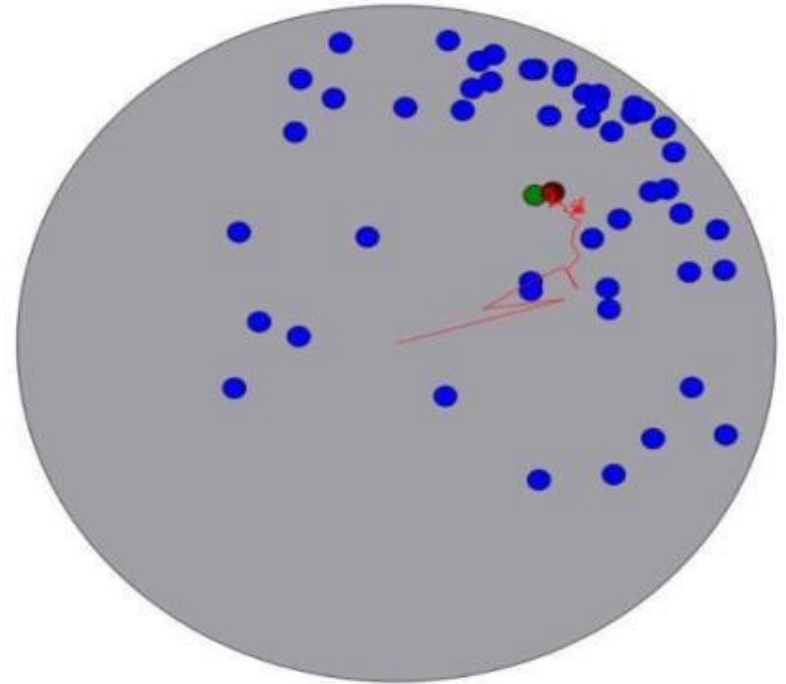


Flot stochastique d'Arnaudon

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the permission of the author. © 2015 All rights reserved.



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



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

Cas du disque de Poincaré

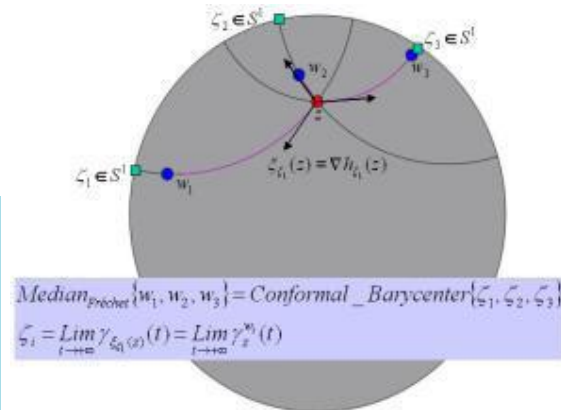
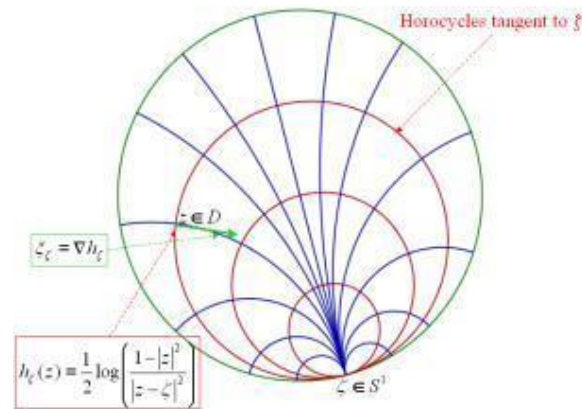
- $\xi_\mu(z)$ peut être écrit en fonction de $\xi_\zeta(z)$ qui est le vecteur tangent de la géodésique en $z \in D$ pointant vers $\zeta \in S^1$:

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

- En géométrie de Poincaré du disque unité, le champ de vecteur ξ_ζ est le gradient d'une fonction h_ζ dont les lignes de niveau sont les horocycles tangents à $\zeta \in S^1$ en S^1 :

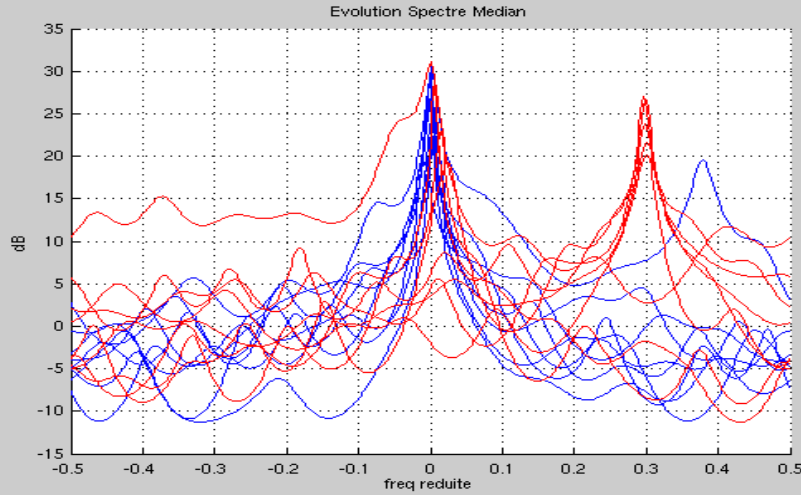
$$\xi_\mu = \nabla h_\mu \quad \text{avec} \quad h_\mu : z \mapsto \int_{S^1} h_\zeta(z) d\mu(\zeta)$$

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

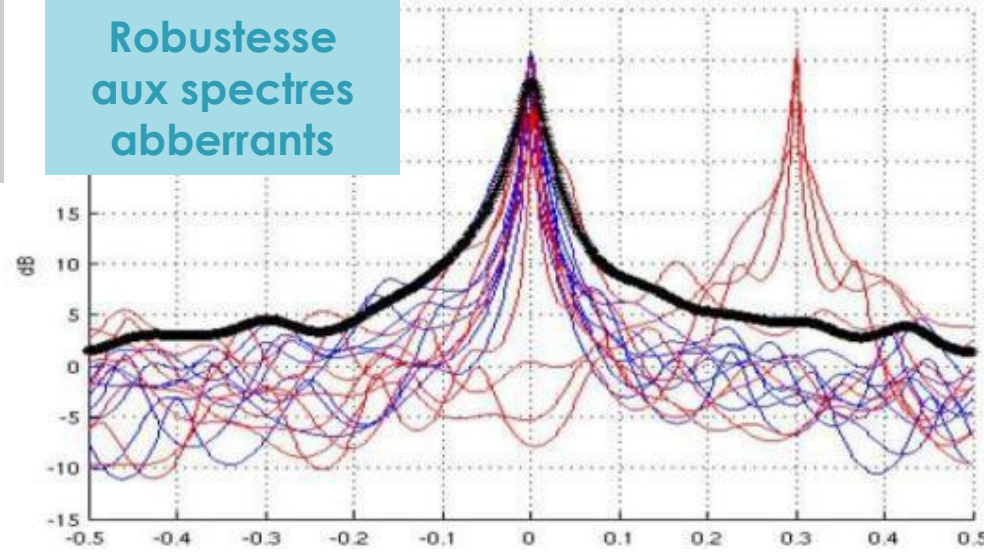


Spectre Doppler "median" par le calcul du barycentre de Fréchet sur les coefficients de réflexion/Verblunsky dans le disque unité

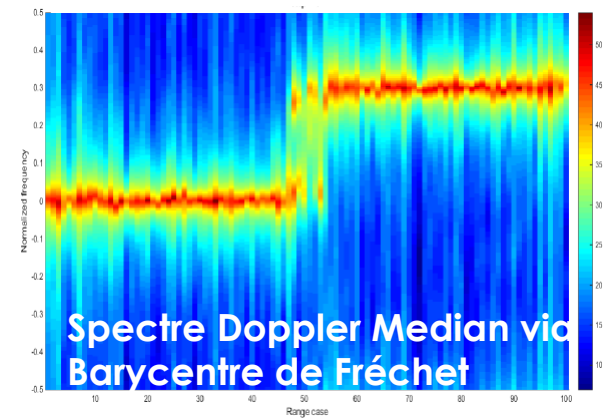
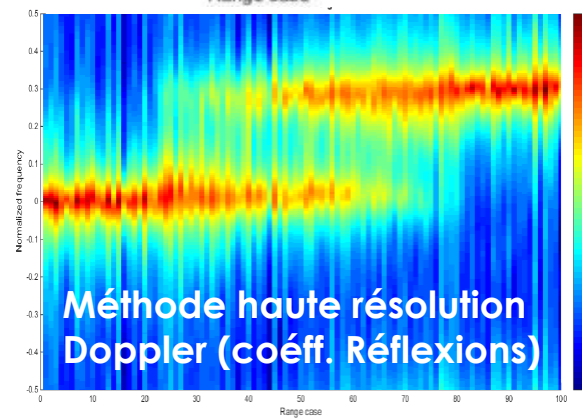
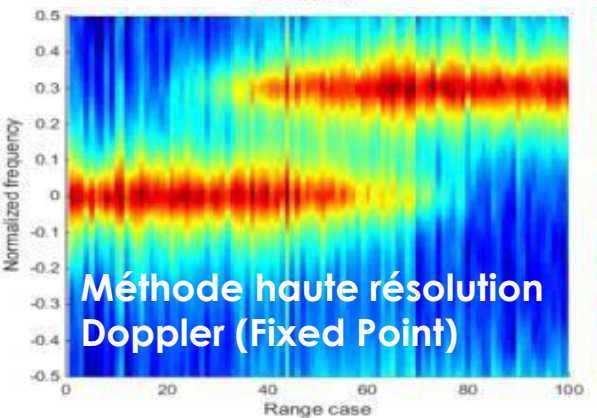
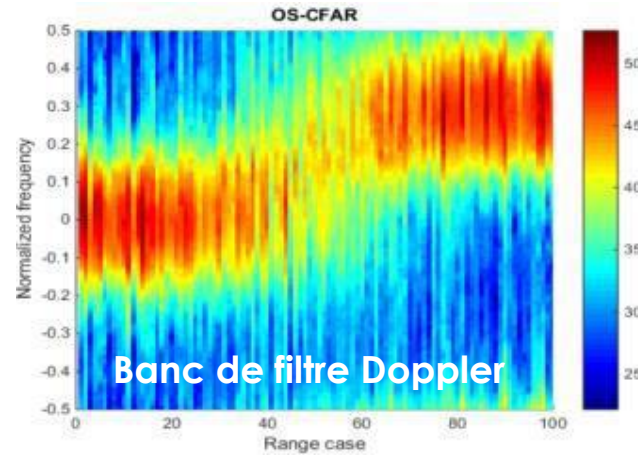
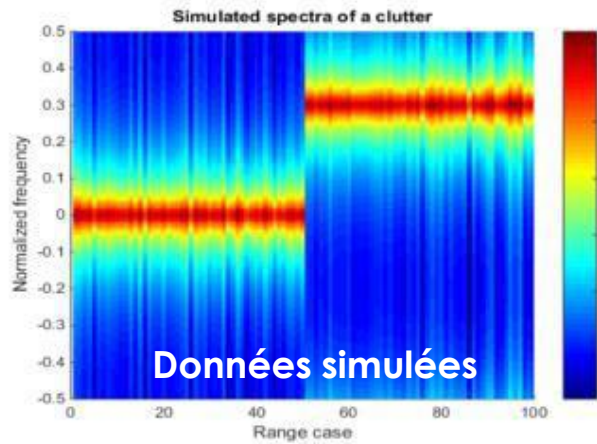
Barycentre géodésique médian



Robustesse aux spectres aberrants

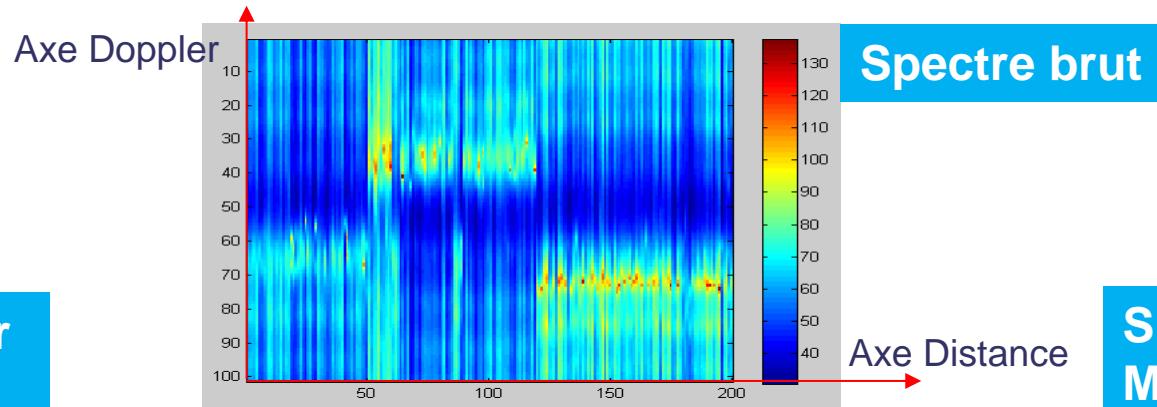


Estimation du spectre Doppler « median » : préserve les discontinuités

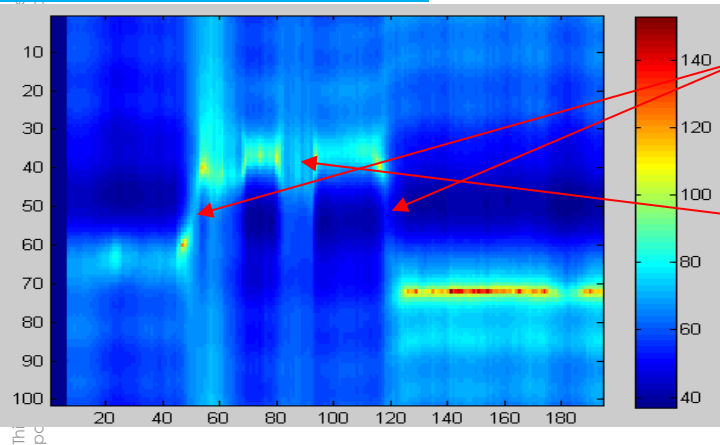


Comparaison spectre Doppler moyen et médian

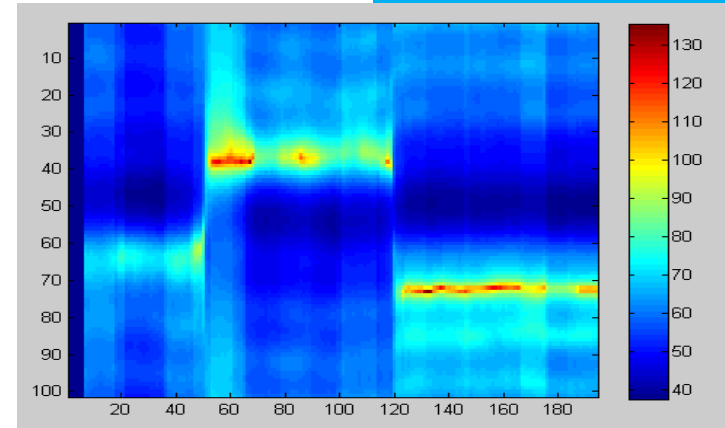
notated, in any way, in whole or in
© Thales 2015 All rights reserved.



Spectre Doppler
Moyen



Spectre Doppler
Median



Equation de la chaleur dans l'espace métrique (matrices HDP)

Equation de diffusion de la chaleur de Fourier (cas scalaire 1D)

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

➤ Approximation par un Laplacien discret : $\hat{u}_n = (u_{n+1} + u_{n-1})/2$



J. Fourier

Equation discrétisée de diffusion de la chaleur de Fourier (cas scalaire 1D)

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

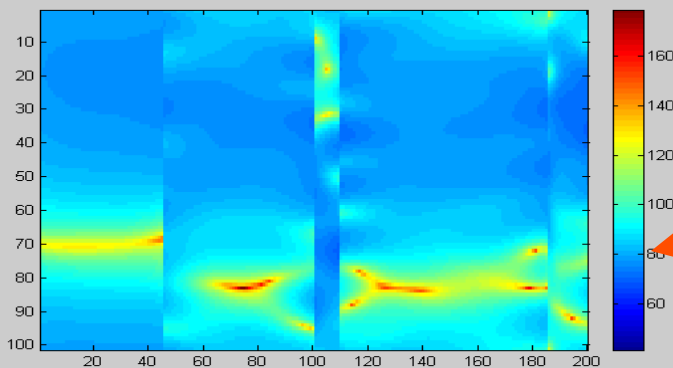
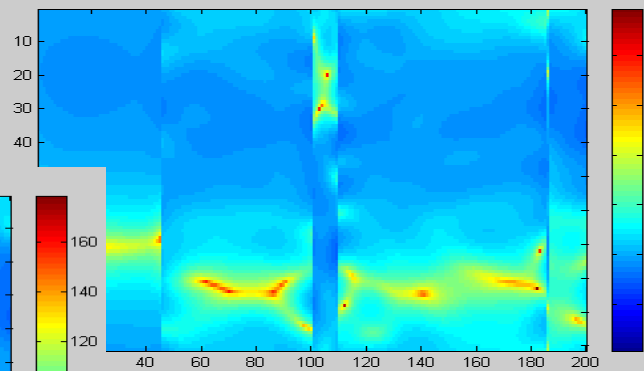
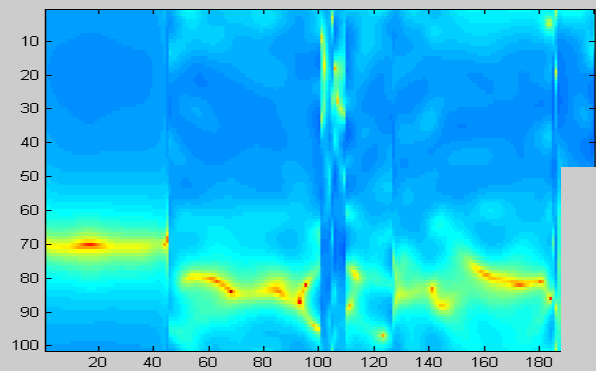
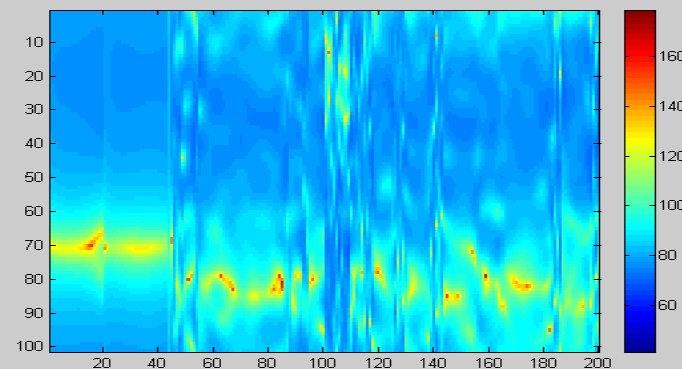
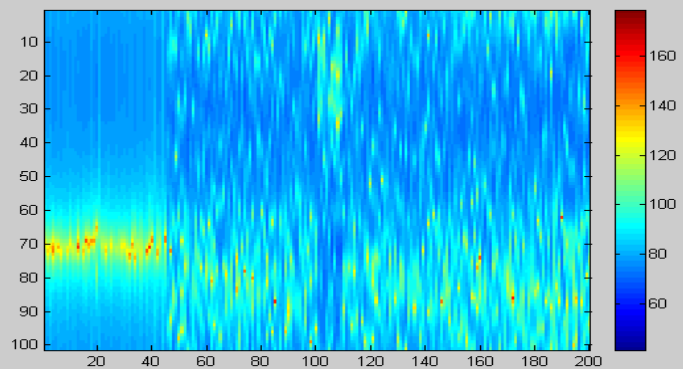
Par analogie, on peut étendre cette équation de diffusion en 1D (axe distance) mais dans le cas de l'espace métrique associé aux matrices Hermitiennes Définies Positives:

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

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

Diffusion anisotrope de Spectre Doppler

ned, translated, in any way, in whole or in
Thales - © Thales 2015 All rights reserved.

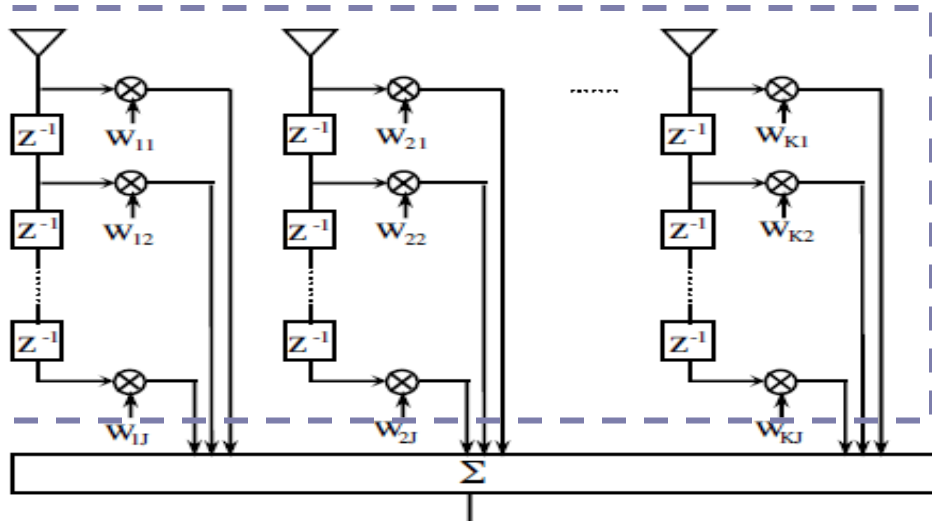


This document
part or disclo:



Extension à la mesure spatio-temporelle de l'onde électromagnétique: structure de Matrice Toeplitz-Bloc-Toeplitz

Nous considérons un vecteur de mesures spatiales et temporelles



Mesures digitale spatio-temporelle (Matrice de covariance spatio-temporelle)

$$z_{i,j} : \begin{cases} i : \text{indice temporel} \\ j : \text{indice spatial} \end{cases}$$

$$R = E[ZZ^+]$$

$$Z = \begin{bmatrix} z_{1,1} \\ \vdots \\ z_{N,1} \\ \vdots \\ z_{1,M} \\ \vdots \\ z_{N,M} \end{bmatrix}$$

$$R_{p,n+1} = \begin{bmatrix} R_0 & R_1 & \cdots & R_n \\ R_1^+ & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1 \\ R_n^+ & \cdots & R_1^+ & R_0 \end{bmatrix} = \begin{bmatrix} R_{p,n} & \tilde{R}_n \\ \tilde{R}_n^+ & R_0 \end{bmatrix}$$

$$\tilde{R}_n = V \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix}^* \text{ avec } V = \begin{bmatrix} 0 & \cdots & 0 & J_p \\ \vdots & \ddots & \ddots & 0 \\ 0 & J_p & \ddots & \vdots \\ J_p & 0 & \cdots & 0 \end{bmatrix}$$

Coding Micro-Doppler Data in Siegel Poly-Disk

$$\{z_1, \dots, z_p, \dots, z_M\} \xRightarrow{\text{split}} \left\{ \{z_1, \dots, z_p\}, \{z_{p+1}, \dots, z_{2p}\}, \dots, \{z_{M-p+1}, \dots, z_M\} \right\}$$

d. in any wa
notes 2015 A

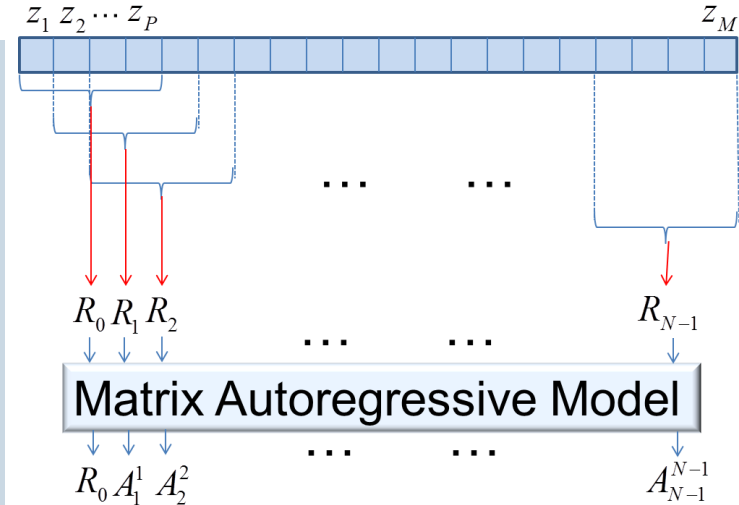
Time series $\{R_0, \dots, R_{N-1}\}$ as TBTHPD matrix

$$R_{N,p} = \begin{bmatrix} R_0 & R_1 & \dots & R_{N-1} \\ R_1^+ & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1 \\ R_{N-1}^+ & \dots & R_1^+ & R_0 \end{bmatrix}$$

document
or disclose

Matrix Verblunsky/Trench Theorem

$$\{R_0, A_1^1, \dots, A_{N-1}^{N-1}\} \text{ with } A_k^k \in SD = \{Z / ZZ^+ < I\}$$



Considering Toeplitz-Block-Toeplitz Matrices :

$$R_{p,n+1} = \begin{bmatrix} R_0 & R_1 & \cdots & R_n \\ R_1^+ & R_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1 \\ R_n^+ & \cdots & R_1^+ & R_0 \end{bmatrix} = \begin{bmatrix} R_{p,n} & \tilde{R}_n \\ \tilde{R}_n^+ & R_0 \end{bmatrix}$$

$$\tilde{R}_n = V \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix}^* \quad \text{with} \quad V = \begin{bmatrix} 0 & \cdots & 0 & J_p \\ \vdots & \ddots & \ddots & 0 \\ 0 & J_p & \ddots & \vdots \\ J_p & 0 & \cdots & 0 \end{bmatrix}$$

Efficient Inversion of Toeplitz-Block Toeplitz Matrix

MATI WAX AND THOMAS KAILATH, FELLOW, IEEE

Extension du théorème de Trench/Verblunsky aux cas des matrices Toeplitz-Bloc-Toeplitz Hermitiennes Définies Positives

Les matrices Toeplitz-Bloc-Toeplitz Hermitiennes Définies Positives peuvent être paramétrées par des coefficients de Verblunsky matriciels:

$$R_{p,n+1}^{-1} = \begin{bmatrix} \alpha_n & \alpha_n \cdot \widehat{A}_n^+ \\ \alpha_n \cdot \widehat{A}_n & R_{p,n}^{-1} + \alpha_n \cdot \widehat{A}_n \cdot \widehat{A}_n^+ \end{bmatrix} \quad \text{with} \quad \alpha_n^{-1} = [1 - A_n^n A_n^{n+}] \alpha_{n-1}^{-1}, \quad \alpha_0^{-1} = R_0$$

$$R_{p,n+1} = \begin{bmatrix} \alpha_n^{-1} + \widehat{A}_n^+ \cdot R_{p,n} \cdot \widehat{A}_n & -\widehat{A}_n^+ \cdot R_{p,n} \\ -R_{p,n} \cdot \widehat{A}_n & R_{p,n} \end{bmatrix} \quad \text{and} \quad \widehat{A}_n = \begin{bmatrix} A_1^1 \\ \vdots \\ A_n^n \end{bmatrix} = \begin{bmatrix} \widehat{A}_{n-1} \\ 0_p \end{bmatrix} + A_n^n \cdot \begin{bmatrix} J_p A_{n-1}^{n-1*} J_p \\ \vdots \\ J_p A_1^{n-1*} J_p \\ I_p \end{bmatrix}$$

Extension du théorème de Trench/Verblunsky au cas matriciel: Existence d'un difféomorphisme φ :

$$\varphi : TBTHPD_{n \times n} \rightarrow THPD_n \times SD^{n-1}$$

$$R \mapsto (R_0, A_1^1, \dots, A_{n-1}^{n-1})$$

$$\text{with } SD = \{Z \in Herm(n) / ZZ^+ < I_n\}$$

Métrie de la géométrie de l'Information: Hessien du potentiel

- L'entropie définie encore un potentiel (Kähleriën) dont le hessien fournit une métrique Riemannienne pour les paramètres matriciels de Verblunsky:

$$\begin{aligned}\tilde{\Phi}(R_{p,n}) &= -\log(\det R_{p,n}) + cste = -Tr(\log R_{p,n})\mu + cste \\ \Rightarrow g_{ij} &= Hess[\phi(R_{p,n})]\end{aligned}$$

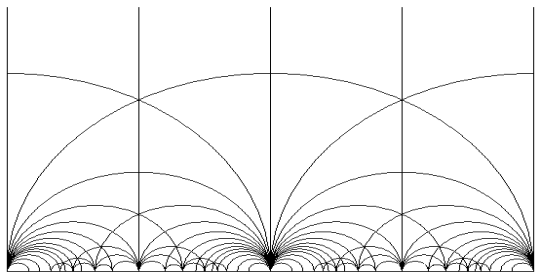
- La structure « Toeplitz-Bloc-Toeplitz » permet d'exprimer l'entropie uniquement à partir des paramètres de Verblunsky traduisant l'information spatiale et d'une matrice R_0 incorporant l'information Doppler étudiée précédemment:

$$\tilde{\Phi}(R_{p,n}) = \sum_{k=1}^{n-1} (n-k) \cdot \log \det [I_n - A_k^k A_k^{k+}] + n \cdot \log [\pi \cdot e \cdot \det R_0]$$

- Le hessien de l'Entropie donne la métrique:

$$ds^2 = n \cdot Tr \left[(R_0^{-1} dR_0)^2 \right] + \sum_{k=1}^{n-1} (n-k) Tr \left[(I_n - A_k^k A_k^{k+})^{-1} dA_k^k (I_n - A_k^{k+} A_k^k)^{-1} dA_k^{k+} \right]$$

Extension des domaines symétriques bornés homogènes: Domaine de Siegel: Demi-espace supérieur et disque de Siegel



Demi-Plan Supérieur

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

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

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

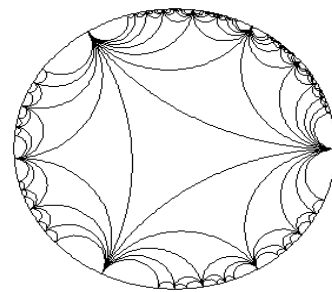
Demi-espace supérieur

$$ds^2 = \text{Tr}(Y^{-1} dZY^{-1} d\bar{Z})$$

with $Z = X + iY$

3. $X \in \text{Herm}(n, \mathbb{C})$ and $Y \in \text{HPD}(n, \mathbb{C})$

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



Disque Unité de Poincaré

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

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

Disque Unité de Siegel

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

Les paramètres de Verblunsky matriciel sont élément du disque de Siegel : $\mathbb{Z}\mathbb{Z}^+ \langle I \rangle$

Les automorphismes du disque de Siegel SD_n sont donnés par:

$$\Sigma = \Phi_{Z_0}(Z) = (I - Z_0 Z_0^+)^{-1/2} (Z - Z_0) (I - Z_0^+ Z)^{-1} (I - Z_0^+ Z_0)^{1/2}$$

L'ensembles des automorphismes sont donnés par:

$$\forall \Psi \in \text{Aut}(SD_n), \exists U \in U(n, \mathbb{C}) / \Psi(Z) = U \Phi_{Z_0}(Z) U^t$$

La distance est:

$$\forall Z, W \in SD_n, d(Z, W) = \frac{1}{2} \log \left(\frac{1 + \|\Phi_Z(W)\|}{1 - \|\Phi_Z(W)\|} \right)$$

L'automorphisme inverse est donné par:

$$G = (I - Z_0 Z_0^+)^{1/2} \Sigma (I - Z_0^+ Z_0)^{-1/2} = (Z - Z_0) (I - Z_0^+ Z)^{-1}$$
$$\Rightarrow \begin{cases} Z = \Phi_{Z_0}^{-1}(\Sigma) = (G Z_0^+ + I)^{-1} (G + Z_0^+) \\ \text{with } G = (I - Z_0 Z_0^+)^{1/2} \Sigma (I - Z_0^+ Z_0)^{-1/2} \end{cases}$$

Groupe Symplectique (Carl-Ludwig Siegel) : Demi-espace SH_n

Métrie de Siegel sur le demi-plan supérieur :

➤ Demi-Plan supérieur : $SH_n = \{Z = X + iY \in \text{Sym}(n, \mathbb{C}) / \text{Im}(Z) = Y > 0\}$

➤ Les isométries de SH_n sont données par le groupe quotient :

$PSp(n, \mathbb{R}) \equiv Sp(n, \mathbb{R}) / \{\pm I_{2n}\}$ avec $Sp(n, \mathbb{F})$ le groupe symplectique :

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

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

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

➤ Seule métrique invariante par $M(Z)$:

$$ds_{\text{Siegel}}^2 = \text{Tr}(Y^{-1}(dZ)Y^{-1}(d\bar{Z})) \quad Z = X + iY$$

Distance dans le demi-plan supérieur de Siegel

Distance dans le demi-plan supérieur de Siegel:

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

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

$$\text{avec} \quad \det(R(Z_1, Z_2) - \lambda I) = 0$$

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

Cas particulier (axe imaginaire: matrices symétriques définies positives)

$$Z = iR \quad \text{avec} \quad R > 0$$

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

$$\det(R_2 - \lambda R_1) = 0$$

Extension des domaines symétriques bornés homogènes: Domaine de Siegel: Demi-espace supérieur et disque de Siegel

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales. © Thales 2015 All rights reserved.



L.K. Hua



C.L. Siegel



F. Berezin

Siegel Upper Half Space

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

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

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

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

$$Z_k = X_k + i.Y_k$$

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

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

Carl Ludwig Siegel

Carl Ludwig Siegel
avec George Polya

« At one point Siegel thought that too many unnecessary things were being published, so he decided not to publish anything at all »

George Polya

The Polya Picture Album, Encounters of a Mathematician, Birkäuser

Décomposition (polaire) de Mostow ou Fibration de Berger pour appliquer le même schéma que dans le disque de Poincaré

➤ Théorème de Mostow :

- Toute matrice M de $GL(n, \mathbb{C})$ peut être décomposé :

où
$$M = Ue^{iA}e^S$$

- U est unitaire
- A est réelle antisymétrique
- S est réelle symétrique

Peut être déduit de



G.D. Mostow



M. Berger

- ## ➤ Lemme : Soit A et B , 2 matrices hermitiennes définies positives, il existe une unique matrice hermitienne définie positive X telle que :

$$XAX = B$$

- ## ➤ Corollaire : Si M est hermitienne définie positive, il existe une unique matrice symétrique réelle S telle que : $M^* = e^S M^{-1} e^S$



Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa



A B S T R A C T

It is shown that every nonsingular square complex matrix Z can be factored as $Z = e^L e^{iT} e^{iK} e^S$, where S and T are real symmetric matrices, and K and L are real skew-symmetric matrices. New proofs are given for some related decomposition theorems for unitary and positive definite matrices.

The bipolar decomposition

Rajendra Bhatia

Indian Statistical Institute, 7, S. I. S. Sansanwal Marg, New Delhi 110016, India

Theorem (The Mostow Decomposition Theorem). Let Z be a nonsingular complex matrix. Then Z can be factored as

$$Z = W e^{iK} e^S$$

where W is unitary, S is real and symmetric, and K is real and skew-symmetric. Such a factorisation is unique.

Acknowledgements

I thank F. Barbaresco, Roger Horn and M.S. Raghunathan for illuminating discussions.

References

- [1] F. Barbaresco, Information geometry of covariance matrix: Cartan–Siegel homogeneous bounded domains, Mostow/Berger fibration and Fréchet median, in: F. Nielsen, R. Bhatia (Eds.), Matrix Information Geometry, Springer, 2013, pp. 199–255.
- [2] R. Bhatia, Positive Definite Matrices, Princeton University Press, 2007.
- [3] R. Bhatia, Matrix factorizations and their perturbations, Linear Algebra Appl. 197/198 (1994) 245–276.
- [4] R. Bhatia, J. Holbrook, Noncommutative geometric means, Math. Intelligencer 28 (2006) 32–39.

Généralisation du codage d'Henri Poincaré

$$(R_0, A_1^1, \dots, A_{n-1}^{n-1}) \in THPD_m \times SD^{n-1}$$

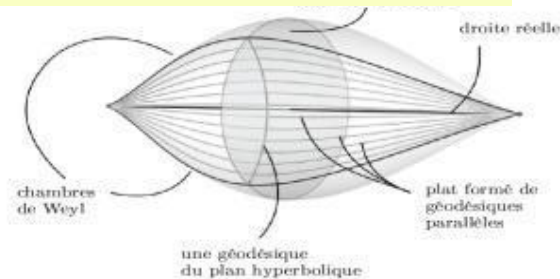
$$SD = \{Z / ZZ^+ < I_m\}$$

$$R_0 \rightarrow (\log(P_0), \mu_1, \dots, \mu_{m-1}) \in R \times D^{m-1}$$

$$D = \{z / zz^* < 1\}$$

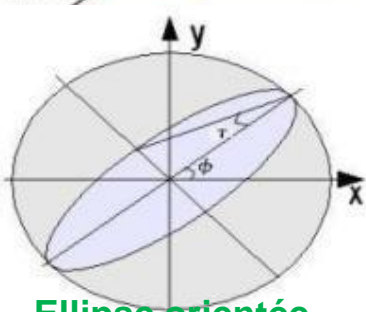
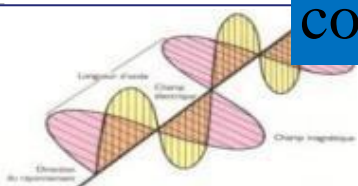
$$\text{codage spatio - Doppler} \in R \times D^{m-1} \times SD^{n-1}$$

Information Spatio-temporelle:
R x Polydisque Poincaré x Polydisque Siegel



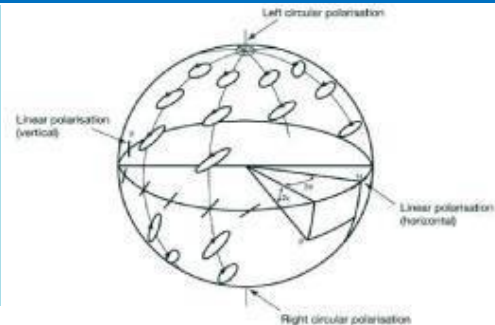
$$\text{codage mesure onde électromagnétique} \in R \times S^1 \times D^{m-1} \times SD^{n-1}$$

This document may not be reproduced, modified, adapted, or distributed in any form without the prior written consent of the author.



Ellipse orientée

$$\vec{S} = \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} s_0 \\ s_0 \cos 2\tau \cos 2\phi \\ s_0 \cos 2\tau \sin 2\phi \\ s_0 \sin 2\tau \end{bmatrix} \text{ avec } \begin{cases} \phi = \frac{1}{2} \arctan\left(\frac{s_2}{s_1}\right) \\ \tau = \frac{1}{2} \arctan\left(\frac{s_3}{\sqrt{s_1^2 + s_2^2}}\right) \end{cases}$$



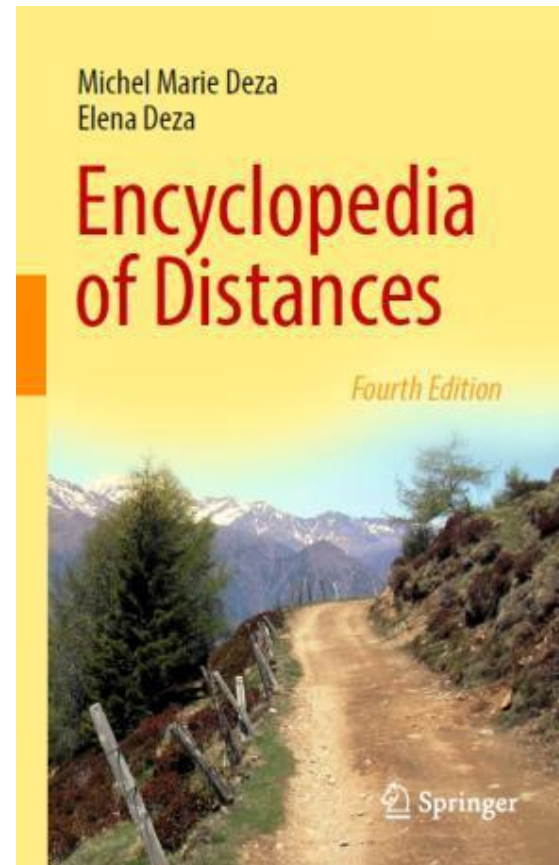
$$\text{codage polarimétrique} \in R \times S^1$$

Information polarimétrique:
 $R_+ \times S^1$

- **Barbaresco metrics**

Let $z(k)$ be a complex temporal (discrete time) *stationary* signal, i.e., its mean value is constant and its *covariance function* $\mathbb{E}[z(k_1)z^*(k_2)]$ is only a function of $k_1 - k_2$. Such signal can be represented by its covariance $n \times n$ matrix $R = ((r_{ij}))$, where $r_{ij} = \mathbb{E}[z(i), z^*(j)] = \mathbb{E}[z(n)z^*(n-i+j)]$. It is a positive-definite *Toeplitz* (i.e. diagonal-constant) Hermitian matrix. In radar applications, such matrices represent the Doppler spectra of the signal. Matrices R admit a parametrization (complex ARM, i.e., m -th order autoregressive model) by *partial autocorrelation coefficients* defined recursively as the complex correlation between the forward and backward prediction errors of the $(m - 1)$ -th order complex ARM.

Barbaresco ([Barb12]) defined, via this parametrization, a **Bergman metric** (Chap. 7) on the bounded domain $\mathbb{R} + xD_n \subset \mathbb{C}^n$ of above matrices R ; here D is a *Poincaré disk*. He also defined a related **Kähler metric** on $M \times S_n$, where M is the set of positive-definite Hermitian matrices and SD_n is the *Siegel disk* (cf. **Siegel distance**). Such matrices represent spatiotemporal stationary signals, i.e., in radar applications, the Doppler spectra and spatial directions of the signal.



Numerical scheme for Fréchet Mean Computation with this distance



KU Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

THE KÄHLER MEAN OF BLOCK-TOEPLITZ MATRICES WITH TOEPLITZ STRUCTURED BLOCKS

B. JEURIS[†] AND R. VANDEBRIL[‡]

Abstract. When computing an average of positive definite (PD) matrices, the preservation of additional matrix structure is desirable for interpretations in applications. An interesting and widely present structure is that of PD Toeplitz matrices, which we endow with a geometry originating in signal processing theory. As an averaging operation, we consider the barycenter, or minimizer of the sum of squared intrinsic distances. The resulting barycenter, the Kähler mean, is discussed along with its origin. Also, a generalization of the mean towards PD (Toeplitz-Block) Block-Toeplitz matrices is discussed. For PD Toeplitz-Block Block-Toeplitz matrices, we derive the generalized barycenter, or generalized Kähler mean, and a greedy approximation. This approximation is shown to be close to the generalized mean with a significantly lower computational cost.

Key words. matrix mean, Toeplitz-Block Block-Toeplitz matrices, differential geometry, optimization on manifolds, Riemannian manifolds

Jeuris, B., Vandebril, R. (2016). The Kähler mean of block-Toeplitz matrices with Toeplitz structured blocks. *SIAM Journal on Matrix Analysis and Applications*

<https://lirias.kuleuven.be/handle/123456789/533797>

$$d_K(\Omega, \Psi) = \frac{1}{2} \log \left(\frac{1 + \|\phi_\Omega(\Psi)\|_2}{1 - \|\phi_\Omega(\Psi)\|_2} \right),$$

$$\left[\phi_\Omega(\Psi) = (I - \Omega\Omega^H)^{-\frac{1}{2}} (\Psi - \Omega) (I - \Omega^H\Psi)^{-1} (I - \Omega^H\Omega)^{\frac{1}{2}}, \right.$$

$$\mathcal{B}_S(A_1, \dots, A_k) = \arg \min_{X \in \mathcal{S}} \frac{1}{2} \sum_{i=1}^k d_S^2(X, A_i).$$

$$\begin{aligned} \mathcal{T}_n^+ &\rightarrow \mathbb{R}^{++} \times \mathbb{D}^{n-1} \\ \mathcal{R}_n &\mapsto (p_0, \mu_1, \dots, \mu_{n-1}), \end{aligned}$$

$$\Phi(\mathcal{R}_n) = \log(\det \mathcal{R}_n^{-1}) - \log(\pi e),$$

$$\Phi(\mathcal{R}_n) = -n \log(p_0) - \sum_{\ell=1}^{n-1} (n-\ell) \log(1 - |\mu_\ell|^2) - \log(\pi e),$$

$$[H]_{i,j} = \frac{\partial^2 \Phi}{\partial \xi_i^{(n)} \partial \xi_j^{(n)}}.$$

$$\begin{aligned} ds^2 &= d\xi^{(n)H} H d\xi^{(n)} \\ &= n \frac{dp_0^2}{p_0^2} + \sum_{\ell=1}^{n-1} (n-\ell) \frac{|d\mu_\ell|^2}{(1 - |\mu_\ell|^2)^2}. \end{aligned}$$

$$\begin{aligned} d_{\mathcal{T}_n^+}^2(T_1, T_2) &= d_{\mathcal{T}_n^+}^2\left((p_{0,1}, \mu_{1,1}, \dots, \mu_{n-1,1}), (p_{0,2}, \mu_{1,2}, \dots, \mu_{n-1,2})\right) \\ &= n \log^2 \left(\frac{p_{0,2}}{p_{0,1}} \right) + \sum_{\ell=1}^{n-1} \frac{n-\ell}{4} \log^2 \left(\frac{1 + \left| \frac{\mu_{\ell,1} - \mu_{\ell,2}}{1 - \mu_{\ell,1}\mu_{\ell,2}} \right|}{1 - \left| \frac{\mu_{\ell,1} - \mu_{\ell,2}}{1 - \mu_{\ell,1}\mu_{\ell,2}} \right|} \right). \end{aligned}$$

$$\begin{aligned} \mathcal{B}_{n,N}^+ &\rightarrow \mathcal{P}_N \times \mathcal{SD}_N^{n-1} \\ \tilde{\mathcal{R}}_n &\mapsto (P_0, \Omega_1, \dots, \Omega_{n-1}), \end{aligned}$$

$$\Phi(\tilde{\mathcal{R}}_n) = -n \log(\det P_0) - \sum_{\ell=1}^{n-1} (n-\ell) \log(\det(I - \Omega_\ell \Omega_\ell^H)) - \log(\pi e),$$

$$\begin{aligned} ds^2 &= n \operatorname{trace}(P_0^{-1} dP_0 P_0^{-1} dP_0) \\ &\quad + \sum_{\ell=1}^{n-1} (n-\ell) \operatorname{trace} \left((I - \Omega_\ell \Omega_\ell^H)^{-1} d\Omega_\ell (I - \Omega_\ell^H \Omega_\ell)^{-1} d\Omega_\ell^H \right). \end{aligned}$$

Algorithm 1 Procedure for translating to the origin

Let $\Omega_1, \dots, \Omega_k$ be k matrices in \mathcal{SD}_N , $X_0 \in \mathcal{SD}_N$ an initial guess

- for $j = 0, 1, \dots$
 - Compute the translated matrices:

$$(\Omega_1^{(j)}, \dots, \Omega_k^{(j)}) = (\phi_{X_j}(\Omega_1), \dots, \phi_{X_j}(\Omega_k));$$

- Compute the gradient of the translated cost function at the origin (4.4):

$$\text{grad } f_{\mathcal{B}_{\mathcal{SD}_N}}(0; \Omega_1^{(j)}, \dots, \Omega_k^{(j)}),$$

and perform a basic descent step to obtain Ψ_{j+1} ;

- Obtain the next iteration point by returning to the original matrices:

$$X_{j+1} = \phi_{-X_j}(\Psi_{j+1});$$

- end for

Return: $\mathcal{B}_{\mathcal{SD}_N}(\Omega_1, \dots, \Omega_k)$

$$\mathcal{B}_{\mathcal{SD}_N}(\Omega_1, \dots, \Omega_k) = \arg \min_{X \in \mathcal{SD}_N} \frac{1}{2} \sum_{i=1}^k \left\| \log \left(\frac{I + C_i^{\frac{1}{2}}}{I - C_i^{\frac{1}{2}}} \right) \right\|_F^2,$$

$$C_i = I - (I - \Omega_i \Omega_i^H)^{\frac{1}{2}} (I - X \Omega_i^H)^{-1} (I - X X^H) (I - \Omega_i X^H)^{-1} (I - \Omega_i \Omega_i^H)^{\frac{1}{2}},$$

$$\text{grad } f_{\mathcal{B}_{\mathcal{SD}_N}}(0; \Omega_1, \dots, \Omega_k) = - \sum_{i=1}^k V_i \Omega_i,$$

$$\left[V_i = \mathcal{L} \left((\Omega_i \Omega_i^H)^{\frac{1}{2}}, \log \left(\frac{I + (\Omega_i \Omega_i^H)^{\frac{1}{2}}}{I - (\Omega_i \Omega_i^H)^{\frac{1}{2}}} \right) \right) \right]$$

Approche de Maurice Fréchet :
Les éléments aléatoires de
nature quelconque dans un
espace distancié de Maurice
Fréchet



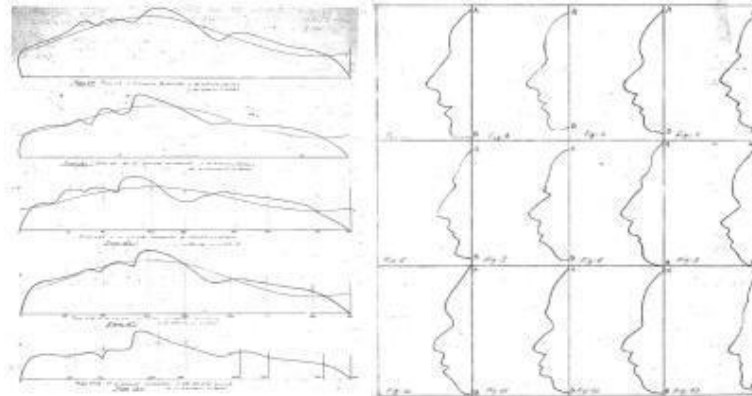
Blaise Pascal

Maurice Fréchet : Les éléments aléatoires de nature quelconque

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)

- 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 : **des transformations aléatoires de courbe en courbe**.

Documents numérisés à partir du Fond Fréchet des Archives de l'Académie des Sciences

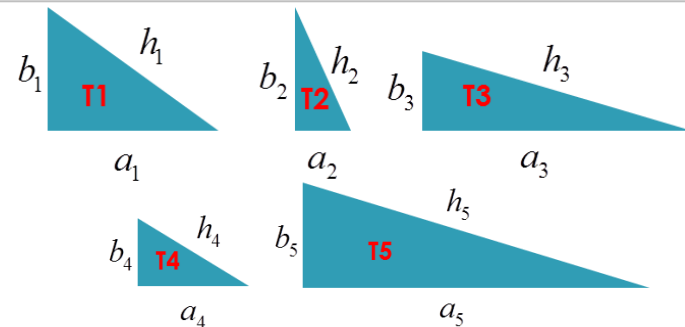


M. Fréchet collaboration avec M. Ozil: Analyse harmonique des profils humains. Les tableaux de M. Ozil ne correspondent pour des formes en fait assez compliquées qu'à 9 harmoniques et représentent seulement la moitié expressive du profil, c'est à dire une fonction non périodique.

Antoine-Augustin Cournot et le triangle rectangle moyen

Joseph Bertrand et l'homme moyen

- Lorsqu'on applique la détermination des moyennes aux diverse parties d'un système compliqué, il faut bien prendre garde que ces valeurs moyennes peuvent ne pas se convenir : en sorte que l'état du système, dans lequel tous les éléments prendraient à la fois les valeurs moyennes déterminées séparément pour chacun d'eux, serait un état impossible.
- Si, par exemple, un triangle est assujéti à rester rectangle pendant que ses côtés varient, il y aura une valeur moyenne pour chacun des trois côtés; mais ces trois moyennes, prises ensemble, ne conviendront pas à un triangle rectangle, ou ne satisferont pas à cette condition si connue, que le carré fait sur l'hypoténuse égale à la somme des carrés fait sur les deux côtés de l'angle droit.
- L'homme moyen: Par un système de moyennes tirées de la mesure de la taille, du poids, des forces, etc., sur des individus en grand nombre. L'homme moyen ainsi défini, bien loin d'être en quelque sorte le type de l'espèce, serait tout simplement un impossible



$$\left(\frac{a_1 + a_2 + a_3 + a_4 + a_5}{5}\right)^2 + \left(\frac{b_1 + b_2 + b_3 + b_4 + b_5}{5}\right)^2 \neq \left(\frac{h_1 + h_2 + h_3 + h_4 + h_5}{5}\right)^2$$



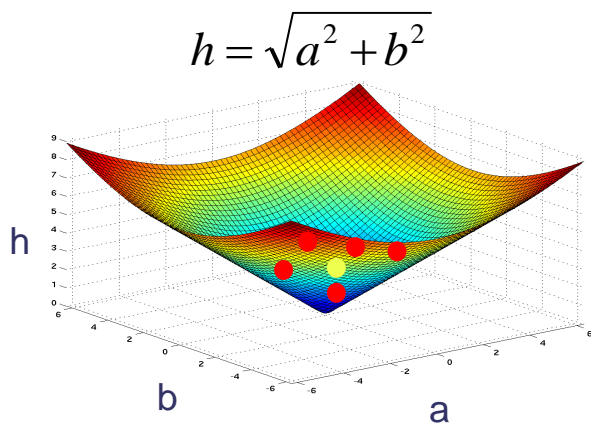
Le poids d'un individu est grosso modo proportionnel à son volume, il varie comme le cube de la taille; or la moyenne des cubes n'est évidemment pas le cube de la moyenne – **Joseph Bertrand**

Moyenne/Médiane: Barycentre de Fréchet dans l'espace métrique

Illustrons l'idée du barycentre de Fréchet pour des « triangles rectangles »

- Considérons N triangles rectangles $\{a_i, b_i, h_i\}$, on cherche le « triangle median rectangle »: La solution est le barycentre « géodésique » de Fréchet, qui minimise la somme des distances géodésiques à tous les triangles rectangles

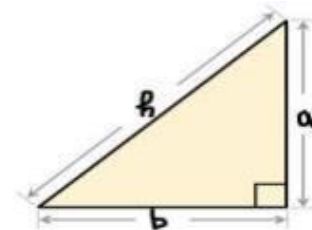
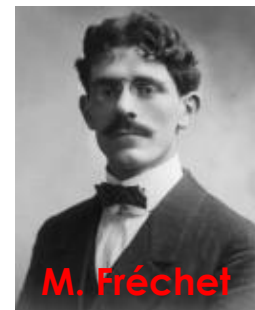
$$\{A, B, H\} = \arg \underset{\{A, B, H\}}{\text{Min}} \sum_{i=1}^N d_{\text{geodesique}}^p(\{a_i, b_i, h_i\}, \{A, B, H\})$$



$p=2$: Moyenne, $p=1$ Median

- Les paramètres d'un triangle $\{a, b, h\}$ rectangle sont contraints par la l'équation de Pythagore $h^2 = a^2 + b^2$.
- Cette contrainte se « déploie » en un cône convexe.
- 1 triangle rectangle est représenté par 1 point sur la le cone convexe

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



Maurice Fréchet, Darmois, Cramer, Fisher, Levy, Blanc-Lapierre

Congrès Calcul des probabilités – Genève 1938 & Lyon 1948

in
ed.



A Congress on Probability in Geneva (Congrès, Calcul des Probabilités, 1938). Along the front you see Lévy, R. A. Fisher, and Georges Darmois. On the stairs are Cramér, Fréchet, and Jean Piaget.



Fig. 9 Colloque International sur le Calcul des Probabilités, Lyon 1948. First row: Paul Lévy and Maurice Fréchet. On the picture one can find among others J. Doob, R. Fortet, D. Van Dantzig, E. Mourier, J. Kampé de Fériet, A. Blanc-Lapierre.... (Photo: © Private collection F. Lederer)

This
part

Maurice Fréchet et l'équation de Clairaut

Travaux précurseurs de Maurice Fréchet

- En 1939, dans son cours de l'IHP, Maurice Fréchet introduit ce qui fut appelée ensuite borne de Cramer-Rao

$$(\sigma_T)^2 \geq \frac{1}{n(\sigma_A)^2} \text{ avec } T = H(X_1, \dots, X_n), \quad A = \frac{1}{f(X, \theta)} \frac{\partial f(X, \theta)}{\partial \theta}$$

$\hat{\theta}$ estimateur de θ , borne de Fréchet : $R_\theta = E\left[(\theta - \hat{\theta})(\theta - \hat{\theta})^T\right] \geq I(\theta)^{-1}$

$$[I(\theta)]_{i,j} = -E\left[\frac{\partial^2 \log P(Z/\theta)}{\partial \theta_i \partial \theta_j}\right] = E\left[\frac{\partial \log P(Z/\theta)}{\partial \theta_i} \frac{\partial \log P(Z/\theta)}{\partial \theta_j}\right]$$

- Dans son article de 1943, Maurice Fréchet s'intéresse aux "**densités distinguées**", densités qui atteignent cette borne. Il montre qu'elles dépendent d'une fonction (logarithme de la fonction de partition) qui vérifie **l'équation de Alexis Clairaut**.

(55)

$$\mu = \theta \mu' - \psi(\mu')$$

$$\Phi^*(x^*) = \left\langle x^*, \frac{d\Phi^*(x^*)}{dx^*} \right\rangle - \Phi\left(\frac{d\Phi^*(x^*)}{dx^*}\right) \text{ et } x^* = \frac{d\Phi(x)}{dx}$$

c'est-à-dire une équation de Clairaut. La solution $\mu' = \text{constante}$ réduirait $f(x, \theta)$, d'après (48) à une fonction indépendante de θ , cas où le problème n'aurait plus de sens. μ est donc donné par la solution singulière de (55), qui est unique et s'obtient en éliminant s entre $\mu = \theta s - \psi(s)$ et $\theta = \psi'(s)$ ou encore entre

2ème partie de l'article de Fréchet: l'équation de Clairaut

- A cette étape, Fréchet cherche les « **densités distinguées** », toute densité de probabilité $p_\theta(x)$ telle que la fonction suivante soit indépendante de θ :

$$h(x) = \theta + \frac{\frac{\partial \log p_\theta(x)}{\partial \theta}}{\int_{-\infty}^{+\infty} \left[\frac{\partial p_\theta(x)}{\partial \theta} \right]^2 \frac{dx}{p_\theta(x)}}$$

- L'objectif de Fréchet est alors de déterminer la fonction minimisante $T = H'(X_1, \dots, X_n)$ qui atteint la borne. L'identité précédente peut se réécrire :

$$\lambda(\theta) \frac{\partial \log p_\theta(x)}{\partial \theta} = h(x) - \theta$$

Inverse de la
matrice de Fisher

$$\lambda(x) = \left[\int_{-\infty}^{+\infty} \left[\frac{\partial p_\theta(x)}{\partial \theta} \right]^2 \frac{dx}{p_\theta(x)} \right]^{-1} = \left[\int_{-\infty}^{+\infty} p_\theta(x) \left[\frac{\partial \ln p_\theta(x)}{\partial \theta} \right]^2 dx \right]^{-1} = \left[E \left[\left[\frac{\partial \ln p_\theta(x)}{\partial \theta} \right]^2 \right] \right]^{-1}$$

2ème partie de l'article de Fréchet: l'équation de Clairaut

- Or comme $\lambda(\theta) > 0$, on peut considérer $\frac{1}{\lambda(\theta)}$ comme la dérivée seconde d'une fonction $\Phi(\theta)$ telle que :

$$\frac{\partial \log p_\theta(x)}{\partial \theta} = \frac{\partial^2 \Phi(\theta)}{\partial \theta^2} [h(x) - \theta] \quad \frac{1}{\lambda(x)} = \int_{-\infty}^{+\infty} p_\theta(x) \left[\frac{\partial \ln p_\theta(x)}{\partial \theta} \right]^2 dx = E \left[\left[\frac{\partial \ln p_\theta(x)}{\partial \theta} \right]^2 \right] = \frac{\partial^2 \Phi(\theta)}{\partial \theta^2}$$

- dont on déduit que la quantité suivante est indépendante de θ :

$$\ell(x) = \log p_\theta(x) - \frac{\partial \Phi(\theta)}{\partial \theta} [h(x) - \theta] - \Phi(\theta)$$

- Une **densité distinguée** sera donc de la forme :

$$p_\theta(x) = e^{\frac{\partial \Phi(\theta)}{\partial \theta} [h(x) - \theta] + \Phi(\theta) + \ell(x)} \quad \text{avec} \quad \int_{-\infty}^{+\infty} p_\theta(x) dx = 1$$

- Ces 2 conditions sont suffisantes

**Fréchet montre que ces densités sont forcément des densités exponentielles.
Fréchet remarque que la matrice de Fisher est égale au hessien d'une fonction
(fonction caractéristique de Massieu)**

Equation de Clairaut(-Legendre) de la géométrie de l'information découverte par Maurice Fréchet

L'équation de Clairaut(-Legendre) de Maurice Fréchet

➤ Partant de l'équation: $p_{\theta}(x) = e^{\frac{\partial\Phi(\theta)}{\partial\theta}[h(x)-\theta]+\Phi(\theta)+\ell(x)}$

➤ On peut prendre arbitrairement $h(x)$ et $\ell(x)$ et alors $\Phi(\theta)$ est déterminée par :

$$\int_{-\infty}^{+\infty} e^{\frac{\partial\Phi(\theta)}{\partial\theta}[h(x)-\theta]+\Phi(\theta)+\ell(x)} dx = 1 \quad \longrightarrow \quad e^{\theta \cdot \frac{\partial\Phi(\theta)}{\partial\theta} - \Phi(\theta)} = \int_{-\infty}^{+\infty} e^{\frac{\partial\Phi(\theta)}{\partial\theta} h(x) + \ell(x)} dx$$

➤ Si on fixe alors arbitrairement $h(x)$ et $\ell(x)$ et soit s une variable arbitraire, la fonction suivante sera une fonction positive connue représentée par $e^{\Psi(s)}$:

$$\int_{-\infty}^{+\infty} e^{s \cdot h(x) + \ell(x)} dx = e^{\Psi(s)}$$

➤ On obtient alors la fonction $\Phi(\theta)$ par l'équation : $\Phi(\theta) = \theta \cdot \frac{\partial\Phi(\theta)}{\partial\theta} - \Psi\left(\frac{\partial\Phi(\theta)}{\partial\theta}\right)$

➤ Fréchet remarque qu'il s'agit de **l'équation d'Alexis Clairaut**. $\Phi(\theta)$ est donné par la solution singulière de cette équation de Clairaut, qui est unique et s'obtient en éliminant s entre :

$$\Phi = \theta \cdot s - \Psi(s) \quad \text{et} \quad \theta = \frac{\partial\Psi(s)}{\partial s} \quad \text{avec} \quad \Psi(s) = \log \int_{-\infty}^{+\infty} e^{s \cdot h(x) + \ell(x)} dx$$

La géométrie de l'information pour les familles de densité exponentielles

ALEAE GEOMETRIA

Géométrie du hasard



Blaise Pascal

La borne de Cramer-Rao-Fréchet-Darmois et la métrique de Fisher

La borne Cramer-Rao –Fréchet-Darmois a été introduite par Fréchet en 1939 et par Rao en 1945 comme l'inverse de la matrice de Fisher : $I(\theta)$

$$R_{\hat{\theta}} = E\left[(\theta - \hat{\theta})(\theta - \hat{\theta})^+\right] \geq I(\theta)^{-1} \quad [I(\theta)]_{i,j} = -E\left[\frac{\partial^2 \log p_{\theta}(z)}{\partial \theta_i \partial \theta_j^*}\right]$$

Rao a proposé d'introduire une métrique Riemannienne dans l'espace des paramètres des densités de probabilité (axiomatisé par N. Chentsov):

$$ds_{\theta}^2 = \text{Kullback_Divergence}(p_{\theta}(z), p_{\theta+d\theta}(z))$$

$$ds_{\theta}^2 = -\int p_{\theta}(z) \log \frac{p_{\theta+d\theta}(z)}{p_{\theta}(z)} dz$$

$$ds_{\theta}^2 \underset{\text{Taylor}}{\approx} \sum_{i,j} g_{ij} d\theta_i d\theta_j^* = \sum_{i,j} [I(\theta)]_{i,j} d\theta_i d\theta_j^* = d\theta^+ \cdot I(\theta) \cdot d\theta$$

$$w = W(\theta)$$

$$\Rightarrow ds_w^2 = ds_{\theta}^2$$

Distance entre gaussiennes au sens de la géométrie de l'information

Matrice de Fisher de l'Information pour le cas gaussien:

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

► Fisher matrix induced the following differential metric :

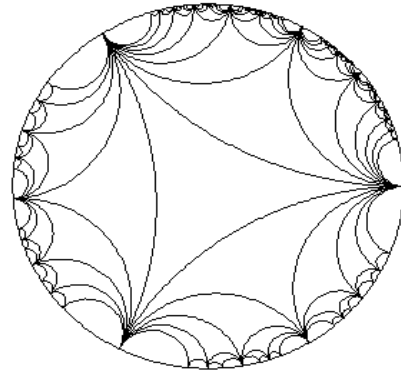
$$ds^2 = d\theta^T \cdot I(\theta) \cdot d\theta = \frac{dm^2}{\sigma^2} + 2 \cdot \frac{d\sigma^2}{\sigma^2} = \frac{2}{\sigma^2} \left[\left(\frac{dm}{\sqrt{2}} \right)^2 + (d\sigma)^2 \right]$$

► Modèle de géométrie hyperbolique de Poincaré

$$z = \frac{m}{\sqrt{2}} + i \cdot \sigma \quad \omega = \frac{z - i}{z + i} \quad (|\omega| < 1)$$

$$\Rightarrow ds^2 = 8 \cdot \frac{|d\omega|^2}{(1 - |\omega|^2)^2}$$

La géométrie paramètres des gaussiennes est la géométrie du disque de Poincaré



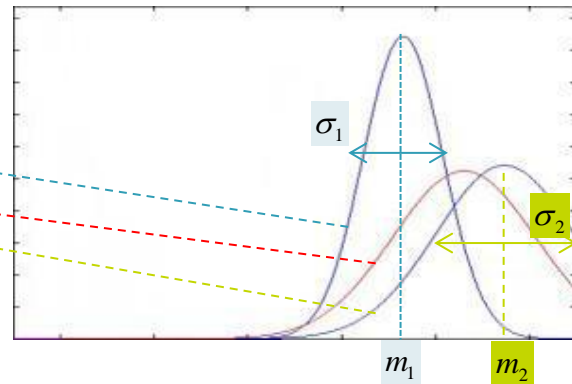
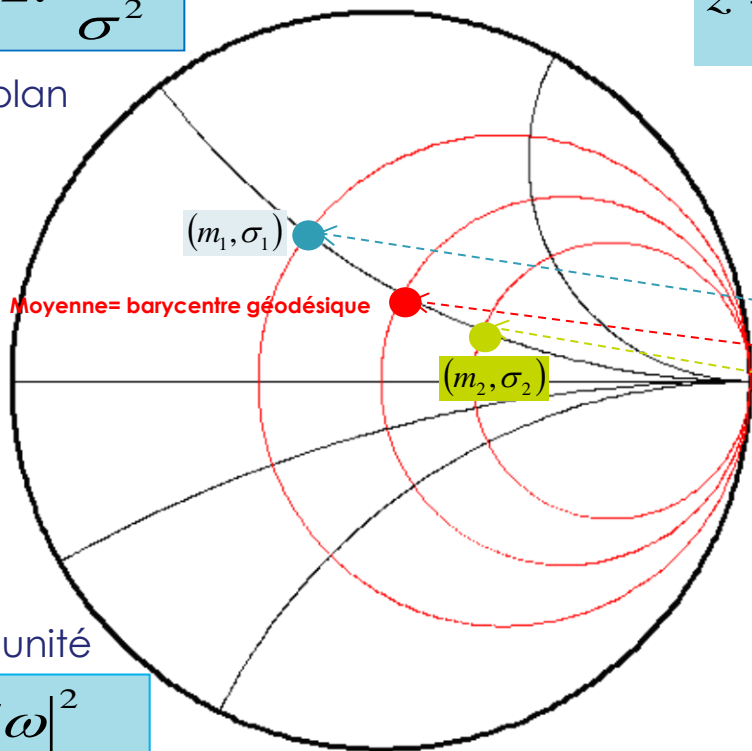
1 gaussiennes monovarié = 1 point dans le disque de Poincaré

$$ds^2 = \frac{dm^2}{\sigma^2} + 2 \cdot \frac{d\sigma^2}{\sigma^2}$$

Métrie demi-plan

$$z = \frac{m}{\sqrt{2}} + i \cdot \sigma$$

$$\omega = \frac{z - i}{z + i} \quad (|\omega| < 1)$$



Métrie disque unité

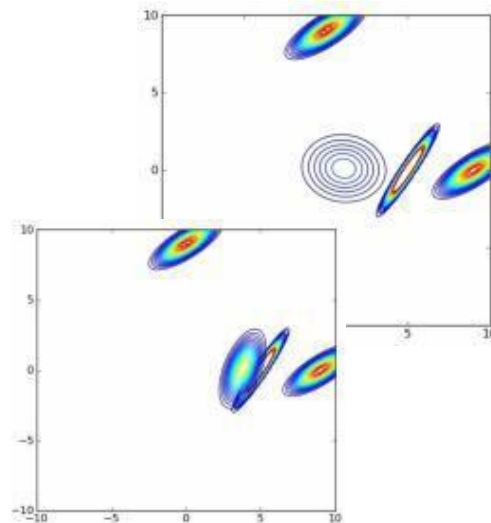
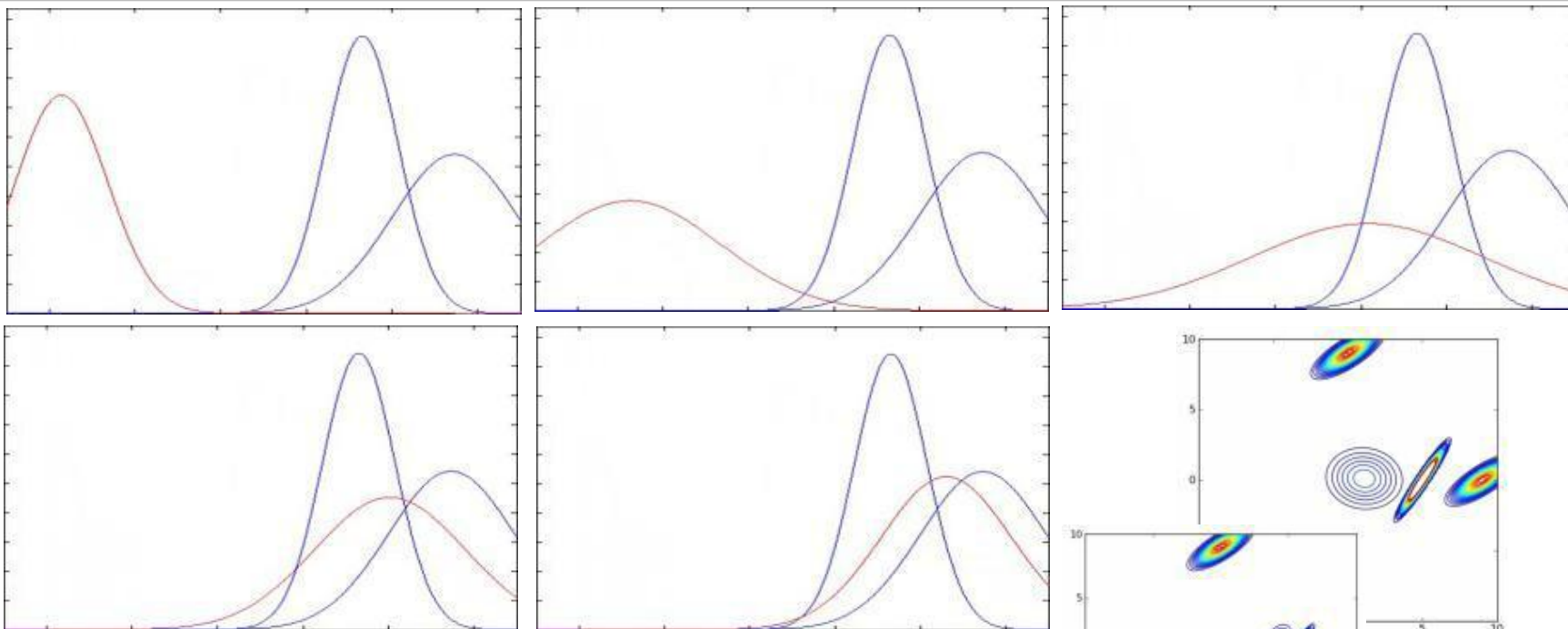
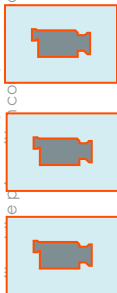
$$ds^2 = 8 \cdot \frac{|d\omega|^2}{(1 - |\omega|^2)^2}$$

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

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

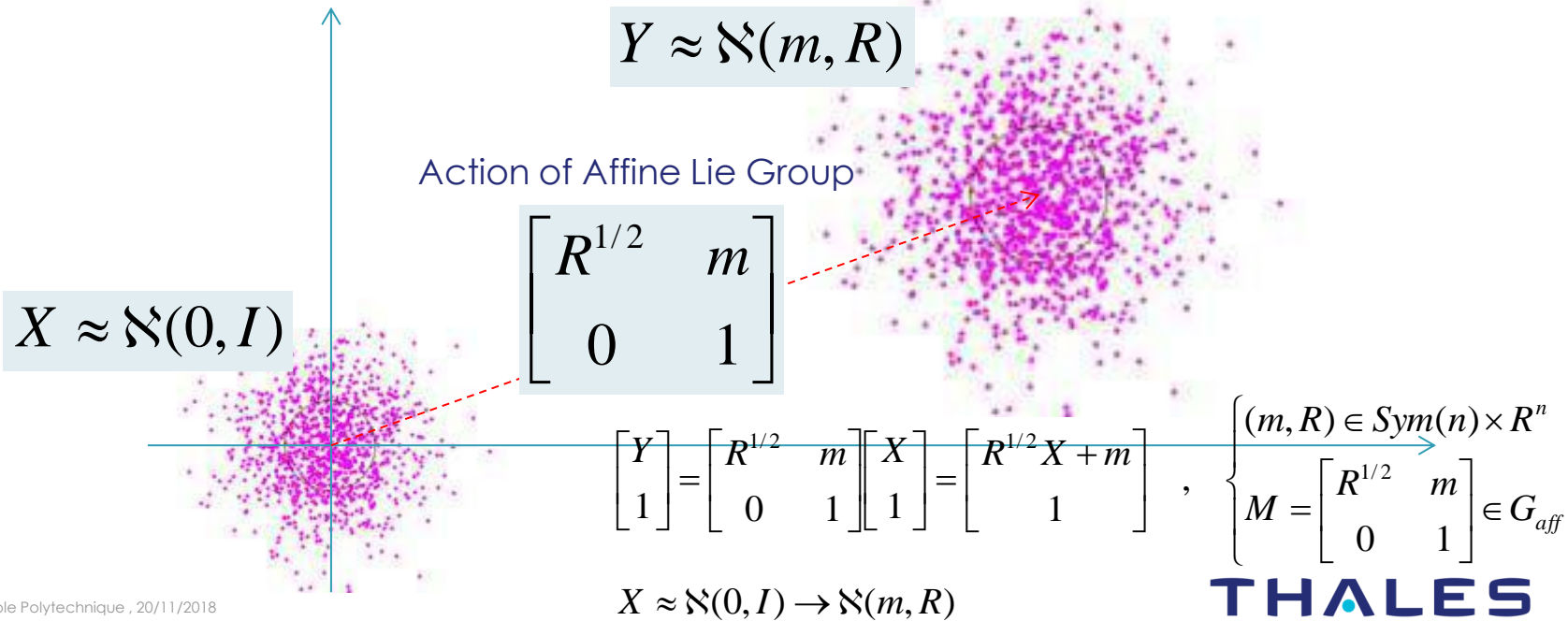
Barycentre de Fréchet entre gaussienne = barycentre dans le disque (voir flot de Karcher dans le disque unité précédent)

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales. © Thales 2015 All rights reserved.



Quel est le groupe qui agit pour les densités multivariées gaussiennes

La racine carrée de la matrice de covariance R symétrique définie positive est donnée par la factorisation de Cholesky $R = LL^T = R^{1/2} (R^{1/2})^T$ qui fait apparaître une matrice triangulaire avec des éléments positifs sur la diagonale (élément d'un groupe)



Géométrie de l'Information pour une gaussienne multivariée

La métrique de Fisher pour une gaussienne multivariée (m, R)

$$ds^2 = \sum_{ij} g_{ij} d\theta_i d\theta_j = dm^T R^{-1} dm + \frac{1}{2} \text{Tr} \left[(R^{-1} dR)^2 \right]$$

Les géodésiques sont données par les équations d'Euler-Lagrange

$$\sum_{i=1}^n g_{ik} \ddot{\theta}_i + \sum_{i,j=1}^n \Gamma_{ijk} \dot{\theta}_i \dot{\theta}_j = 0 \quad , \quad k = 1, \dots, n$$

avec $\Gamma_{ijk} = \frac{1}{2} \left[\frac{\partial g_{jk}}{\partial \theta_i} + \frac{\partial g_{jk}}{\partial \theta_j} + \frac{\partial g_{ij}}{\partial \theta_k} \right]$

$$\Rightarrow \begin{cases} \ddot{R} + \dot{m} \dot{m}^T - \dot{R} R^{-1} \dot{R} = 0 \\ \ddot{m} - \dot{R} R^{-1} \dot{m} = 0 \end{cases}$$

Avec les symétries, les invariants (invariant de Noether ou application moment de Souriau) réduisent les équations à celles d'Euler-Poincaré:

$$\frac{d\Pi_R}{dt} = \left[\begin{array}{c} \frac{d(R^{-1} \dot{R} + R^{-1} \dot{m} \dot{m}^T)}{dt} \\ 0 \end{array} \quad \frac{d(R^{-1} \dot{m})}{dt} \right] = 0 \Rightarrow \begin{cases} R^{-1} \dot{R} + R^{-1} \dot{m} \dot{m}^T = B = cste \\ R^{-1} \dot{m} = b = cste \end{cases} \Rightarrow \begin{cases} \dot{m} = Rb \\ \dot{R} = R(B - b \dot{m}^T) \end{cases}$$

Tirs géodésiques pour calculer les distances entre 2 gaussiennes multivariées

Principe du tir géodésique:

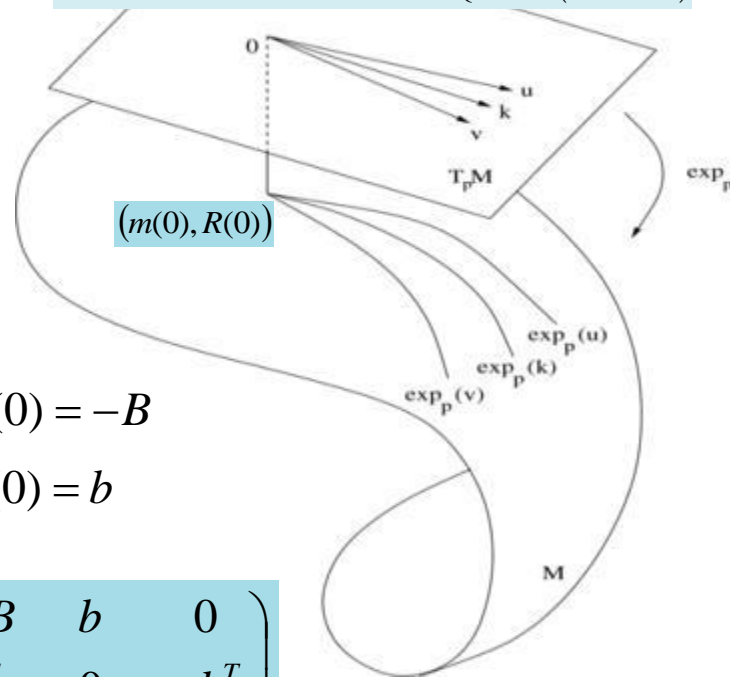
- Fixer un vecteur tangent au point $(m(0), R(0))$
- Corriger le tir, par transport parallèle de l'erreur
- Après convergence la distance est

$$d = \sqrt{\dot{m}(0)^T R^{-1}(0) \dot{m}(0) + \frac{1}{2} \text{Tr} \left[\left(R^{-1}(0) \dot{R}(0) \right)^2 \right]}$$

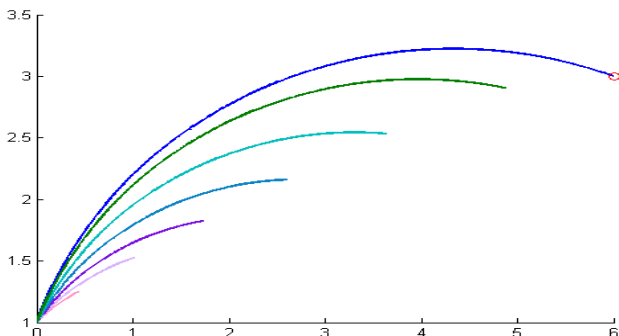
$$\begin{cases} \Delta(t) = R^{-1}(t) \\ \delta(t) = R^{-1}(t)m(t) \end{cases} \Rightarrow \begin{cases} \dot{\Delta} = -B\Delta + bm^T \\ \dot{\delta} = -B\delta + (1 + \delta^T \Delta^{-1} \delta) b \\ \Delta(0) = I_p, \delta(0) = 0 \end{cases} \text{ with } \begin{cases} \dot{\Delta}(0) = -B \\ \dot{\delta}(0) = b \end{cases}$$

$$\Lambda(t) = \exp(tA) = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!} = \begin{pmatrix} \Delta & \delta & \Phi \\ \delta^T & \varepsilon & \gamma^T \\ \Phi^T & \gamma & \Gamma \end{pmatrix} \text{ with } A = \begin{pmatrix} -B & b & 0 \\ b^T & 0 & -b^T \\ 0 & -b & B \end{pmatrix}$$

Equations d'Euler - Poincaré $\begin{cases} \dot{m} = Rb \\ \dot{R} = R(B - bm^T) \end{cases}$



Tirs géodésiques pour calculer les distances entre 2 gaussiennes multivariées



Principe de corrections du tir géodésique

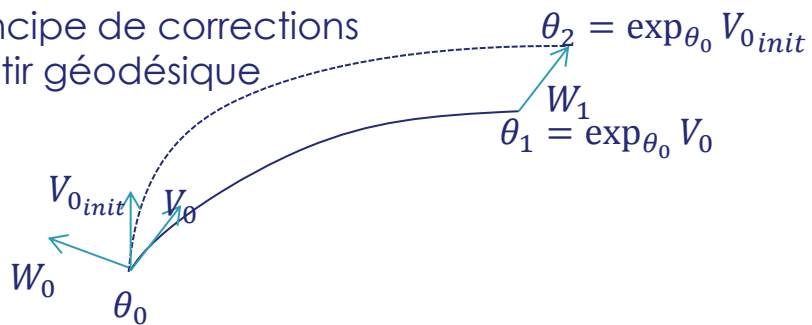
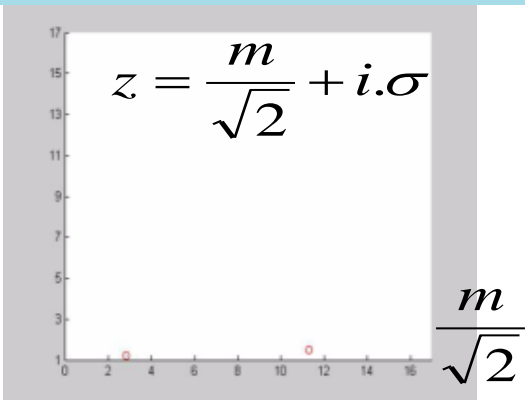


Illustration du tir géodésique pour des gaussiennes de dimension 1 et 2

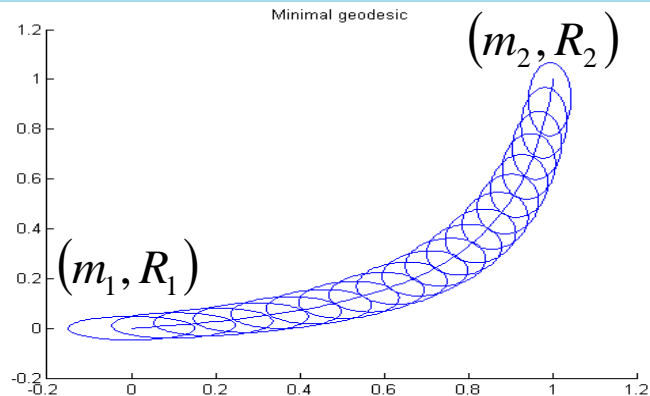
Tir dans le demi-plan de Poincaré pour gaussienne dim 1

$i \cdot \sigma$

$$z = \frac{m}{\sqrt{2}} + i \cdot \sigma$$



Géodésique pour gaussienne dim 2



Machine Learning Approach: Unsupervised Method by « Mean-Shift » extended on Riemannian Manifold

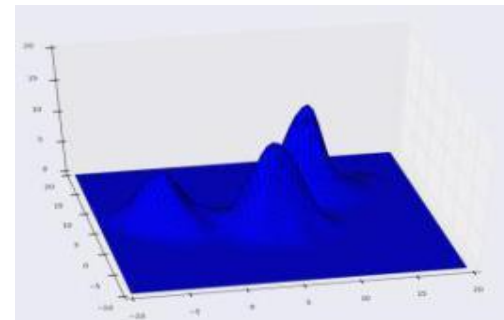
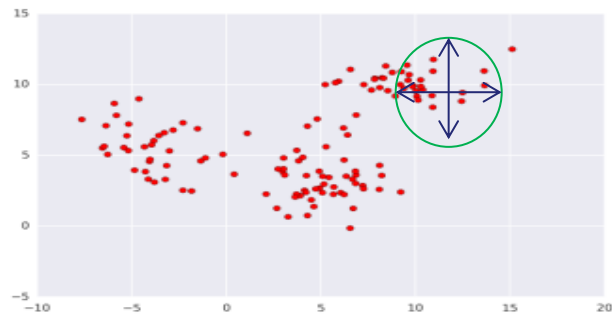


Blaise Pascal

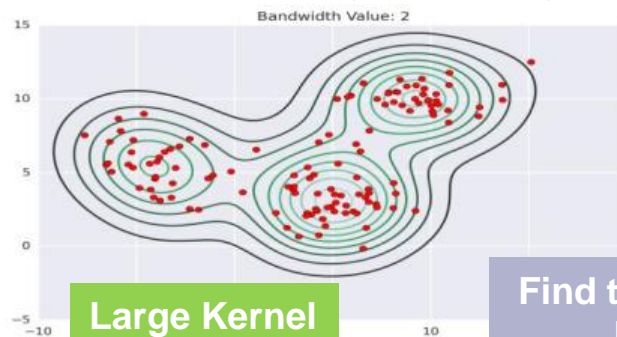
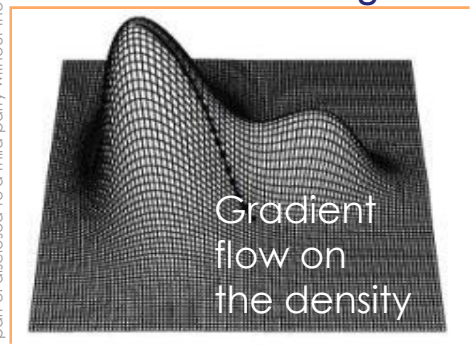
Classical « Kernel Method » for density estimation & « Mean-Shift » for clustering

Kernel Method for non-parametric density estimation

- Kernel Method: selection of an optimal kernel size and counting of points per cell

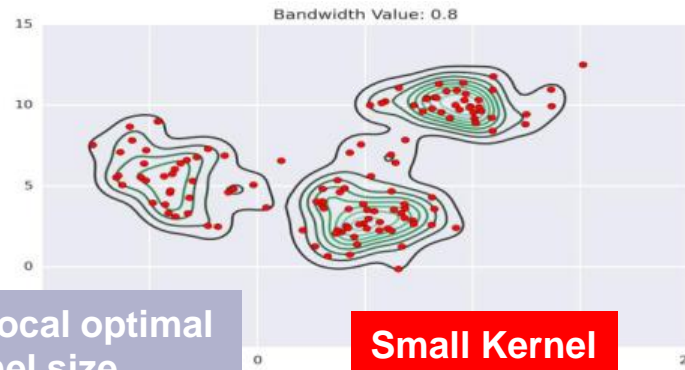


- Mean-Shift: gradient flot on the density to identify mode



Large Kernel
(3 classes)

Find the local optimal
kernel size



Small Kernel
(7 classes)

Riemannian Kernel Density Estimation of Doppler Spectrum Statistics

- We introduce a kernel density estimation for Riemannian manifold on the elements of the product $(P_0, \mu_1, \dots, \mu_n) \in \mathbb{R}^+ \times D_n$, to estimate density for Doppler Spectrum.
- The specificity of the hyperbolic space enables to adapt the different density estimation methods at a reasonable cost.
- Recently convergence rates for the kernel density estimation without the compact assumption have been introduced, which enables the use of Gaussian-type kernels.
- Reference:
 - Emmanuel Chevallier, Probability density estimation on the hyperbolic space applied to radar processing, SPRINGER, GSI'15 Lecture Note
 - <https://hal.archives-ouvertes.fr/hal-01121090/document>

Riemannian Kernel Method of Density Estimation

Let $K : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be a map which verifies the following properties:

$$\int_{\mathbb{R}^n} K(\|x\|).dx = 1, \int_{\mathbb{R}^n} x.K(\|x\|).dx = 0, K(x > 1) = 0, \text{Sup}(K(x)) = 0$$

Given a point $p \in H_n$ (the hyperbolic space of dimension n ; $H_2=D$), the exponential map \exp_p defines a new injective parametrization of H_n . The Lebesgue measure of the tangent space is noted Leb_p . The function $\theta_p : H_n \rightarrow \mathbb{R}^+$ defined by:

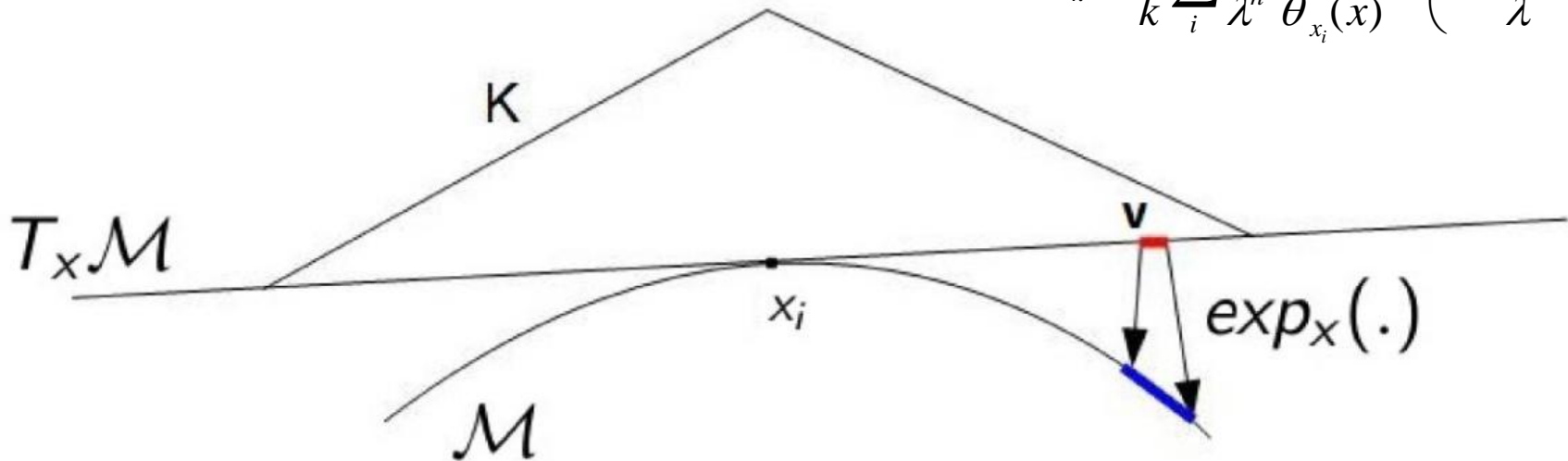
$$\theta_p : q \rightarrow \theta_p(q) = \frac{dvol}{d \exp_p^* (Leb_p)}(q)$$

is the density of the Riemannian measure with respect to the image of the Lebesgue measure of $T_p H_n$ by \exp_p . Given K and a scaling parameter λ , the estimator of f proposed by Pelletier is defined by:

$$\hat{f}_k = \frac{1}{k} \sum_i \frac{1}{\lambda^n} \frac{1}{\theta_{x_i}(x)} K\left(\frac{d(x, x_i)}{\lambda}\right)$$

Volume change due to Exponential map

$$\hat{f}_k = \frac{1}{k} \sum_i \frac{1}{\lambda^n} \frac{1}{\theta_{x_i}(x)} K\left(\frac{d(x, x_i)}{\lambda}\right)$$



$$\theta_p : q \rightarrow \theta_p(q) = \frac{d\text{vol}}{d \exp_p^*(\text{Leb}_p)}(q)$$

Riemannian Kernel Method of Density Estimation

Given $p_{ref} \in H_n$, μ_k the empirical measure and “*” the natural convolution on homogeneous spaces, let:

$$\tilde{K}(q) = \frac{1}{k\lambda^n} \frac{1}{\theta_{p_{ref}}(q)} K\left(\frac{d(p_{ref}, q)}{\lambda}\right) \quad \text{then} \quad \hat{f}_k = \mu_k * \tilde{K}$$

One still needs to obtain an explicit expression of θ_p . Given a reference point p , the point of polar coordinates (r, α) of the hyperbolic space H_n is defined as point at distance r of p on the geodesic with initial direction $\alpha \in S^{n-1}$. Since H_n is isotropic the expression the length element in polar coordinates depends only on r . Expressed in polar coordinates the hyperbolic metric expression is:

$$g_{H_n} = dr^2 + \sinh(r)^2 \cdot g_{S^{n-1}}$$

The polar coordinates are a polar expression of the exponential map at p . In an adapted orthonormal basis of the tangent plane the metric takes the following form:

$$G = \begin{pmatrix} 1 & 0 \\ 0 & \sinh(r)^2 \frac{1}{r^2} I_{n-1} \end{pmatrix}$$

Riemannian Kernel Method of Density Estimation

the metric takes the following form:

$$G = \begin{pmatrix} 1 & 0 \\ 0 & \sinh(r)^2 \frac{1}{r^2} I_{n-1} \end{pmatrix}$$

where G is the matrix of the metric and I_{n-1} is the identity matrix of size $n-1$. The volume $dvol$ is given by:

$$dvol = \sqrt{|G|} \cdot d \exp_p^* (Leb_p) = \left(\frac{1}{r} \sinh(r) \right)^{n-1} d \exp_p^* (Leb_p)$$

where $r = d(p, q)$. Finally, one obtains:

$$\theta_p(q) = \left(\frac{1}{r} \sinh(r) \right)^{n-1} \hat{f}_k = \frac{1}{k} \sum_i \frac{1}{\lambda^n} \frac{d(x, x_i)^{n-1}}{\sinh(d(x, x_i))^{n-1}} K\left(\frac{d(x, x_i)}{\lambda} \right)$$

Extension of Kernel Method for Siegel Unit Disk

Density in Siegel Unit Disk using Cartan Coordinates:

$$\hat{f}_k = \frac{1}{k} \sum_j \frac{1}{r^n} \prod_{i < j} \frac{\tau_i - \tau_j}{\sinh(\tau_i - \tau_j)} \prod_{i \leq j} \frac{\tau_i + \tau_j}{\sinh(\tau_i + \tau_j)} K \left(\frac{\left(2 \sum_i \tau_j^2 \right)^{1/2}}{r} \right)$$

Deduced from Hua-Cartan Decomposition on Siegel Unit Disk (Lemma of Hua for radial coordinates in Siegel Disk):

$$\tau = [\tau_1 \quad \tau_2 \quad \dots \quad \tau_n] \text{ with } 0 \leq \tau_n \leq \tau_{n-1} \leq \dots \leq \tau_1$$

$$A_0(\tau) = \text{diag}[ch(\tau_1) \quad ch(\tau_2) \quad \dots \quad ch(\tau_n)]$$

$$B_0(\tau) = \text{diag}[sh(\tau_1) \quad sh(\tau_2) \quad \dots \quad sh(\tau_n)]$$

$$g = \begin{bmatrix} A & B \\ B^* & A^* \end{bmatrix} \in Sp(n), g = \begin{bmatrix} U^t & 0 \\ 0 & U^+ \end{bmatrix} \begin{bmatrix} A_0 & B_0 \\ B_0 & A_0 \end{bmatrix} \begin{bmatrix} V^* & 0 \\ 0 & V \end{bmatrix}$$

there exist U and V unitary complex matrices of order n

$$\begin{cases} A = U^t A_0(\tau) V^* \\ B = U^t B_0(\tau) V \end{cases} \Rightarrow \begin{pmatrix} A_0(\tau) & B_0(\tau) \\ B_0(\tau) & A_0(\tau) \end{pmatrix} = \exp \begin{pmatrix} 0 & Z_2(\tau) \\ Z_2(\tau) & 0 \end{pmatrix}$$

$$\text{with } Z_2(\tau) = \text{diag}[\tau_1 \quad \tau_2 \quad \dots \quad \tau_n]$$

$$\text{Let } Z = B(A^*)^{-1} = U^t P U, P^2 = B_0^2(A_0^{-1})^2 = \text{diag}[\text{eigen}(ZZ^+)]$$

$$P = \text{diag}[th(\tau_1) \quad th(\tau_2) \quad \dots \quad th(\tau_n)]$$

Density Estimation with Riemannian Kernel Method

$$\begin{aligned} (x_i)_{i=1:k} &\in M \\ r &\in R_+^* \end{aligned}$$

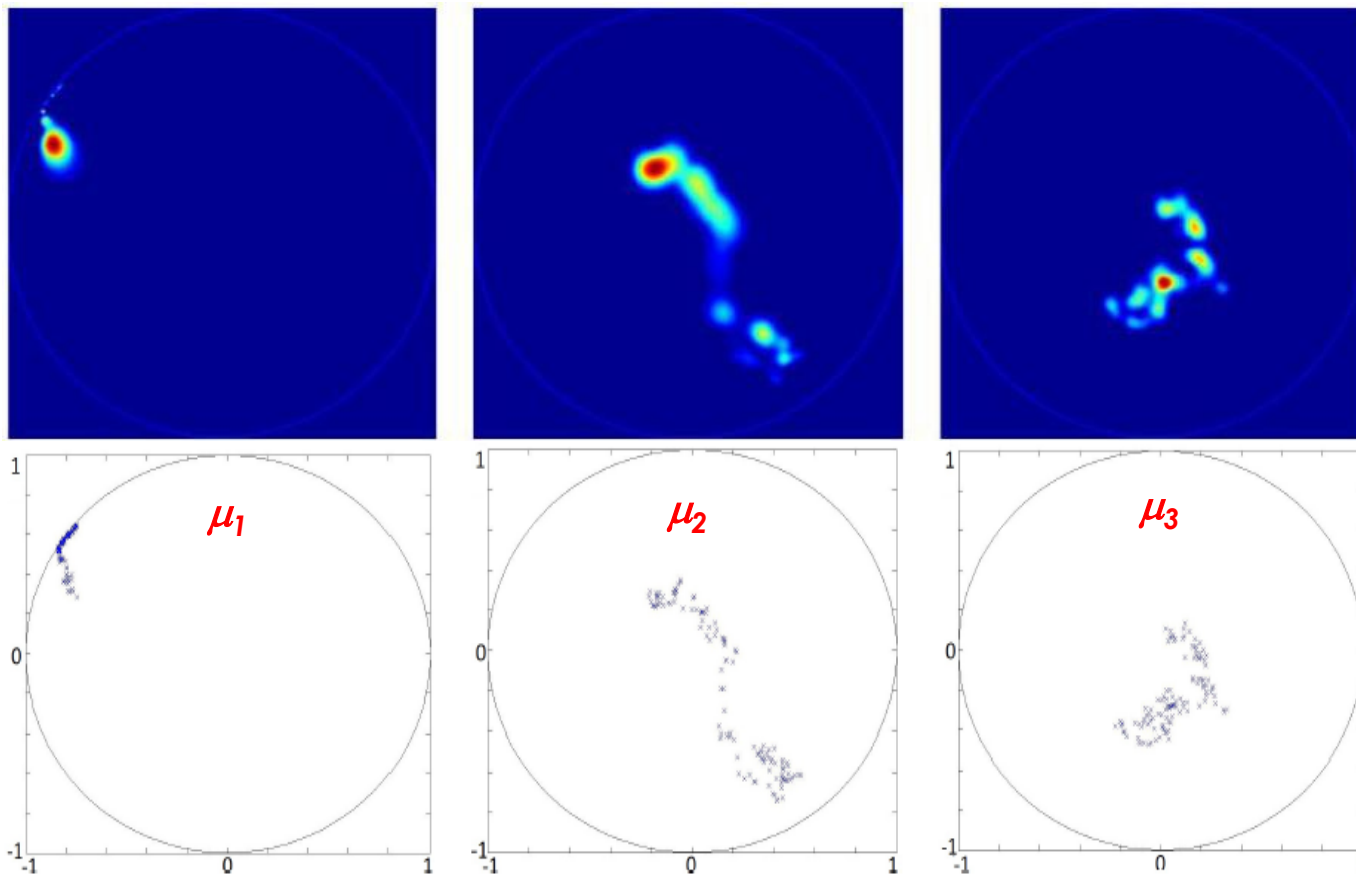
$$\hat{f}_k^r(x) = \frac{1}{k} \sum_{i=1:k} \frac{1}{r^n} \frac{1}{\theta_{x_i}(x)} K\left(\frac{\text{dist}(x, x_i)}{r}\right)$$

How to optimise Kenel size « r »

Minimise the product of densities evaluated for each point by cross-validation:

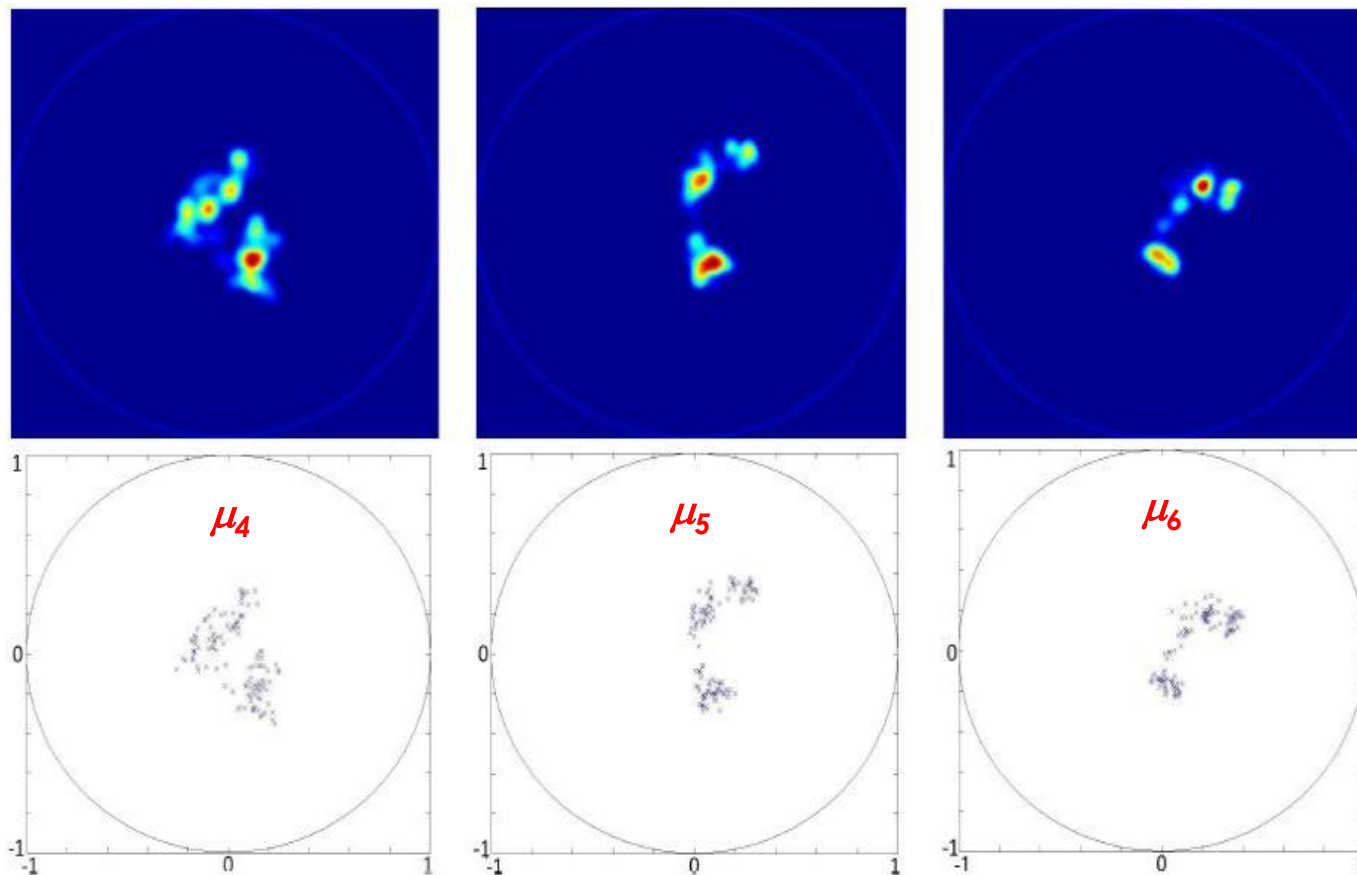
$$\begin{aligned} &\min_r \prod_{i=1:k} \hat{f}_{-i}^r(\Sigma_i) \\ \text{or} &\min_r \sum_{i=1:k} \log(\hat{f}_{-i}^r(\Sigma_i)) \end{aligned}$$

Riemannian Kernel Method of Density Estimation in Poincaré Disk 1/2



This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - © Thales 2015 All rights reserved.

Riemannian Kernel Method of Density Estimation in Poincaré Disk 2/2



This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - © Thales 2015 All rights reserved.

Mean-Shift on Riemannian Manifold

- The original mean shift algorithm is widely applied for nonparametric clustering of data in vector spaces. In this section, we will generalize it to data points lying on Riemannian manifolds of reflection coefficients.
- This allows us to extend mean shift based clustering to Doppler Data by reflection coefficient.
- Mean shift is provided by following gradient equation where the $\log_y(x_i)$ terms lie in the tangent space, and the kernel terms K are scalars. The mean shift vector is a weighted sum of tangent vectors, and is itself a tangent vector. The mean shift iteration is: $y_{j+1} \exp_{y_j}(m_\lambda(y_j))$

■ with : $g(\cdot) = -K'(\cdot)$

$$m_\lambda(y) = \left[\sum_{i=1}^n g\left(\frac{d(y, x_i)}{\lambda}\right) \right]^{-1} \sum_{i=1}^n g\left(\frac{d(y, x_i)}{\lambda}\right) \log_y(x_i)$$

Non-Linear Mean-Shift

For $i=1:n$

- $y=x_i$

- Repeat:

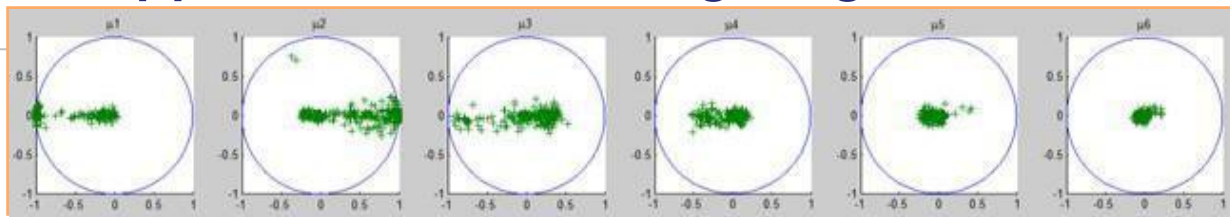
- $$m_r(y) \leftarrow \frac{\sum_{i=1}^n g(d^2(y, x_i)/r^2) \log_y(x_i)}{\sum_{i=1}^n g(d^2(y, x_i)/r^2)}$$
 with $g = -(K)'$

- $$y \leftarrow \exp_y(m_r(y))$$

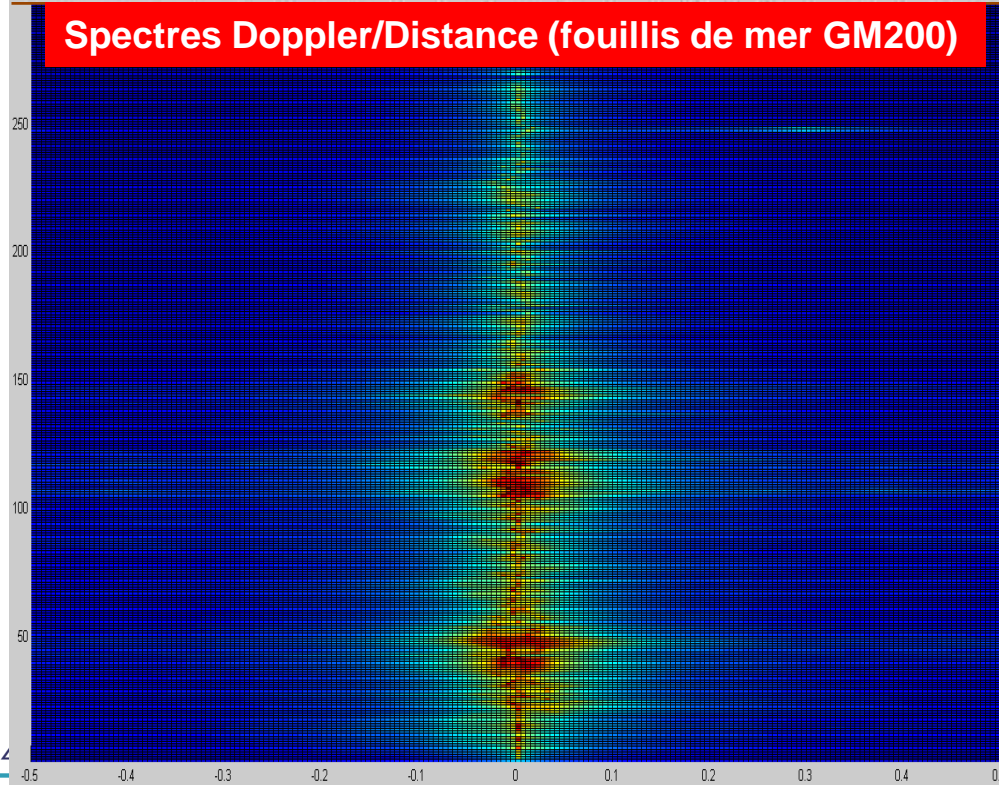
until $\|m_r(y)\| \leq \varepsilon$

- y is the local mode

Doppler Fluctuation along range axis

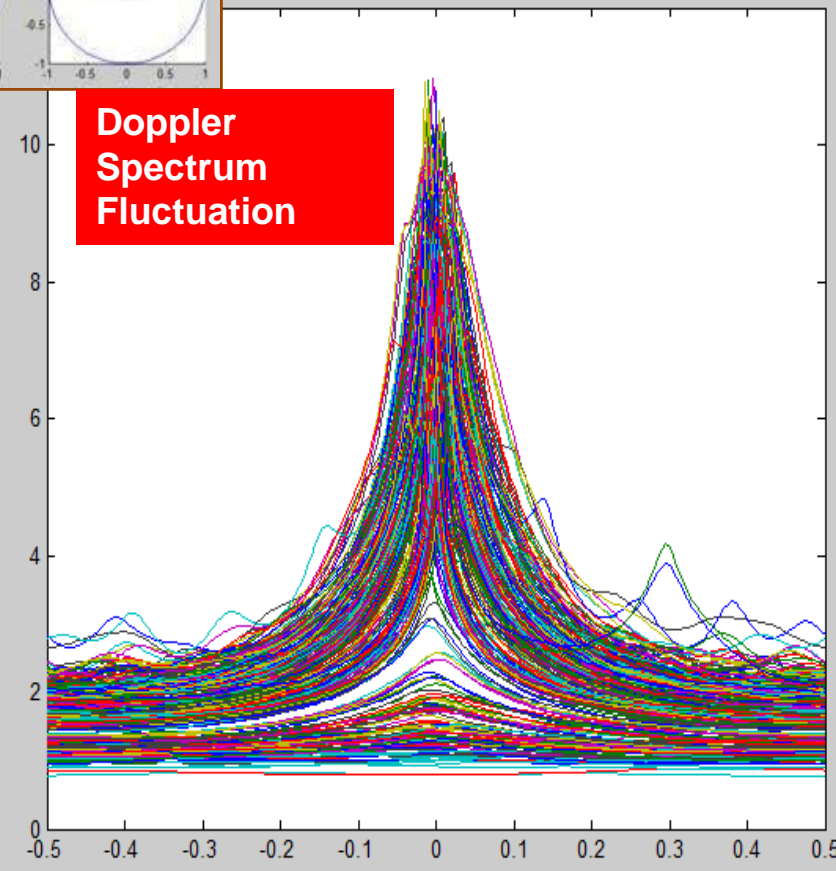


Spectres Doppler/Distance (fouillis de mer GM200)



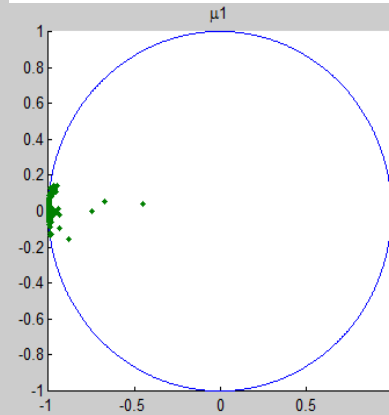
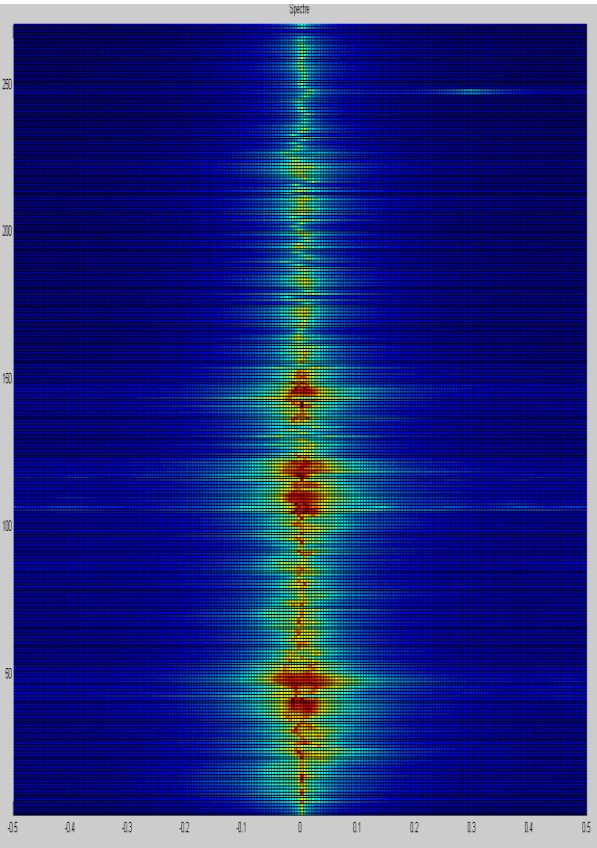
Spectre=f(r)

Doppler Spectrum Fluctuation

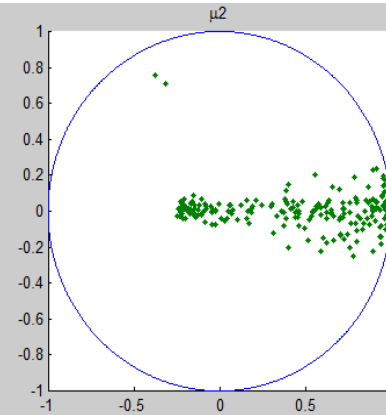


This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - © Thales 2015 All rights reserved.

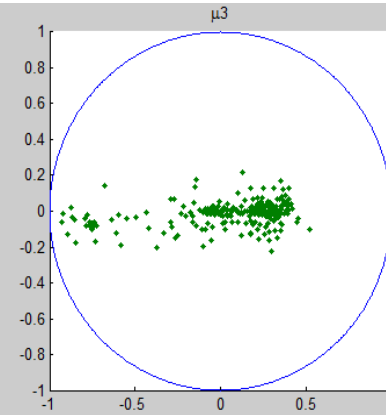
Doppler Spectrum and Reflection Coefficients



1st reflection coefficient



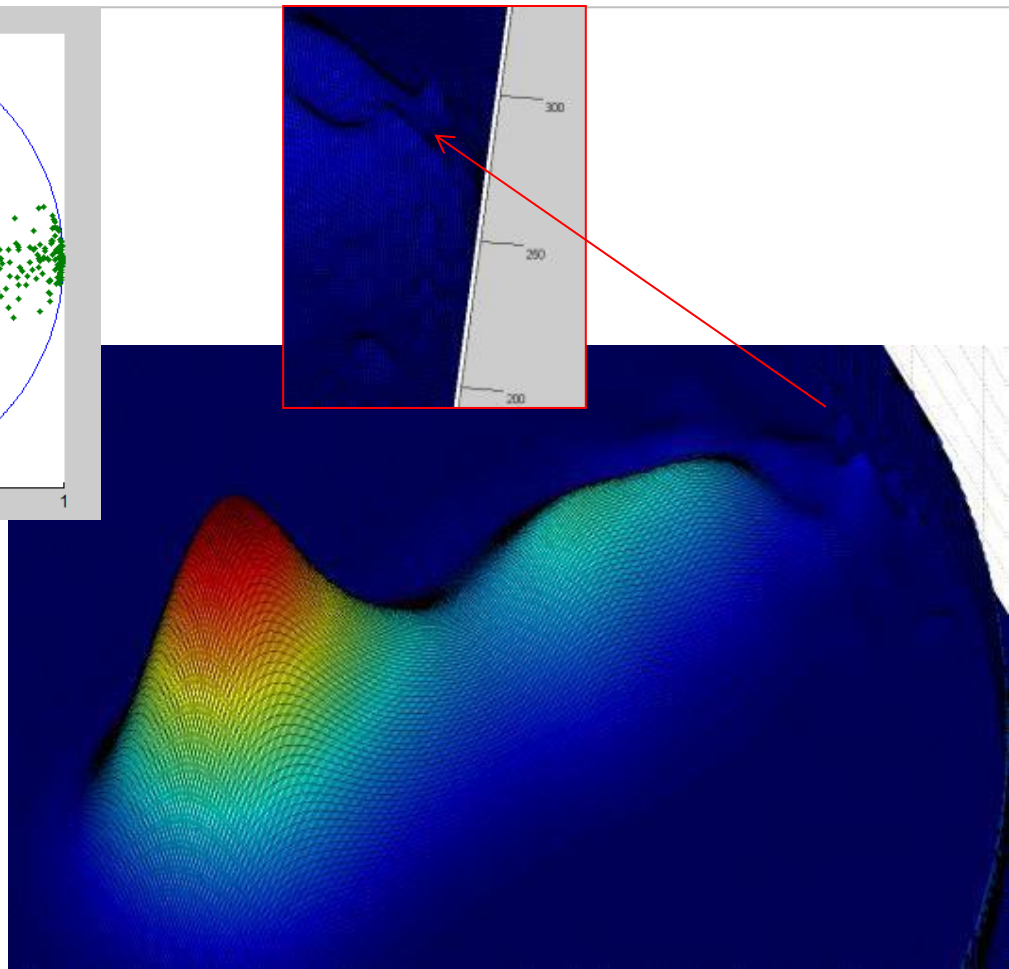
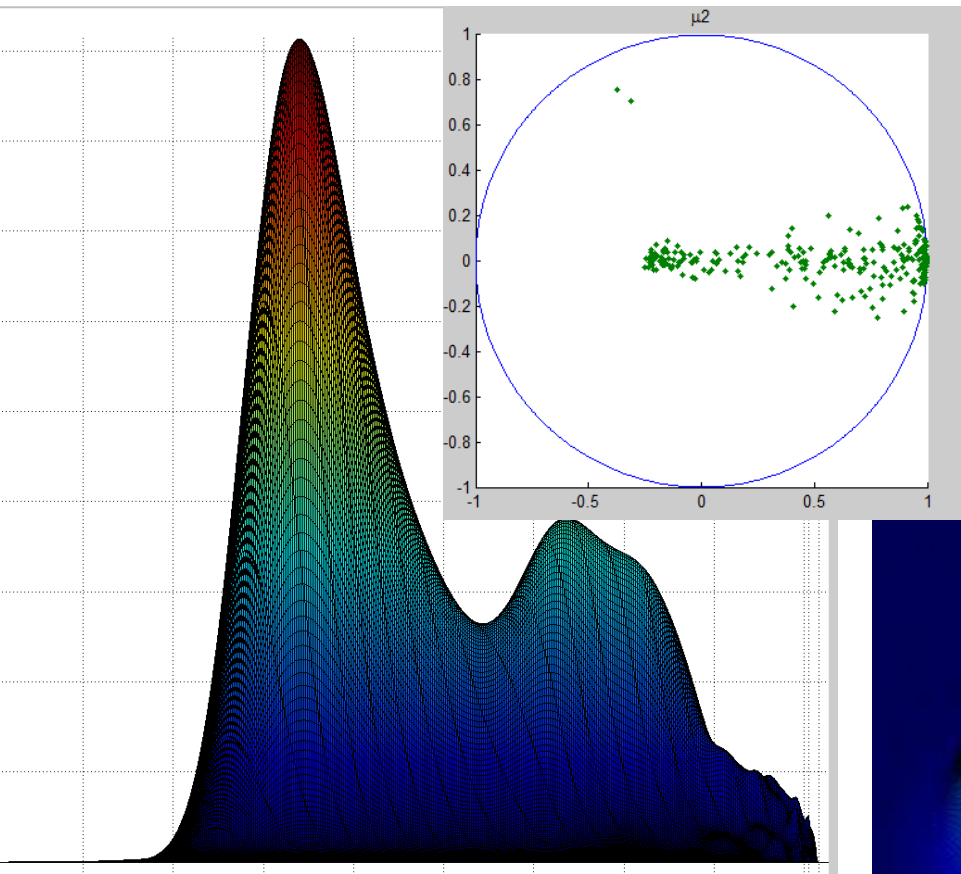
2nd reflection coefficient

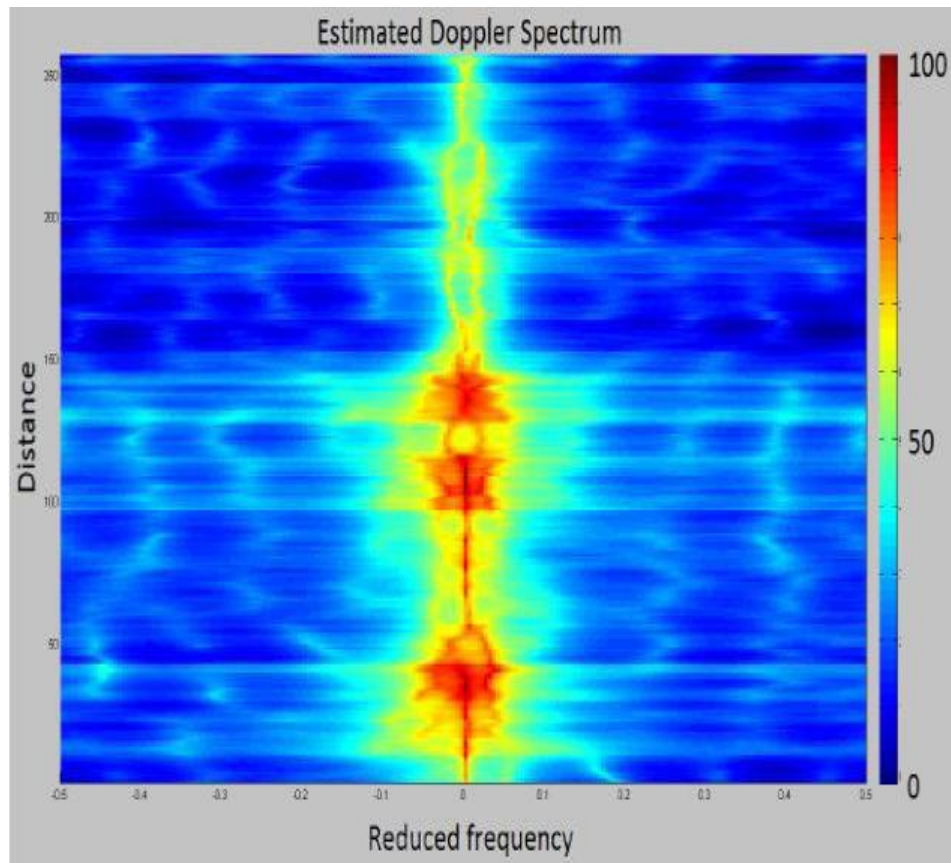
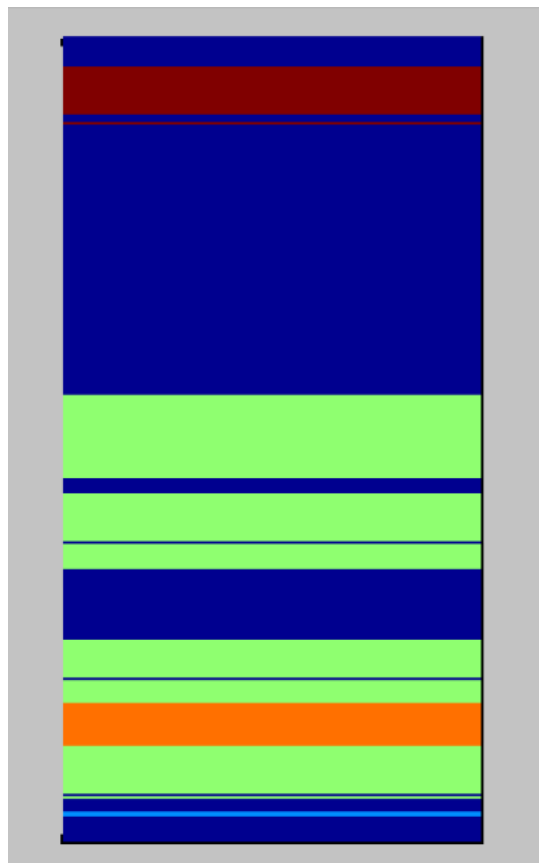


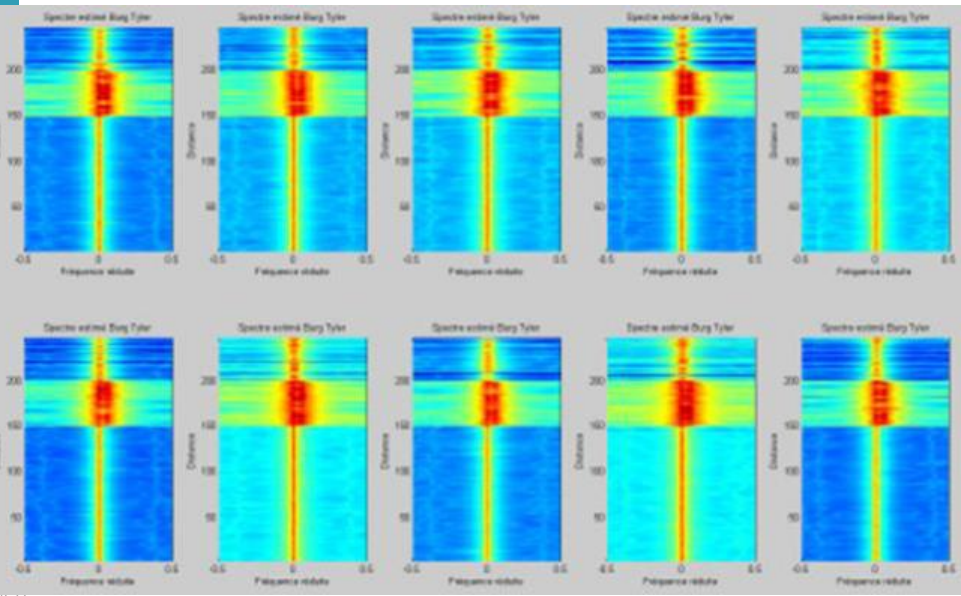
3rd reflection coefficient

Density of 2nd reflection coefficient estimated by kernel method

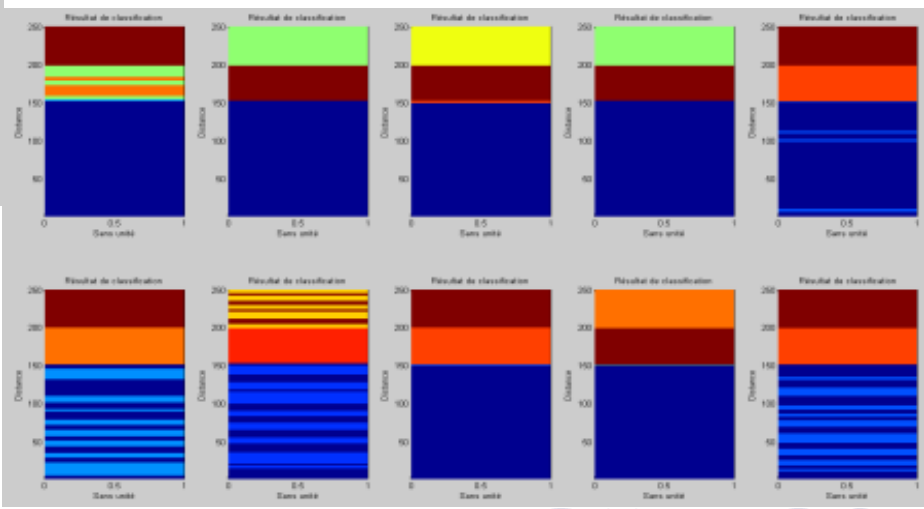
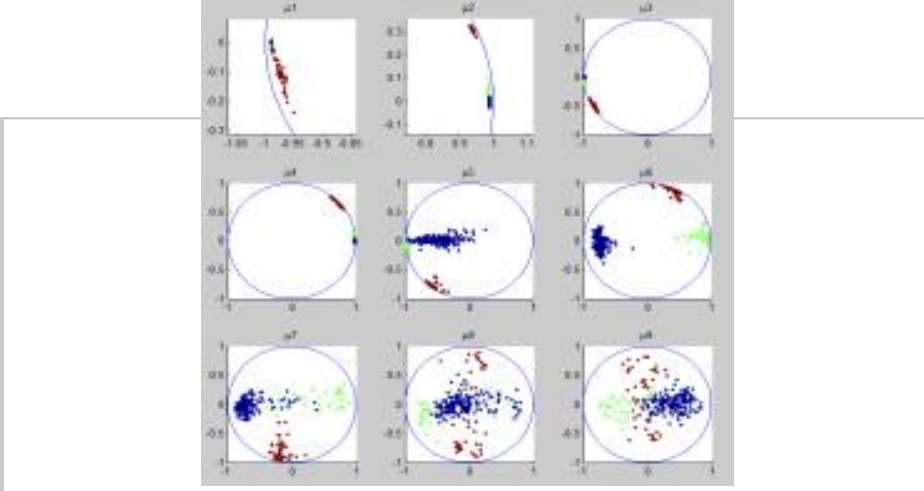
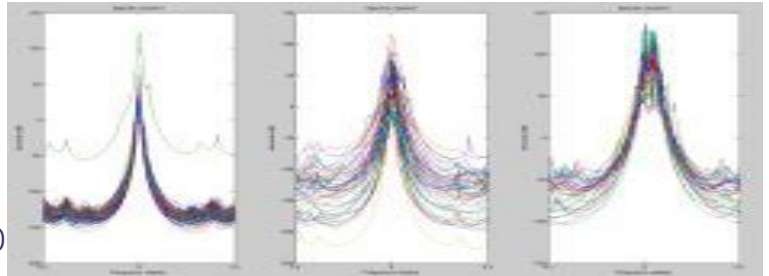
This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part, without the prior written permission of the Institut Polytechnique de Paris.

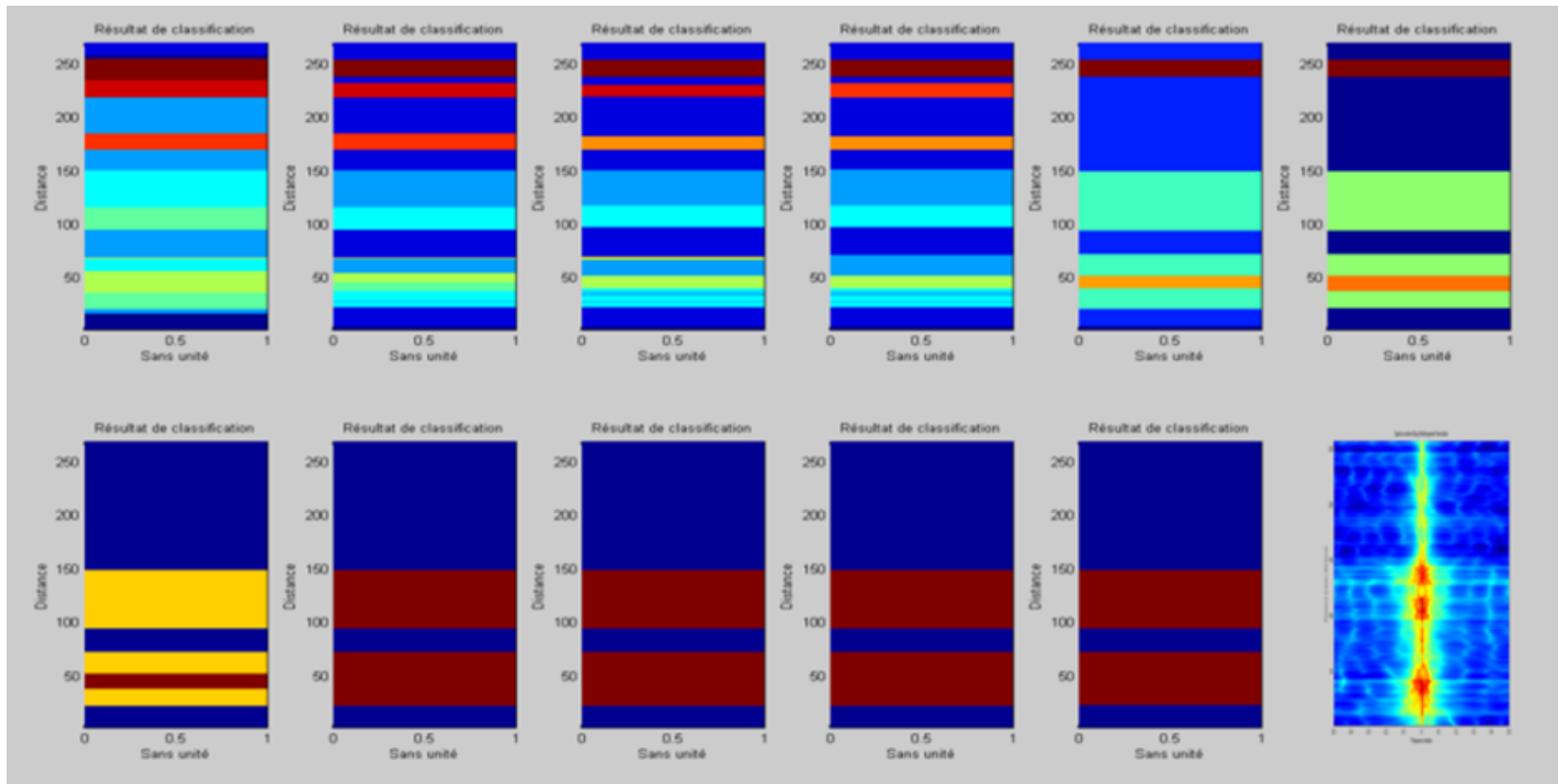




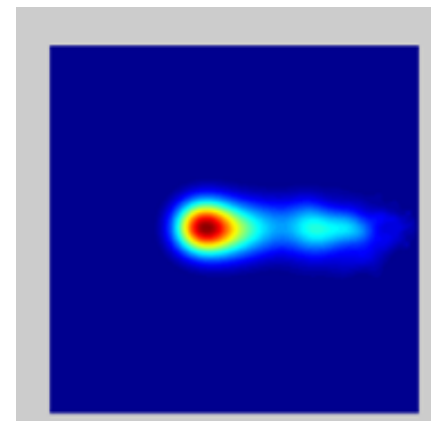
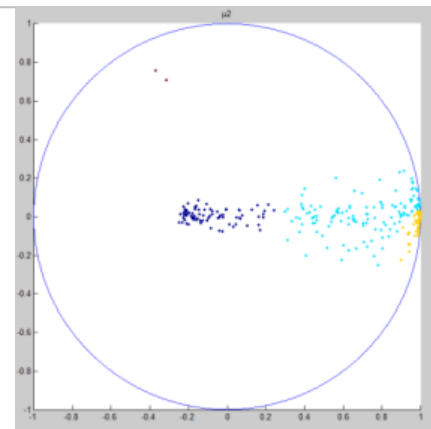
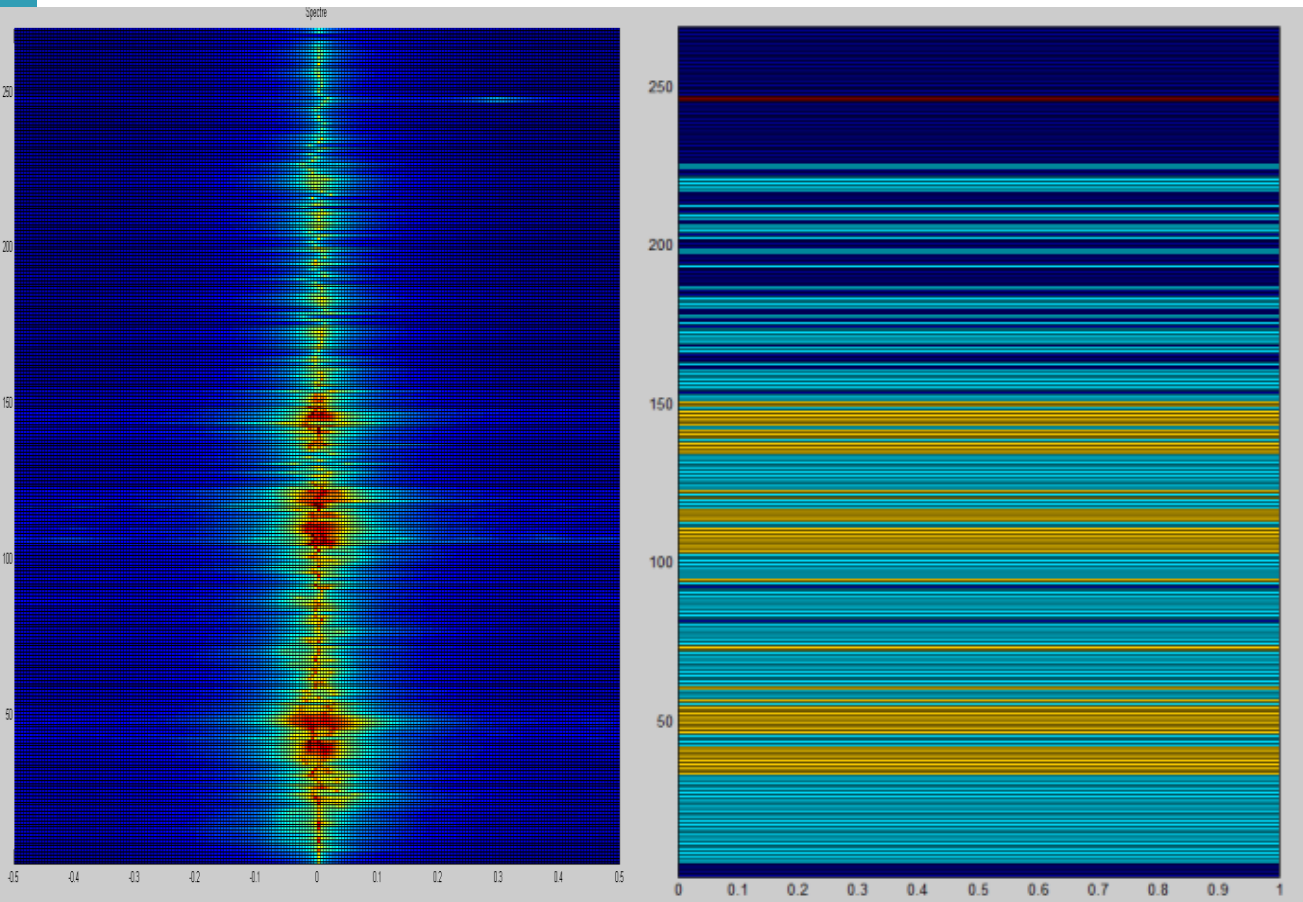


Doppler Spectrum by Burg Tyler on 5 cases of simulated data (Ground clutter on 50 cases, Ground+Sea clutters on 50 cases, and Sea clutter on 50 cases)



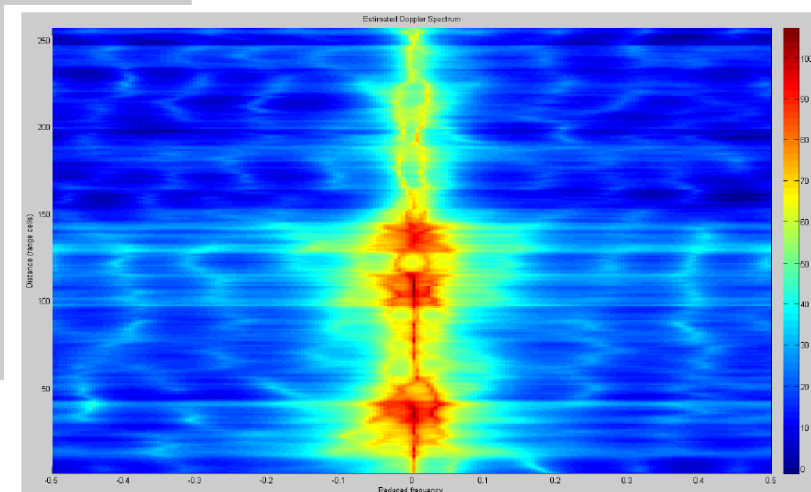
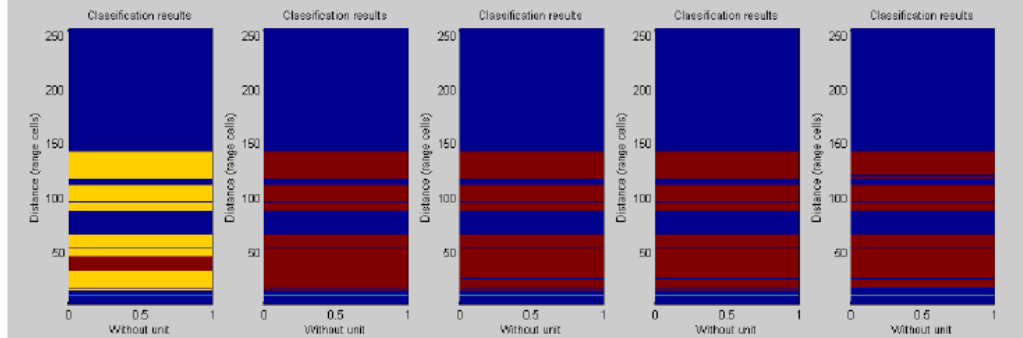
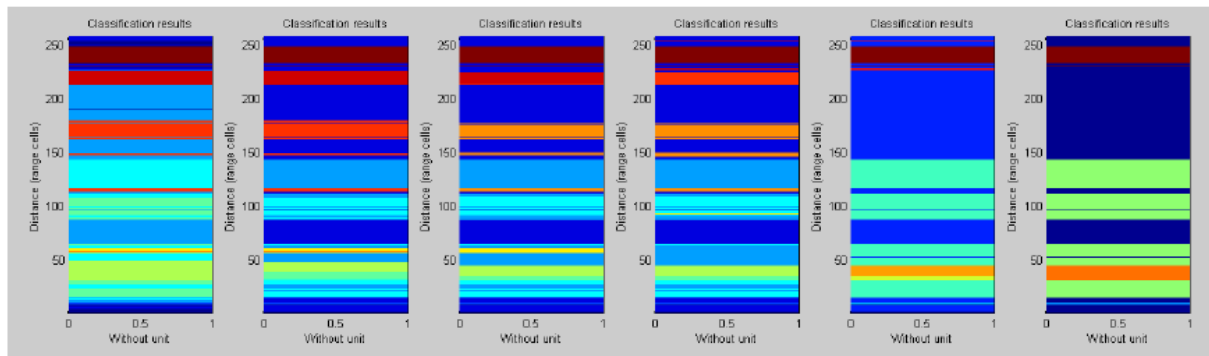


Mean-Shift Segmentation on 2nd reflection coefficients: 4 classes



Mean-Shift Segmentation of Real Range/Doppler Data

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales - © Thales 2015 All rights reserved.



Future Research : Statistical Depth for THPD and TBTHPD matrices

Exploration of statistical depth for Covariances matrices

➤ Existing works on HPD matrices:

- Intrinsic data depth for Hermitian positive definite matrices, Joris Chau, Hernando Ombao, Rainer von Sachs
- <https://arxiv.org/abs/1706.08289>

➤ Extension for THPD and TBTHPD matrices, the notion of statistical depth

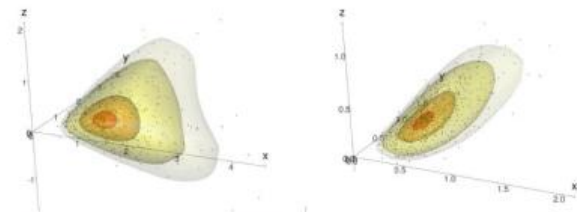
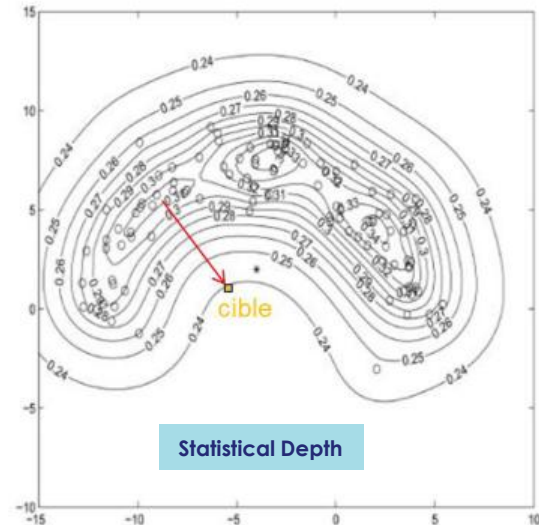


Figure 1: $100(1 - \alpha)\%$ intrinsic zonoid depth regions with $\alpha = \{0.025, 0.25, 0.75, 0.9\}$ for random SPD matrices, ($d = 2, n = 500$), from a Riemannian log-normal distribution (left) and a rescaled Wishart distribution (right), with intrinsic mean μ shown by the green cube.



THALES

Synthèse

www.thalesgroup.com

ALEAE GEOMETRIA

Géométrie du hasard



Blaise Pascal

Que faut-il retenir après avoir tout oublié

Devant une bonne choucroute au jambon, ils oublièrent le pudding de graisse de phoque farci aux myrtilles ! — (Jean-Baptiste Charcot, Dans la mer du Groenland, 1928)

- La découverte de la borne inférieure sur la variance de tout estimateur est à attribuer à Maurice Fréchet lors de l'hiver 1939 (cours de l'IHP), 6 ans avant Rao. Borne que nous appellerons dorénavant **Borne de Fréchet**.
- L'article séminale de 1943 n'introduit pas seulement la borne de Fréchet, mais l'étude des **densités distinguées**, densités dont les paramètres atteignent cette borne. Fréchet montre que ces densités sont forcément des densités exponentielles.
- Fréchet remarque que la **matrice de Fisher est égale au hessien d'une fonction** intervenant dans son équation de Clairaut. Cette fonction c'est le logarithme de la fonction de partition (c'est la **fonction caractéristique de François Massieu**).
- Fréchet montre que ces densités distinguées sont définies par l'intermédiaire de l'**Equation de Clairaut(-Legendre)**, qui met en dualité 2 fonctions (entropie et fonction caractéristique).
- Les structures des densités distinguées et l'équation de Clairaut-Fréchet sont les structures fondamentales de la **Géométrie de l'Information**, basée sur la **géométrie hessienne de J.L. Koszul**
- Jean-Marie Souriau a généralisé cette structure dans le cas d'une variété homogène en introduisant une **« Thermodynamique des groupes de Lie »**. La Densité de Gibbs est covariante et la métrique est invariante sous l'action du groupe. La métrique de Fisher est lié à la **2-forme de Souriau-Kostant-Kirillov**.

Que faut-il retenir après avoir tout oublié

Devant une bonne choucroute au jambon, ils oublièrent le pudding de graisse de phoque farci aux myrtilles ! — (Jean-Baptiste Charcot, Dans la mer du Groenland, 1928)

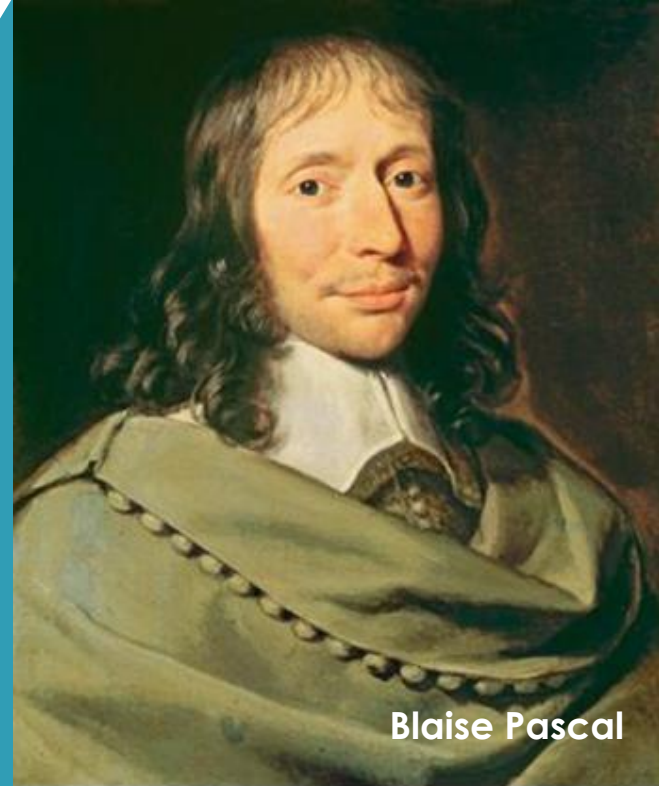
- On peut étendre les probabilités et les statistiques dans les **espaces abstraits** grâce aux travaux de **Maurice Fréchet** (**centralité = barycentre géodésique**)
- Pour faire des probabilités/statistiques sur des mesures de l'expérience dans les **espaces métriques**, la **métrique de Fisher** (celle associée à la borne de Fréchet-Darmois-Cramer-Rao) est la plus naturelle (**invariance par reparamétrisation**): **métrique = Hessien de l'Entropie**
- Pour étendre la **notion de lois gaussiennes**, il faut pouvoir étendre la définition d'une **densité à Maximum d'Entropie (de Gibbs)** dans les espaces abstraits. Pour les **espaces homogènes** (sur lequel un groupe agit), les équations sont données par la **Thermodynamique des groupes de Lie** de **Jean-Marie Souriau**.
- Les structures géométriques élémentaires de l'information de l'onde Electromagnétique sont codées dans $R \times S^1 \times D^{m-1} \times SD^{n-1}$

Le dernier mot à Souriau - Questions

« Il est évident que l'on ne peut définir de valeurs moyennes que sur des objets appartenant à un espace vectoriel (ou affine); donc - si bourbakiste que puisse sembler cette affirmation - que l'on n'observera et ne mesurera de valeurs moyennes que sur des grandeurs appartenant à un ensemble possédant physiquement une structure affine. Il est clair que cette structure est nécessairement unique sinon les valeurs moyennes ne seraient pas bien définies. » - J.M. Souriau

« Les différentes versions de la science mécanique peuvent se classer par la géométrie que chacune implique pour l'espace et le temps ; géométrie qui se détermine par le groupe de covariance de la théorie. Ainsi la mécanique newtonienne est covariante par le groupe de Galilée; la relativité restreinte par le groupe de Lorentz-Poincaré ; la relativité générale par le groupe « lisse » (le groupe des difféomorphismes de l'espace-temps). Il existe cependant une partie des énoncés de la mécanique dont la covariance appartient à un quatrième groupe – rarement envisagé : le groupe affine. ... Comment se fait-il qu'un point de vue unitaire, (qui serait nécessairement une véritable Thermodynamique), ne soit pas encore venu couronner le tableau ? Mystère...» - J.M. Souriau

Reference Bibliographique



Blaise Pascal

- [1] Balian, R., Alhassid, Y. and Reinhardt, H. Dissipation in many-body systems: a geometric approach based on information theory. Phys. Reports, 131, 1986, 1-146.
- [2] Balian, R., Valentin P., Hamiltonian structure of thermodynamics with gauge, Eur. Phys. J. B 21, 2001, pp. 269-282.
- [3] Balian, R. The entropy-based quantum metric, Entropy, Vol.16, n°7, 2014, pp.3878-3888.
- [4] Balian, R., François Massieu et les potentiels thermodynamiques, Évolution des disciplines et histoire des découvertes, Académie des Sciences, Avril 2015.
- [5] Barbaresco, F. Koszul information geometry and Souriau Lie group thermodynamics, AIP Conf. Proc. n°1641, n° 74, 2015, Proceedings of MaxEnt'14 conference, Amboise, Septembre 2014
- [6] Barbaresco, F. Koszul Information Geometry and Souriau Geometric Temperature/Capacity of Lie Group Thermodynamics. Entropy, vol. 16, 2014, pp. 4521-4565. Published in the book Information, Entropy and Their Geometric Structures, MDPI Publisher, September 2015.
- [7] Barbaresco, F. Symplectic Structure of Information Geometry: Fisher Metric and Euler-Poincaré Equation of Souriau Lie Group Thermodynamics. In Geometric Science of Information, Second International Conference GSI 2015 Proceedings, Lecture Notes in Computer Science vol. 9389, Springer 2015, pp. 529–540.
- [8] Barbaresco, F., Geometric Theory of Heat from Souriau Lie Groups Thermodynamics and Koszul Hessian Geometry: Applications in Information Geometry for Exponential Families. Preprint soumis au Special Issue "Differential Geometrical Theory of Statistics", MDPI, Entropy, 2016.
- [9] Casalis, M. Familles exponentielles naturelles invariantes par un groupe. Thèse de l'Univ. Paul Sabatier, Toulouse, 1990
- [10] Casalis, M. Familles Exponentielles Naturelles sur \mathbb{R}^d Invariantes par un Groupe, International Statistical Review (Revue Internationale de Statistique), Vol. 59, No. 2, pp. 241-262, Aug. 1991

- [11] Duhem, P. Sur les équations générales de la thermodynamique, Annales scientifiques de l'ENS, 3ème série, tome 9, 1891, pp.231-266.
- [12] Eriksen, P.S. Geodesics connected with the Fisher metric on the multivariate normal manifold. Lancaster : Proceedings of the GST Workshop, 1987
- [13] Fréchet, M.R. Sur l'extension de certaines évaluations statistiques au cas de petits échantillons. Revue de l'Institut International de Statistique 1943, vol. 11, n° 3/4, pp. 182–205.
- [14] Gromov, M. In a Search for a Structure, Part 1: On Entropy, 6 July 2012, Available online: <http://www.ihes.fr/~gromov/PDF/structre-serch-entropy-july5-2012.pdf> (accessed on 6 August 2015).
- [15] Gromov, M. Six Lectures on Probability, Symmetry, Linearity. Available online: <http://www.ihes.fr/~gromov/PDF/probability-huge-Lecture-Nov-2014.pdf> (accessed on 6 August 2015).
- [16] Koszul, J.L. Sur la forme hermitienne canonique des espaces homogènes complexes. Can. J. Math. 1955, 7, 562–576.
- [17] Koszul, J.L. Exposés sur les Espaces Homogènes Symétriques; Publicação da Sociedade de Matematica de São Paulo: São Paulo, Brazil, 1959.
- [18] Koszul, J.L. Domaines bornées homogènes et orbites de groupes de transformations affines. Bull. Soc. Math. Fr. 1961, 89, 515–533.
- [19] Koszul, J.L. Ouverts convexes homogènes des espaces affines. Math. Z. 1962, 79, 254–259.
- [20] Koszul, J.L. Variétés localement plates et convexité. Osaka. J. Math. 1965, 2, 285–290.

- [21] Koszul, J.L. Déformations des variétés localement plates. Ann. Inst. Fourier 1968, 18, 103–114.
- [22] Koszul, J.L. Trajectoires Convexes de Groupes Affines Unimodulaires. In Essays on Topology and Related Topics; Springer: Berlin, Germany, 1970; pp. 105–110.
- [23] Libermann, P.; Marle, C.M. Symplectic Geometry and Analytical Mechanics; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1987.
- [24] Marle, C.M. Symmetries of Hamiltonian Systems on Symplectic and Poisson Manifolds, 15th Int. Conf. on Geometry, Integ. and Quant., Sofia, 2014, pp.1-86.
- [25] Marle, C.M. From Tools in Symplectic and Poisson Geometry to Souriau's theories of Statistical Mechanics and Thermodynamics, MDPI Entropy, special issue on "Differential Geometrical Theory of Statistics", 2016
- [26] Massieu, F. Sur les Fonctions caractéristiques des divers fluides. Comptes Rendus de l'Académie des Sciences 1869, 69, 858–862.
- [27] Massieu, F. Addition au précédent Mémoire sur les Fonctions caractéristiques. Comptes Rendus de l'Académie des Sciences 1869, 69, 1057–1061.
- [28] Massieu, F. Exposé des principes fondamentaux de la théorie mécanique de la chaleur (note destinée à servir d'introduction au Mémoire de l'auteur sur les fonctions caractéristiques des divers fluides et la théorie des vapeurs), 31 p., S.I. - s.n., 1873
- [29] Massieu, F. Thermodynamique: Mémoire sur les Fonctions Caractéristiques des Divers Fluides et sur la Théorie des Vapeurs; Académie des Sciences: Paris, France, 1876; p. 92.
- [30] Poincaré, H. Sur les tentatives d'explication mécanique des principes de la thermodynamique. Comptes rendus de l'Académie des sciences 1889, 108, 550–553.

- [31] Poincaré, H. Thermodynamique, Cours de Physique Mathématique; Carré, G., Ed.; 1892; en ligne: <http://gallica.bnf.fr/ark:/12148/bpt6k2048983>
- [32] Poincaré, H. Calcul des Probabilités; Gauthier-Villars: Paris, France, 1896.
- [33] Poincaré, H. Réflexions sur la théorie cinétique des gaz. J. Phys. Theor. Appl. 1906, 5, 369–403.
- [34] de Saxcé, G., Entropy and Structure for the Thermodynamic Systems. In Geometric Science of Information, Second International Conference GSI 2015 Proceedings, (F. Nielsen and F. Barbaresco, editors), Lecture Notes in Computer Science, vol. 9389, Springer 2015, pp. 519–528.
- [35] de Saxcé, G. Link between Lie Group Statistical Mechanics and Thermodynamics of Continua, Special Issue MDPI ENTROPY "Differential Geometrical Theory of Statistics", 2016, 18, 254
- [36] de Saxcé, G.; Vallée, C. Galilean Mechanics and Thermodynamics of Continua; Wiley-ISTE: London, UK, 2016.
- [37] Shima, H. The Geometry of Hessian Structures; World Scientific: Singapore, Singapore, 2007.
- [38] Shima, H. Geometry of Hessian Structures. In Springer Lecture Notes in Computer Science; Nielsen, F., Frederic, B., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8085, pp. 37–55.
- [39] Souriau, J.M. Géométrie de l'espace de phases. Comm. Math. Phys. 1966, 374, 1–30
- [40] Souriau, J.M. Définition covariante des équilibres thermodynamiques. Suppl. Nuov. Cimento 1966, 1, pp.203–216. (In French)

- [41] Souriau, J.M. Structure des systèmes dynamiques; Editions Jacques Gabay: Paris, France, 1970.
- [42] Souriau, J.-M., Mécanique statistique, groupes de Lie et cosmologie, Colloques internationaux du CNRS numéro 237, Géométrie symplectique et physique mathématique, 1974, pp. 59–113.
- [43] Souriau, J.M. Géométrie Symplectique et Physique Mathématique; Éditions du C.N.R.S.: Paris, France, 1975.
- [44] Souriau, J.M. Thermodynamique Relativiste des Fluides; Centre de Physique Théorique: Marseille, France, 1977.
- [45] Souriau, J.M. Thermodynamique et géométrie. In Differential Geometrical Methods in Mathematical Physics II; Bleuler, K., Reetz, A., Petry, H.R., Eds.; Springer: Berlin/Heidelberg, Germany, 1978; pp. 369–397.
- [46] Souriau, J.M. Dynamic Systems Structure (chap.16 convexité, chap. 17 Mesures, chap. 18 Etats statistiques, Chap. 19 Thermodynamique), available in Souriau archive (document sent by C. Vallée), unpublished technical notes, 1980.
- [47] Souriau, J.M. Mécanique classique et géométrie symplectique. CNRS Marseille. Cent. Phys. Théor. 1984, Report ref. CPT-84/PE-1695.
- [48] Souriau, J.M. On Geometric Mechanics. Discret. Cont. Dyn. Syst. J. 2007, 19, 595–607.
- [49] Vallée, C. Relativistic thermodynamics of continua. Int. J. Eng. Sci. 1981, 19, 589–601.
- [50] Vallée, C.; Lerintiu, C., Convex analysis and entropy calculation in statistical mechanics, Proceedings of A. Razmadze mathematical Institute 137, 2005, pp. 111-129.
- [51] Vey, J. Sur les Automorphismes Affines des Ouverts Convexes Saillants. Annali della Scuola Normale Superiore di Pisa, Classe di Science 1970, 24, 641–665.
- [52] Yoshizawa, S.; Tanabe, K. Dual differential geometry associated with the Kullback-Leibler information on the Gaussian distributions and its 2-parameters deformations. SUT Journal of Mathematics, vol.35, n°1, 1999, 113-137

Finesse et géométrie dans le style raffiné de Joseph de Maistre et de Erich Kähler



Pour la théorie de la connaissance mais aussi pour les sciences est fondamentale la notion de perspective.

Or, les expériences faites dans la géométrie algébriques, dans la théorie des nombres, et dans l'algèbre abstraite m'induisent à tenter une formulation mathématique de cette notion pour **surmonter ainsi au moyen de raisonnements d'origine géométrique la géométrie**. Il me semble en effet, que la tendance vers l'abstraction observée dans les mathématiques d'aujourd'hui, loin d'être l'ennemi de l'intuition ait le sens profond de **quitter l'intuition pour la faire renaître dans une alliance entre « esprit de géométrie » et « esprit de finesse »**, alliance rendue possible par les réserves énormes des mathématiques pures dont Pascal et Goethe ne pouvaient pas encore se douter.

Erich Kähler – Sur la théorie des corps purement algébriques, 1952

Si on ajoute que la critique qui accoutume l'esprit, surtout en matière de faits, à recevoir de simples **probabilités** pour des preuves, est, par cet endroit, moins propre à le former, que ne le doit être la **géométrie** qui lui fait contracter l'habitude de n'acquiescer qu'à l'évidence; nous répliquerons qu'à la rigueur on pourrait conclure de cette différence même, que la critique donne, au contraire, plus d'exercice à l'esprit que la **géométrie**: parce que l'évidence, qui est une et absolue, le fixe au premier aspect sans lui laisser ni la liberté de douter, ni le mérite de choisir; au lieu que les **probabilités** étant susceptibles du plus et du moins, il faut, pour se mettre en état de prendre un parti, les comparer ensemble, les discuter et les peser. Un genre d'étude qui rompt, pour ainsi dire, l'esprit à cette opération, est certainement d'un usage plus étendu que celui où tout est soumis à l'évidence; parce que les occasions de se déterminer sur des vraisemblances ou **probabilités**, sont plus fréquentes que celles qui exigent qu'on procède par démonstrations: pourquoi ne dirions – nous pas que souvent elles tiennent aussi à des objets beaucoup plus importants ?

Joseph de Maistre

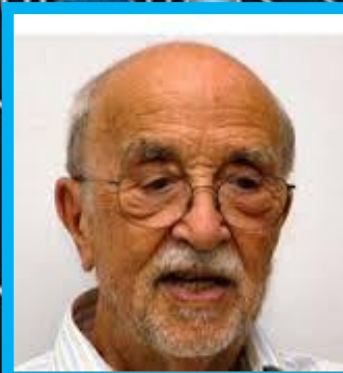
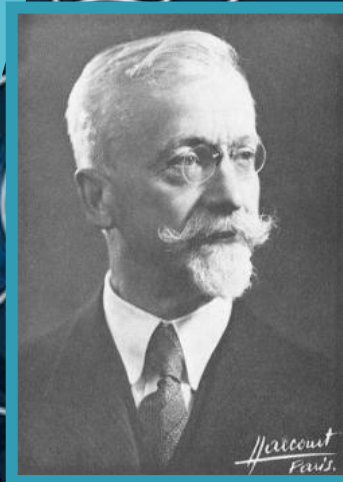
« La Physique mathématique, en incorporant à sa base la notion de groupe, marque la suprématie rationnelle...Chaque géométrie – et sans doute plus généralement chaque organisation mathématique de l'expérience – est caractérisée par un groupe spécial de transformations.... Le groupe apporte la preuve d'une mathématique fermée sur elle-même. Sa découverte clôt l'ère des conventions, plus ou moins indépendantes, plus ou moins cohérentes » -
Gaston Bachelard, Le nouvel esprit scientifique, 1934

Les sources thermodynamiques de la Géométrie de l'Information: de la Philosophie naturelle d'Aristote et l'épistémologie de Blaise Pascal à la métrique de Koszul-Souriau

Frédéric BARBARESCO

Tout mathématicien sait qu'il est impossible de comprendre un cours élémentaire en thermodynamique.

Vladimir Arnold

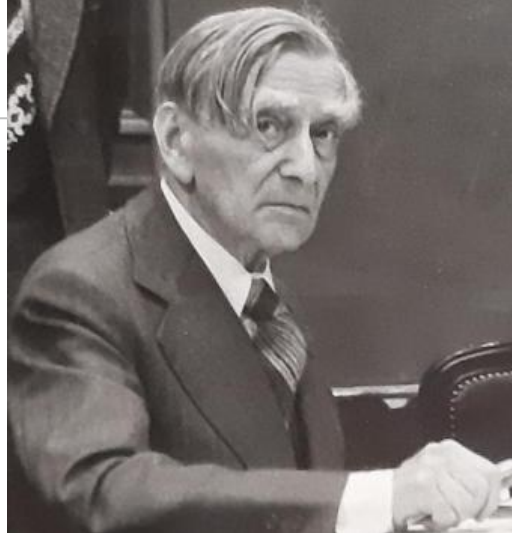


Vladimir Jankélévitch & la Physique d'Aristote

VLADIMIR JANKÉLÉVITCH. NOTES POUR UN COURS SUR LA *PHYSIQUE* D'ARISTOTE (1976) ET NOTES ANCIENNES D'ÉTUDIANT (S.D.) AYANT SERVI POUR CE COURS

Manuscrits autographes

BNF, Manuscrits, NAF 28313



Aristote, "Physique"

De 1 à 4 : la Nature
- 5 à 8 : le Mouvement ("Περί κινήσεως")
l'authenticité de 7 a été contestée : simples notes ? notes d'auditeur ?
8 : le Premier Moteur. 1 : opinions des philosophes antérieurs.
La Nature = principe interne du mouvement et du repos. Mouvement local = variation de
changement : changement de lieu (cf. altération, génération, corruption, accroissement - décroissement
déplacement). Cause locale du mouvement. Les êtres mobiles font partie d'un monde fermé.
Mouvement de la sphère des fixes lui-même doit avoir cause transcendente : 1^{er} moteur → 1^{er} mobile → 1^{er}
cf. Hamelin "Le Système d'Aristote".
voir "Métaphysique" I : où il y a question de Dieu, au f^o de vue théologique, non physique
cf. Commentaires de Simplicius sur la Physique (6^{es} ap. J.-C.) - et de Philopon (Jean
d'Alexandrie, chrétien) - de Themistios (rhéteur de 4^{es} ap. J.-C., auteur d'un paraphrase
Simplicius cite sous Alexandre d'Aphrodise (début du 3^e s. ap. J.-C.) - Ed. Viels.
Brandis in Bekker

HENRI BERGSON,
de l'Académie Française
et de l'Académie des Sciences morales et politiques,
regrette de ne pouvoir fournir à Monsieur
V. Jankélévitch les indications qu'il
BnF
MSS
32, rue Viala.

Bibliothèque François-Mitterrand – Exposition Vladimir Jankelevitch - 2019

Blaise Pascal: ALEA GEOMETRIA / Geometry of Chance

In 1654, Blaise Pascal submitted a paper to « *Celeberrimae matheseos Academiae Parisiensi* » entitled « *ALEAE GEOMETRIA : De compositione aleae in ludis ipsi subjectis* »

- « ... et sic matheseos demonstrationes cum aleae incertitudine jugendo, et quae contraria videntur conciliando, ab utraque nominationem suam accipiens, stupendum hunc titulum jure sibi arrogat: **Aleae Geometria** »
- « ... par l'union ainsi réalisée entre les démonstrations des mathématiques et l'incertitude du hasard, et par la conciliation entre les contraires apparents, elle peut tirer son nom de part et d'autre et s'arroger à bon droit ce titre étonnant: **Géométrie du Hasard** »
- « ... by the union thus achieved between the demonstrations of mathematics and the uncertainty of chance, and by the conciliation between apparent opposites, it can take its name from both sides and arrogate to right this amazing title: **Geometry of Chance** »



(¹) « Novissima autem ac penitus intractatae materiae tractatio, « scilicet de compositione aleae in ludis ipsi subjectis (quod « gallico nostro idiomate dicitur *faire les partis des jeux*): ubi « anceps fortuna aequitate rationis ita reprimatur ut utriusque « lusorum quod jure competit exacte semper assignetur. Quod « quidem eo fortius ratiocinando quaerendum, quo minus tentando « investigari possit: ambigui enim sortis eventus fortuitae contingentiae potius quam naturali necessitati merito tribuuntur. Ideo « res hactenus erravit incerta; nunc autem quae *experimento* « *rebellis fuerat*, rationis dominium effugere non potuit: eam « quippe tanta securitate in artem per geometriam reduximus, « ut, certitudinis ejus particeps facta, jam audacter prodeat; et « sic, matheseos demonstrationes cum aleae incertitudine jungendo, « et quae contraria videntur conciliando, ab utraque nominationem « suam accipiens, stupendum hunc titulum jure sibi arrogat: « *aleae geometria*. » (*Œuvres de Pascal*, t. IV, p. 358 de l'édition de 1819.)

Aristotle's natural philosophy and Pascal's epistemology

- ***“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”*** - Pierre Duhem, Sur les équations générales de la Thermodynamique, 1891
- ***“Nous prenons le mot mouvement pour désigner non seulement un changement de position dans l'espace, mais encore un changement d'état quelconque, lors même qu'il ne serait accompagné d'aucun déplacement . . . De la sorte, le mot mouvement s'oppose non pas au mot repos, mais au mot équilibre.”*** - Pierre Duhem, Commentaire aux principes de la Thermodynamique, 1894
- ***“This theoretical design led Duhem to rediscover and reinterpret the tradition of Aristotle's natural philosophy and Pascal's epistemology . . . This outcome was surprising and clearly echoed the Aristotelian language and concept of motion as change and transformation: within the framework of Aristotelian natural philosophy, motion in the modern physical sense was actually a special case of the general concept of motion... Pascal had criticised Descartes's mechanism and had stressed the need for a theoretical practice, a synthetic and intuitive one, besides the geometric formalisation”*** – S.

Bordoni, From thermodynamics to philosophical tradition: Pierre Duhem's research between 1891 and 1896. *Lettera Matematica* 2017, 5, 261–266.





At fourteen



At seventeen



At nineteen

Duhem, P. (1894), *Commentaire aux principes de la Thermodynamique – Troisième partie, Journal de Mathématiques pures et appliquées, 4e série, tome 10, 207-285.*

- “Il nous semble qu’une conclusion générale se dégage de cette étude: si la science des mouvements cesse d’être, dans l’ordre logique, la première des Sciences physiques, pour devenir seulement un cas particulier **d’une science plus générale embrassant dans ses formules toutes les modifications des corps**, la tentation sera moindre, pensons-nous, de ramener l’étude de tous les phénomènes physiques à l’étude du mouvement; on comprendra mieux que **le changement de lieu dans l’espace n’est pas une modification plus simple que le changement de température ou de quelque autre qualité physique**; on fuira dès lors plus volontiers ce qui a été jusqu’ici le plus dangereux écueil de la Physique théorique, la recherche d’une explication mécanique de l’Univers.”

Pierre Duhem: La théorie Physique: son objet, sa structure

Pierre Duhem

- « Il est une Cosmologie avec laquelle la Thermodynamique générale présente une analogie non-méconnaissable; cette Cosmologie, **c'est la Physique péripatéticienne** ... Parmi les attributs de la substance, la Physique péripatéticienne confère une **égale importance à la catégorie de la quantité et à la catégorie de la qualité**; or, par ses symboles numériques, la Thermodynamique générale représente également les diverses grandeurs des quantités et les diverses intensités des qualités. **Le mouvement local n'est, pour Aristote, qu'une des formes du mouvement général**, tandis que les Cosmologies cartésienne, atomistique et newtonienne concordent en ceci que le seul mouvement possible est le changement de lieu dans l'espace. Et voici que la Thermodynamique générale traite, en ses formules, d'une **foule de modifications telles que les variations de températures, les changements d'état électrique ou d'aimantation, sans chercher le moins du monde à réduire ces variations au mouvement local** »



Motion & Aristotle's natural philosophy

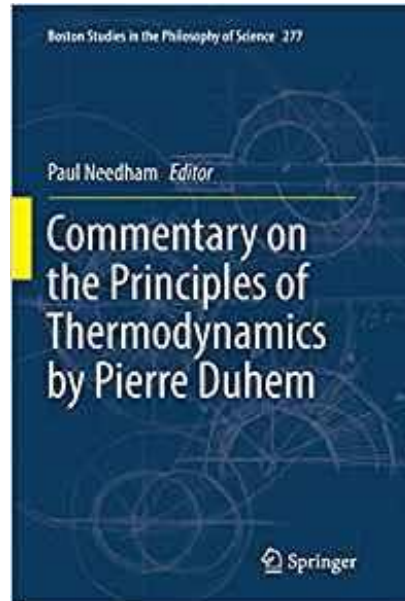
From Theoretical Physics to Meta-theoretical Commitments, S. Bordoni

- **Duhem transformed “dynamics into a specific instance of thermodynamics”**; in other words, “under the name of thermodynamics” he had put forward a more general theory that encompassed within common principles **“every transformation of a body.” Both changes of place and changes of physical qualities found room in that generalized mechanics.** The traditional science of motion became a specific instance of a more general science: it had to be understood that **the change of position in space was not “a simpler modification than the change of temperature or any other physical quality.”** This generalization could bypass “the most dangerous stumbling block on the path of theoretical physics,” namely “the search for a mechanical explanation of the universe.” From the mathematical point of view, his design corresponded to a reduction of physics to the language of analytical mechanics, but from the theoretical point of view it was an anti-reductionist design that involved a generalisation of that language. **In Duhem’s “more general science” we find the coexistence of a mechanical approach, in the sense of a generalization of Lagrange’s mathematical physics, and the rejection of mechanical models and mechanical explications in the sense of the traditional mechanistic worldview** [Duhem 1894a, pp. 284-5].

ÉQUATIONS GÉNÉRALES DE LA THERMODYNAMIQUE,

PAR P. DUHEM,

CHARGÉ D'UN COURS COMPLÉMENTAIRE A LA FACULTÉ DES SCIENCES DE LILLE.



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**

Pierre Duhem (1861-1916): Les équations générales de la thermodynamique

Publications de Pierre Duhem:

- Duhem, P. **Sur les équations générales de la thermodynamique**. In Annales scientifiques de l'École Normale Supérieure; Volume 8, pp. 231–266.
- Duhem, P. **Commentaire aux principes de la Thermodynamique—Première partie**. J. Math. Appl. 1892, 8, 269–330.
- Duhem, P. **Commentaire aux principes de la Thermodynamique—Troisième partie**. J. Math. Appl. 1894, 10, 207–286.
- Duhem, P. **Les théories de la chaleur**. Revue des deux Mondes 1895, 130, 851–868.



<http://www.duhem2016.info/>



Pierre Duhem (1861-1916) et ses contemporains

Institut Henri Poincaré, 14 Septembre 2016

Amphithéâtre Hermite

organisée par Hervé Le Ferrand (Dijon) - Laurent Mazliak (Paris)



Pierre Duhem: General Equations of Thermodynamics

ÉQUATIONS GÉNÉRALES DE LA THERMODYNAMIQUE,

PAR P. DUHEM,

CHARGÉ D'UN COURS COMPLÉMENTAIRE A LA FACULTÉ DES SCIENCES DE LILLE.

Au premier rang, il convient de citer M. F. Massieu ⁽²⁾; il a obtenu un résultat capital, à savoir que toutes les équations de la Thermodynamique peuvent être écrites au moyen d'une seule *fonction caractéristique* et de ses dérivées partielles, cette fonction changeant d'ailleurs avec les variables indépendantes adoptées.

M. Gibbs ⁽³⁾, dans le Travail célèbre où il a démontré que les fonctions caractéristiques de M. Massieu pouvaient jouer le rôle de potentiels dans la détermination des états d'équilibre du système, a fourni également de profondes idées sur les équations de la Thermodynamique prises sous la forme la plus générale.

M. H. von Helmholtz ⁽¹⁾ a développé de son côté des idées analogues.

Enfin, M. Arthur von Oettingen ⁽²⁾ a donné un exposé de la Thermodynamique d'une remarquable généralité; il a cherché, dans cet exposé, à mettre nettement en évidence le caractère dualistique que présente le développement de la Thermodynamique, caractère déjà marqué par M. Massieu.

(1) R. CLAUSIUS, *Sur diverses formes des équations fondamentales de la Thermodynamique, qui sont commodes dans l'application (Théorie mécanique de la chaleur. Trad. Folio, Mémoire IX).*

(2) F. MASSIEU, *Sur les fonctions caractéristiques (Comptes rendus, t. LXIX, p. 858 et 1077; 1869).* — *Mémoire sur les fonctions caractéristiques des divers fluides et sur la théorie des vapeurs (Savants étrangers, t. XXII; 1876).*

(3) J. WILLARD GIBBS, *On the equilibrium of heterogeneous substances (Transactions of the Connecticut Academy, t. III; 1875-1876).*

(1) H. VON HELMHOLTZ, *Zur Thermodynamik chemischer Vorgänge (Sitzungsber. der Berl. Akademie, t. I, p. 23; 1882).*

(2) ARTHUR VON OETTINGEN, *Die thermodynamischen Beziehungen, autathetisch entwickelt (Mémoires de l'Académie de Saint-Petersbourg, t. XXIII; 1885).*

Duhem developed a general theory and **extend the concept of « calorific capacity »**

(Souriau will give a geometrical status to this calorific capacity).
Duhem and Souriau theories are compliant

Pierre Duhem: General Equations of Thermodynamics

➤ P. DUHEM. Commentaire aux principes de la Thermodynamique (Troisième Partie).
Journal de mathématiques pures et appliquées 4e série, tome 10 (1894), p. 207-286.

Le potentiel thermodynamique interne. — Soient

$$U(\alpha, \beta, \dots, \lambda, \vartheta) \quad \text{et} \quad S(\alpha, \beta, \dots, \lambda, \vartheta)$$

l'énergie interne et l'entropie d'un système.

Posons

$$(1) \quad \mathcal{F}(\alpha, \beta, \dots, \lambda, \vartheta) = E[U(\alpha, \beta, \dots, \lambda, \vartheta) - F(\vartheta)S(\alpha, \beta, \dots, \lambda, \vartheta)].$$

SECONDE RESTRICTION. — Imaginons qu'à partir d'un certain état $(\alpha, \beta, \dots, \lambda, \vartheta)$ du système, on lui impose une modification virtuelle

$$\delta\alpha, \delta\beta, \dots, \delta\lambda, \delta\vartheta.$$

Les actions des corps extérieurs qui le maintiendraient en équilibre en l'état $(\alpha, \beta, \dots, \lambda, \vartheta)$ effectuent un travail virtuel

$$A\delta\alpha + B\delta\beta + \dots + L\delta\lambda + \Theta\delta\vartheta.$$

$$\left\{ \begin{array}{l} A = f_\alpha(\alpha, \beta, \dots, \lambda, \vartheta), \\ B = f_\beta(\alpha, \beta, \dots, \lambda, \vartheta), \\ \dots \\ L = f_\lambda(\alpha, \beta, \dots, \lambda, \vartheta), \\ \Theta = f_\vartheta(\alpha, \beta, \dots, \lambda, \vartheta). \end{array} \right. \quad \rightarrow \quad \left\{ \begin{array}{l} R_\alpha = \frac{\partial U}{\partial \alpha} - \frac{f_\alpha}{E}, \\ R_\beta = \frac{\partial U}{\partial \beta} - \frac{f_\beta}{E}, \\ \dots \\ R_\lambda = \frac{\partial U}{\partial \lambda} - \frac{f_\lambda}{E}, \\ C = \frac{\partial U}{\partial \vartheta} - \frac{f_\vartheta}{E}. \end{array} \right.$$

$$\left\{ \begin{array}{l} R_\alpha = \frac{\partial U}{\partial \alpha} - \frac{A}{E}, \\ R_\beta = \frac{\partial U}{\partial \beta} - \frac{B}{E}, \\ \dots \\ R_\lambda = \frac{\partial U}{\partial \lambda} - \frac{L}{E}, \\ C = \frac{\partial U}{\partial \vartheta} - \frac{\Theta}{E}. \end{array} \right. \quad \rightarrow \quad \left\{ \begin{array}{l} R_\alpha = F(\vartheta) \frac{\partial S}{\partial \alpha}, \\ R_\beta = F(\vartheta) \frac{\partial S}{\partial \beta}, \\ \dots \\ R_\lambda = F(\vartheta) \frac{\partial S}{\partial \lambda}, \\ C = F(\vartheta) \frac{\partial S}{\partial \vartheta}. \end{array} \right. \quad \rightarrow \quad \left\{ \begin{array}{l} A = \frac{\partial \mathcal{F}}{\partial \alpha}, \\ B = \frac{\partial \mathcal{F}}{\partial \beta}, \\ \dots \\ L = \frac{\partial \mathcal{F}}{\partial \lambda}, \\ \Theta = \frac{\partial \mathcal{F}}{\partial \vartheta} + ES \frac{dF(\vartheta)}{d\vartheta}. \end{array} \right.$$

$$ES = \frac{1}{\frac{dF(\vartheta)}{d\vartheta}} \left(\Theta - \frac{\partial \mathcal{F}}{\partial \vartheta} \right) \quad \quad EU = \mathcal{F} + \frac{F(\vartheta)}{\frac{dF(\vartheta)}{d\vartheta}} \left(\Theta - \frac{\partial \mathcal{F}}{\partial \vartheta} \right)$$

$$\left\{ \begin{array}{l} R_\alpha = \frac{F(\vartheta)}{\frac{dF(\vartheta)}{d\vartheta}} \left(\frac{\partial \Theta}{\partial \alpha} - \frac{\partial^2 \mathcal{F}}{\partial \vartheta \partial \alpha} \right), \\ R_\beta = \frac{F(\vartheta)}{\frac{dF(\vartheta)}{d\vartheta}} \left(\frac{\partial \Theta}{\partial \beta} - \frac{\partial^2 \mathcal{F}}{\partial \vartheta \partial \beta} \right), \\ \dots \\ R_\lambda = \frac{F(\vartheta)}{\frac{dF(\vartheta)}{d\vartheta}} \left(\frac{\partial \Theta}{\partial \lambda} - \frac{\partial^2 \mathcal{F}}{\partial \vartheta \partial \lambda} \right), \\ C = \frac{F(\vartheta)}{\frac{dF(\vartheta)}{d\vartheta}} \left(\frac{\partial \Theta}{\partial \vartheta} - \frac{\partial^2 \mathcal{F}}{\partial \vartheta^2} \right) - \frac{F(\vartheta) F''(\vartheta)}{[F'(\vartheta)]^2} \left(\Theta - \frac{\partial \mathcal{F}}{\partial \vartheta} \right) \end{array} \right.$$

1° L'expression du potentiel thermodynamique interne du système;

2° L'expression $\Theta = f_\vartheta(\alpha, \beta, \dots, \lambda, \vartheta)$

de la quantité Θ relative au système en équilibre, On peut déterminer :

1° L'énergie interne et l'entropie du système dans un état quelconque;

2° Les conditions nécessaires et suffisantes de l'équilibre du système;

3° Les coefficients calorifiques du système en équilibre.

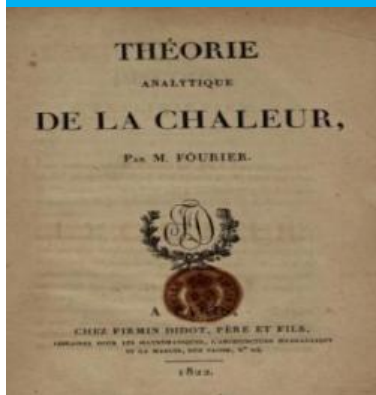
C'est la généralisation d'une proposition bien connue de M. Massieu.

Référence à Massieu

This document may not be reproduced, modified, published, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales. - © Thales 2015 All rights reserved.

From Analytical/Mechanical and Mathematical to Geometrical Theory of Heat

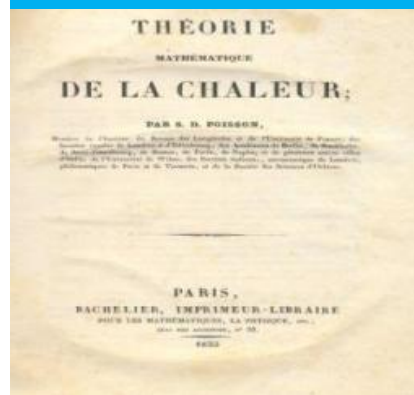
Fourier: Analytical Theory of Heat



Clausius: Mechanical Theory of Heat

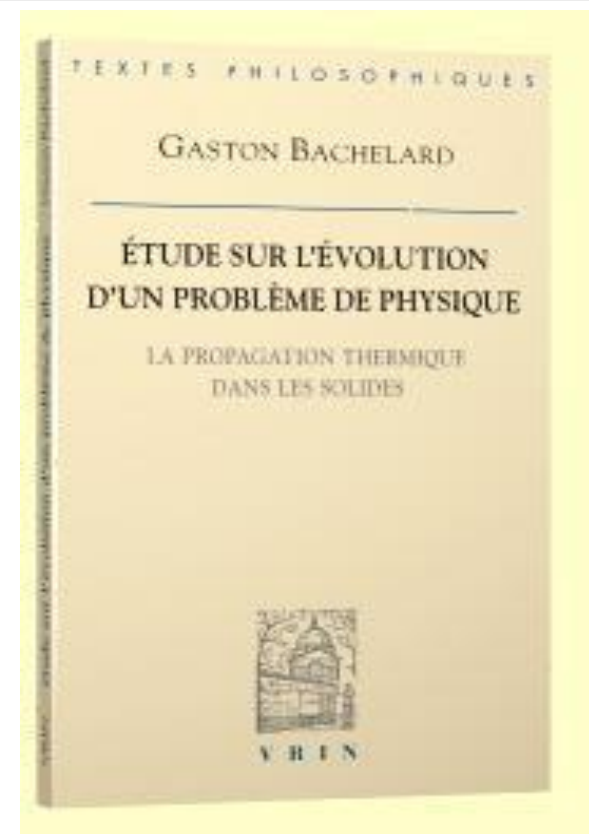


Poisson: Mathematical Theory of Heat



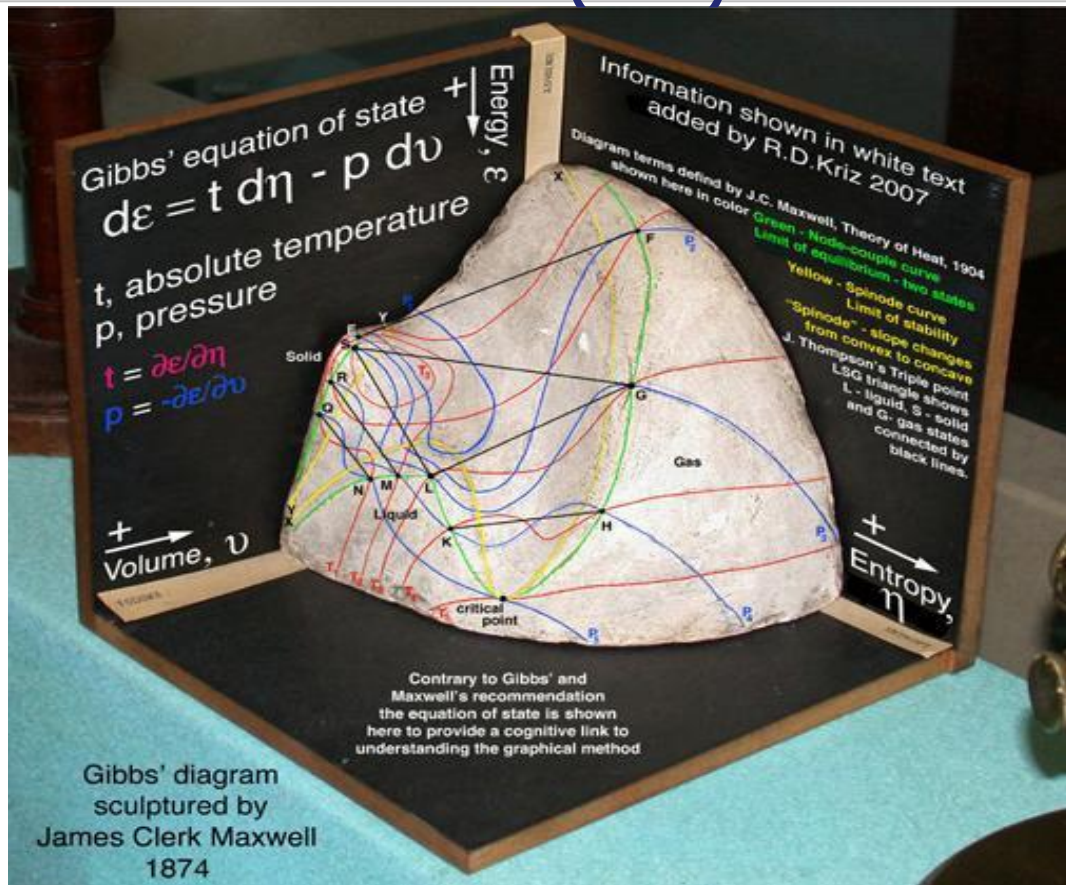
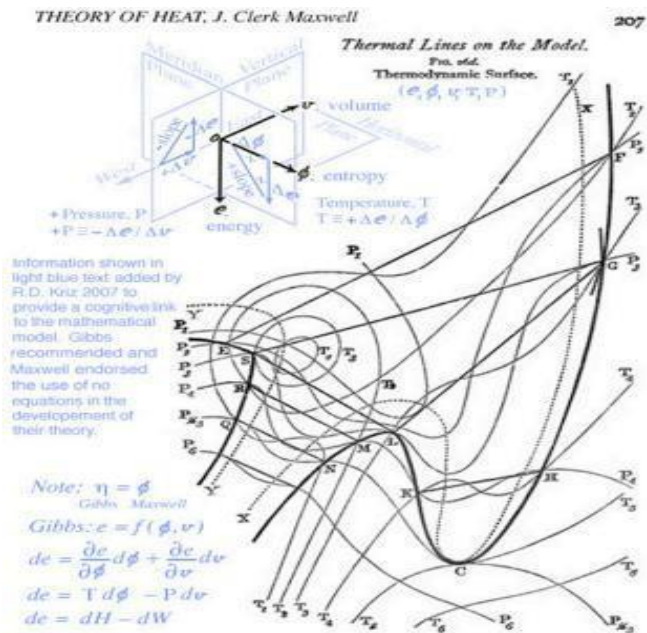
Gaston Bachelard Book

- Étude sur l'évolution d'un problème de physique: La propagation thermique dans les solides; Vrin - Bibliothèque des Textes Philosophiques, 1973
- Evolution of ideas : Biot, Fourier, Poisson, Lamé, Boussinesq
- http://www.vrin.fr/book.php?code=9782711600434&search_bac_k=bachelard&editor_back=10



Geometric Theory of Heat: Gibbs Diagrams sculpted by James Clerk Maxwell (1874)

This document may not be reproduced, modified, adapted, published, translated, in any way, in whole or in part or disclosed to a third party without the prior written consent of Thales. - © Thales 2015 All rights reserved.



Poincaré unsuccessful to build a Mechanical Theory of Heat: « bon sens » paysan d'Henri Poincaré ou la fable du grain d'avoine et du tas de blé

333. Toutes les tentatives de cette nature doivent donc être abandonnées; les seules qui aient quelque chance de succès sont celles qui sont fondées sur l'intervention des lois statistiques comme, par exemple, la théorie cinétique des gaz.

Ce point de vue, que je ne puis développer ici, peut se résumer d'une façon un peu vulgaire comme il suit :

Supposons que nous voulions placer un grain d'avoine au milieu d'un tas de blé; cela sera facile; supposons que nous voulions ensuite l'y retrouver et l'en retirer; nous ne pourrons y parvenir. Tous les phénomènes irréversibles, d'après certains physiciens, seraient construits sur ce modèle.

FIN.

**Conclusion du cours de
Thermodynamique d'Henri Poincaré**

Supposons que nous voulions placer un grain d'avoine au milieu d'un tas de blé; cela sera facile; supposons que nous voulions ensuite l'y retrouver et l'en retirer; nous ne pourrons y parvenir.

