THALES

« Data Science sur un plateau » Seminar
22/01/2018 ENSAE – Paris-Saclay Campus

Elementary Structures of Information Geometry: Koszul-Souriau Characteristic Function, Fisher-Souriau Metric & Higher Order Maximum Entropy Density

Frédéric BARBARESCO
Representative of Key Technology Domain PCC (Processing, Control & Cognition) Board
THALES LAND & AIR SYSTEMS

THALES

Geometric Science of Information: Interdisciplinary community on geometric concepts that are universal in Mathematics, Physics and Information Theory

OPEN

Geometric Science of Information: GSI

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: Ecole des Mines de Paris
150 attendees from 38 countries

Videos:
https://www.youtube.com/channel/UCnE9-LbfFRqtaes49cN2DVg/videos
Website: www.gsi2017.org

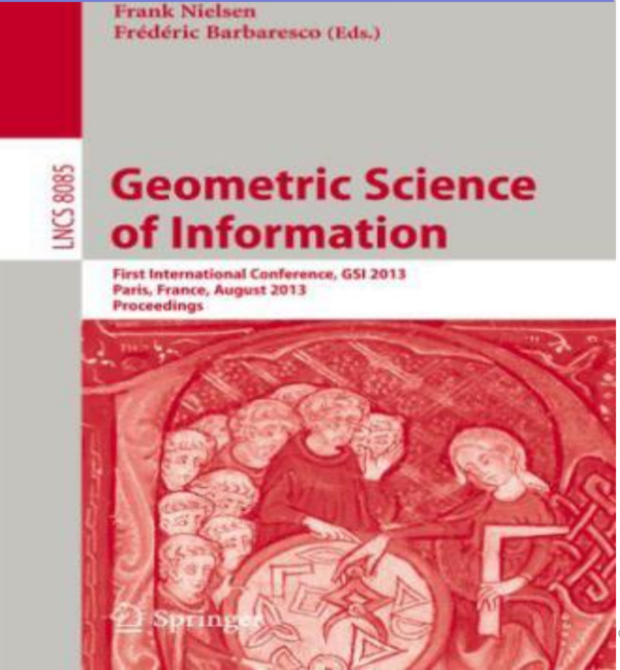
Société Mathématique de France

SMF

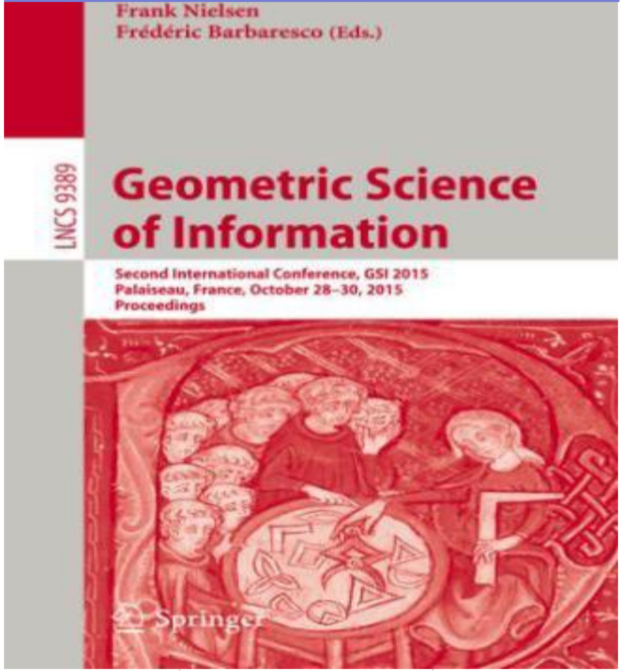
THALES

GSI SPRINGER PROCEEDINGS
LECTURE NOTES IN COMPUTER SCIENCE

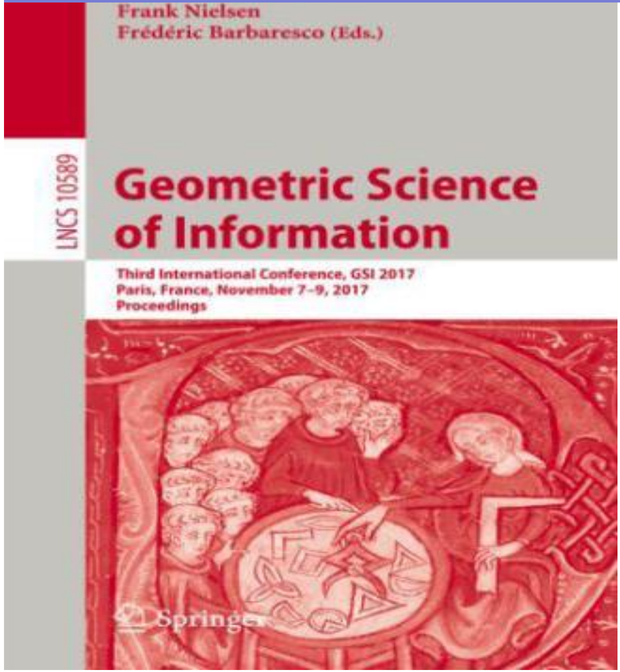
GSI’13 Springer Proceedings: http://www.springer.com/us/book/9783642400193



GSI’15 Springer Proceedings: http://www.springer.com/la/book/9783319250397

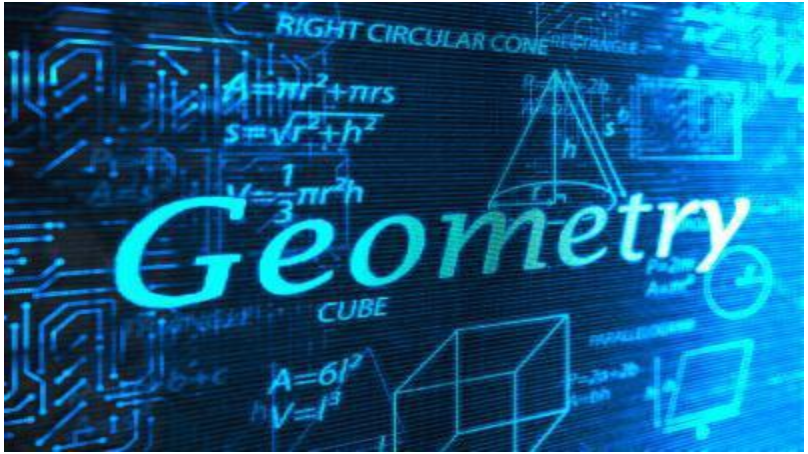


GSI’17 Springer Proceedings: http://www.springer.com/cn/book/9783319684444



GSI'17: Ecole des Mines de Paris
19 sessions

- Computational Information Geometry
- Geometrical Structures of Thermodynamics
- Geometry of Tensor-Valued Data
- Probability on Riemannian Manifolds
- Information Structure in Neuroscience
- Geometric Mechanics & Robotics
- Optimization on Manifold
- Geometric Robotics & Tracking
- Probability Density Estimation
- Applications of Distance Geometry
- Statistics on non-linear data
- Shape Space
- Divergence Geometry
- Geodesic Methods with Constraints
- Optimal Transport & Applications
- Monotone Embedding in Information Geometry
- Non-parametric Information Geometry
- Optimal Transport & Applications
- Statistical Manifold & Hessian Information Geometry

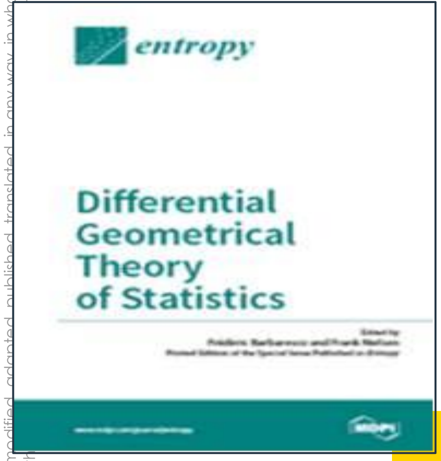


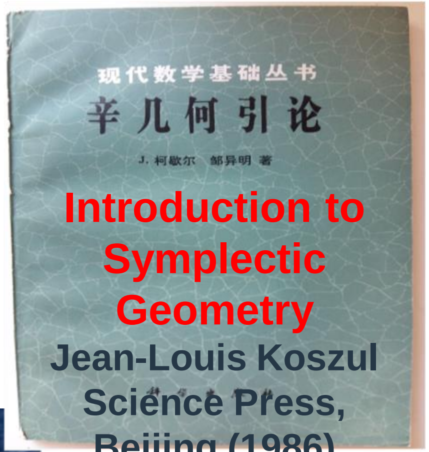
THALES

Last Publications on Geometric Science of Information

See More on <u>UNESCO UNITWIN website GSI « Geometric Science of Information »</u>

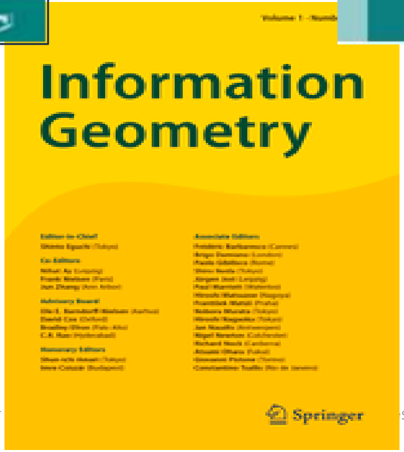
http://forum.cs-dc.org/topic/369/geometric-science-of-information-presentation-organisation-subscription

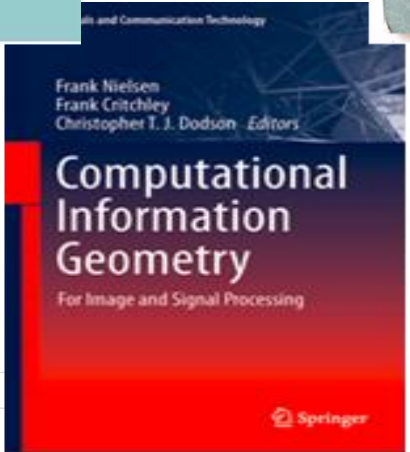




Introduction to Symplectic Geometry

Jean-Louis Koszul
Science Press, Beijing (1986)
(in Chinese)
(with reference to Souriau work)
Translation by SPRINGER in 2018





THALES

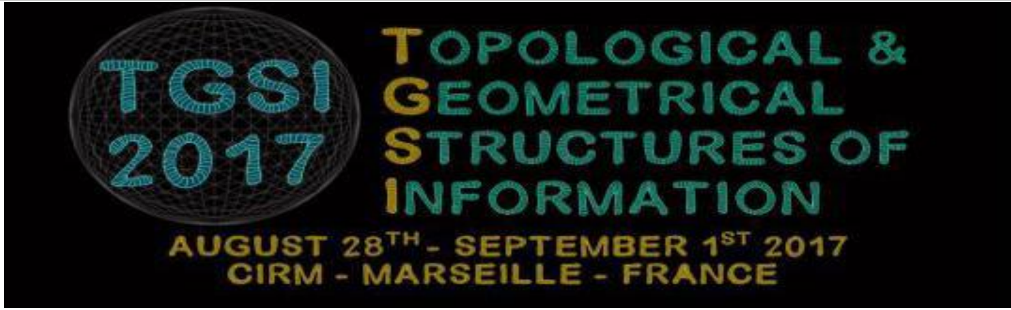
CIRM Seminar, August 2017
TGSI’17 « Topological & Geometrical Structures of Information »

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

entropy
an open access journal by MDPI



Talk on Koszul-Souriau Characteristic Function:
https://www.youtube.com/watch?v=VXxiMCn-tsE&feature=youtu.be

Max Planck Institut für Mathematik
9th GSO-2018 – Information Geometry & Statistical Physics

Max-Planck-Institut für **Mathematik** in den **Naturwissenschaften**

MAX-PLANCK-GESELLSCHAFT

Home | The Institute | Research | Career | IMPRS | Calendar | People | Publications | Library | Services | Internal

Conferences | Seminars | Colloquiums | For students | For grammar pupils | Leipzig University | Other lectures

2018 2017 2016 2015 2014 2013 2012 2011 2010 2009 2008 2007 2006 2005 2004 2003 2002 2001 2000 1999 1998 1997 199

The Ninth International Conference on Guided Self-Organisation (GSO-2018) : Information Geometry and Statistical Physics

https://www.mis.mpg.de/calendar/conferences/2018/gso18.html

March 26 - 28, 2018
Max Planck Institute for Mathematics in the Sciences

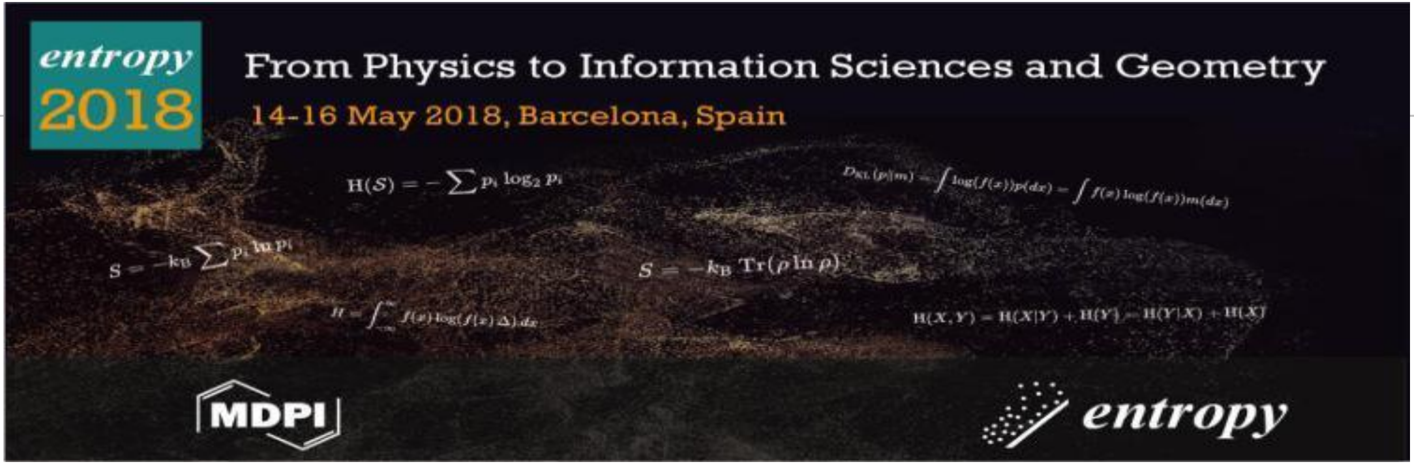
The goal of Guided Self-Organization (GSO) is to leverage the strengths of self-organization (i.e., its simplicity, parallelization, adaptability, robustness, scalability) while still being able to direct the outcome of the self-organizing process. GSO typically has the following features:

(i) An increase in organization (i.e., structure and/or functionality) over time;

(ii) Local interactions that are not explicitly guided by any external agent;

(iii) Task-independent objectives that are combined with task-dependent constraints.

GSO-2018 is the 9th conference in a bi-annual series on GSO. Recent research is starting to indicate that **information geometry**, **nonequilibrium statistical physics** in general, and the **thermodynamics of computation** in particular, all play a key role in GSO. Accordingly, a particular focus of this conference will be the interplay of those three topics as revealed by their relationship with GSO.



14–16 May 2018

From Physics to Information Sciences and Geometry

Barcelona, Spain

https://sciforum.net/conference/Entropy2018-1

The main topics and sessions of the conference cover:

- **Physics: classical Thermodynamics and Quantum**
- **Statistical physics and Bayesian computation**
- **Geometrical science of information, topology and metrics**
- **Maximum entropy principle and inference**
- **Kullback and Bayes or information theory and Bayesian inference**
- **Entropy in action (applications)**

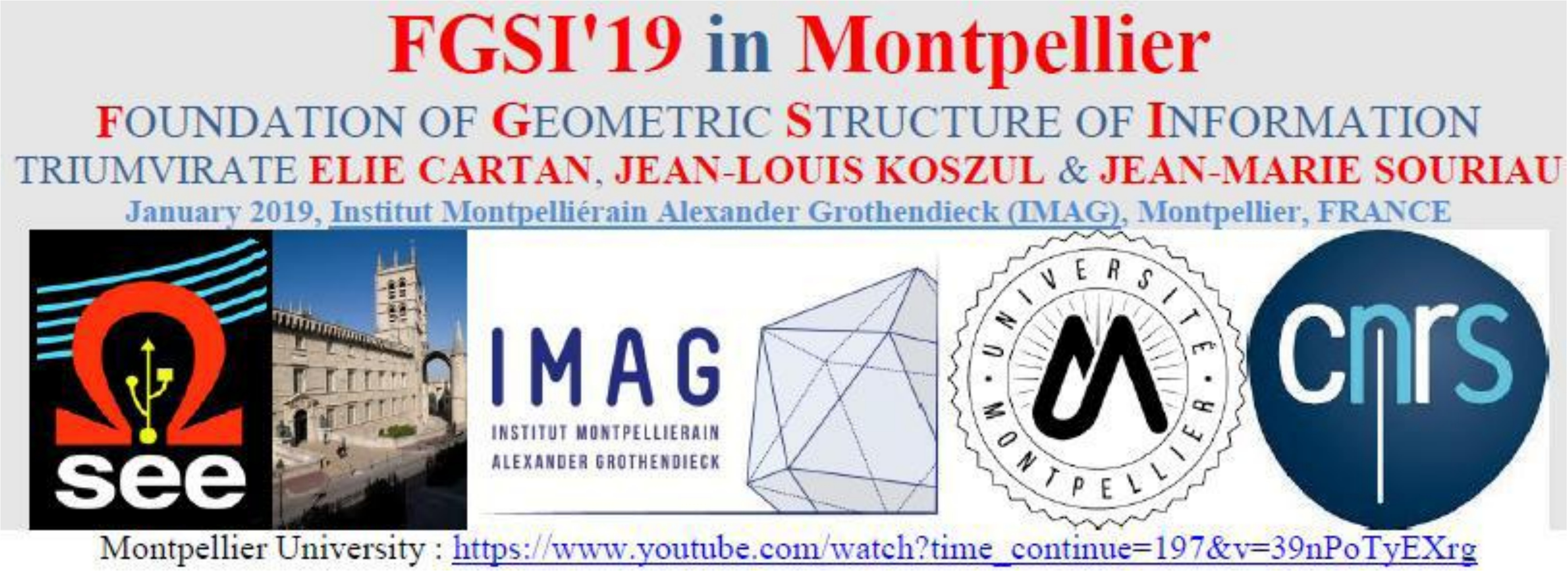
Abstract Submission: 23 Feb. 2018

The inter-disciplinary nature of contributions from both theoretical and applied perspectives are very welcome, including papers addressing conceptual and methodological developments, as well as new applications of entropy and information theory.

FGSI'19: Foundation of Geometric Structure of Information Dedicated to Koszul and Souriau Works

50th birthday of Souriau Book: « Structures des systèmes dynamiques »

SPRINGER translation of Koszul Lecture at Beijing « Introduction to Symplectic Geometry »



OPEN

THALES

Orbituary: Talk dedicated to Jean-Louis Koszul (1921-2018)

Jean-Louis Koszul passed away Friday January 12th 2018


Koszul, Shima, Boyom @ GSI'13


Shima talk on « Koszul Geometry of Hessian Structure » @ GSI'13


Shima, Koszul & Amari @ GSI'13

Jean-Louis Koszul
- Henri Cartan's PhD student
- Member of Bourbaki
- Strasbourg University and Institut Joseph Fourier (Grenoble)
- Creator of CIRM in Luminy


Koszul & Amari @ GSI'13

THALES

Motivations to use Information Geometry in Data Science:
Best in class Deep Learning Algorithms are based on IG Natural Gradient

OPEN

Information Geometry Motivation: Yann Ollivier's works
Natural Gradient & Natural Langevin Dynamics

GSI'17 Best paper Winners are:

Yann Ollivier (FACEBOOK IA Lab, Paris) &
Gaétan Marceau-Caron (MILA Lab, Montreal)

Natural Langevin Dynamics for Neural Networks

Gaétan Marceau-Caron[1](✉) and Yann Ollivier[2](✉)

[1] MILA, Université de Montréal, Montréal, Canada
gaetan.marceau.caron@umontreal.ca
[2] CNRS, Université Paris-Saclay, Paris, France
contact@yann-ollivier.org

Preconditioned Stochastic Gradient Langevin Dynamics

$$\theta \leftarrow \theta - \eta\, C\, \hat{\mathbb{E}}_{(x,y)\in\mathcal{D}}\, \partial_\theta \left(\ell_\theta(y|x) - \frac{1}{N} \ln \alpha(\theta) \right) + \sqrt{\frac{2\eta}{N}}\, C^{1/2} \mathcal{N}(0, \mathrm{Id})$$

Natural Langevin Dynamics (use of Fisher Matrix as in Natural Gradient from Information Geometry)

$$C \propto J(\theta^*)^{-1}$$

$$J(\theta) := \mathbb{E}_{(x,y)\in\mathcal{D}}\, \mathbb{E}_{\tilde{y}\sim p_\theta(\tilde{y}|x)} \left[(\partial_\theta \ln p_\theta(\tilde{y}|x)) (\partial_\theta \ln p_\theta(\tilde{y}|x))^\top \right]$$

EN

Information Geometry Natural Gradient & its Dual Entropic Gradient

Natural Gradient

$$\boxed{\theta^{k+1} = \theta^k - \eta\, M(\theta)^{-1} \frac{\partial L(\theta^k)}{\partial \theta}}$$

$$\theta^{k+1} = \theta^k - \eta\, (\text{Hess } L)^{-1} \frac{\partial L(\theta^k)}{\partial \theta}$$

(Natural gradient: $M(\theta) =$ Fisher information matrix.)

Dual Entropic Gradient (Mirror Descent, Balian Gradient)

$$\theta_{t+1} = \theta_t - \alpha_t \nabla f_t(\theta_t) \qquad \theta_{t+1} = \arg\min_{\theta \in \Theta} \left\{ \langle \theta, \nabla f_t(\theta_t) \rangle + \frac{1}{2\alpha_t} \|\theta - \theta_t\|_2^2 \right\}$$

$$\theta_{t+1} = \arg\min_{\theta \in \Theta} \left\{ \langle \theta, \nabla f_t(\theta_t) \rangle + \Psi(\theta, \theta_t) \right\} \qquad \Psi(\theta, \theta') = \frac{1}{2} \|\theta - \theta'\|_2^2$$

$$B_G(\theta, \theta') = G(\theta) - G(\theta') - \langle \nabla G(\theta'), \theta - \theta' \rangle \qquad \theta_{t+1} = \arg\min_{\theta} \left\{ \langle \theta, \nabla f_t(\theta_t) \rangle + \frac{1}{\alpha_t} B_G(\theta, \theta_t) \right\}$$

$$H(\eta) := \sup_{\theta \in \Theta} \{ \langle \theta, \eta \rangle - G(\theta) \}$$

$$h = \nabla H,\ g = h^{-1}$$

$$\eta = g(\theta)$$

$G(\theta)$	$H(\eta)$	$B_H(\eta, \eta')$
$\frac{1}{2}\|\theta\|_2^2$	$\frac{1}{2}\|\eta\|_2^2$	$\frac{1}{2}\|\eta - \eta'\|_2^2$
$\exp(\theta)$	$\langle \eta, \log \eta \rangle - \eta$	$\eta \log \frac{\eta}{\eta'}$
$\log(1 + \exp(\theta))$	$\eta \log \eta + (1 - \eta) \log(1 - \eta)$	$(1 - \eta) \log\left(\frac{1-\eta}{1-\eta'}\right) + \eta \log \frac{\eta}{\eta'}$

$$B_H(\eta, \eta') = H(\eta) - H(\eta') - \langle \nabla H(\eta'), \eta - \eta' \rangle$$

Dual Entropic Gradient

$$\theta_{t+1} = \arg\min_{\theta} \left\{ \langle \theta, \nabla f_t(\theta_t) \rangle + \frac{1}{\alpha_t} B_G(\theta, \theta_t) \right\}$$

$$\boxed{g(\theta_{t+1}) = g(\theta_t) - \alpha_t \nabla_\theta f_t(\theta_t)}$$

Natural Gradient

$$\eta = g(\theta) = \nabla G \qquad \theta = h(\eta) = \nabla H(\eta)$$

$$\eta_{t+1} = \eta_t - \alpha_t \nabla_\theta f_t(h(\eta_t))$$

$$\nabla_\eta f_t(h(\eta)) = \nabla_\eta h(\eta) \nabla_\theta f_t(h(\eta))$$

$$\nabla_\theta f_t(h(\eta_t)) = [\nabla_\eta h(\eta_t)]^{-1} \nabla_\eta f_t(h(\eta_t))$$

$$\boxed{\eta_{t+1} = \eta_t - \alpha_t [\nabla^2 H(\eta_t)]^{-1} \nabla_\eta f_t(h(\eta_t))}$$

Dual Entropic Gradient
H: Shannon Entropy

Yann Ollivier's TANGO for Best in Class Deep Learning Algorithm
TANGO (True Asymptotic Natural Gradient Optimization)

TANGO algorithm

Abstract

We introduce a simple algorithm, True Asymptotic Natural Gradient Optimization (TANGO), that converges to a true natural gradient descent in the limit of small learning rates, without explicit Fisher matrix estimation.

For quadratic models the algorithm is also an instance of averaged stochastic gradient, where the parameter is a moving average of a "fast", constant-rate gradient descent. TANGO appears as a particular de-linearization of averaged SGD, and is sometimes quite different on non-quadratic models. This further connects averaged SGD and natural gradient, both of which are arguably optimal asymptotically.

In large dimension, small learning rates will be required to approximate the natural gradient well. Still, this shows it is possible to get arbitrarily close to exact natural gradient descent with a lightweight algorithm.

Definition 1 (TANGO). *Let $\delta t_k \leqslant 1$ be a sequence of learning rates and let $\gamma > 0$. Set $v_0 = 0$. Iterate the following:*

- *Select a sample (x_k, y_k) at random in the dataset $\mathcal{D}$.*
- *Generate a pseudo-sample $\tilde{y}_k$ for input x_k according to the predictions of the current model, $\tilde{y}_k \sim p_\theta(\tilde{y}_k|x_k)$ (or just $\tilde{y}_k = y_k$ for the "outer product" variant). Compute gradients*

$$g_k \leftarrow \frac{\partial \ell(y_k|x_k)}{\partial \theta}, \qquad \tilde{g}_k \leftarrow \frac{\partial \ell(\tilde{y}_k|x_k)}{\partial \theta} \tag{2}$$

- *Update the velocity and parameter via*

$$v_k = (1 - \delta t_{k-1}) v_{k-1} + \gamma g_k - \gamma (1 - \delta t_{k-1}) (v_{k-1}^\top \tilde{g}_k) \tilde{g}_k \tag{3}$$

$$\theta_k = \theta_{k-1} - \delta t_k v_k \tag{4}$$

Proposition 2. *Assume that for each sample (x, y), the log-loss $\ell(y|x)$ is a quadratic function of θ whose Hessian does not depend on y (e.g., linear regression $\ell(y|x) = \frac{1}{2}\|y - \theta^\top x\|^2$).*

Then TANGO is identical to the following trajectory averaging algorithm:

$$\theta_k^{\text{fast}} = \theta_{k-1}^{\text{fast}} - \gamma \frac{\partial \ell(y_k|x_k)}{\partial \theta_{k-1}^{\text{fast}}} + \gamma \xi_k \tag{8}$$

$$\theta_k = (1 - \delta t_k)\theta_{k-1} + \delta t_k \theta_k^{\text{fast}} \tag{9}$$

where ξ_k is some centered random variable whose law depends on $\theta_{k-1}^{\text{fast}}$ and θ_{k-1}. The identification with TANGO is via $v_k = \theta_{k-1} - \theta_k^{\text{fast}}$.

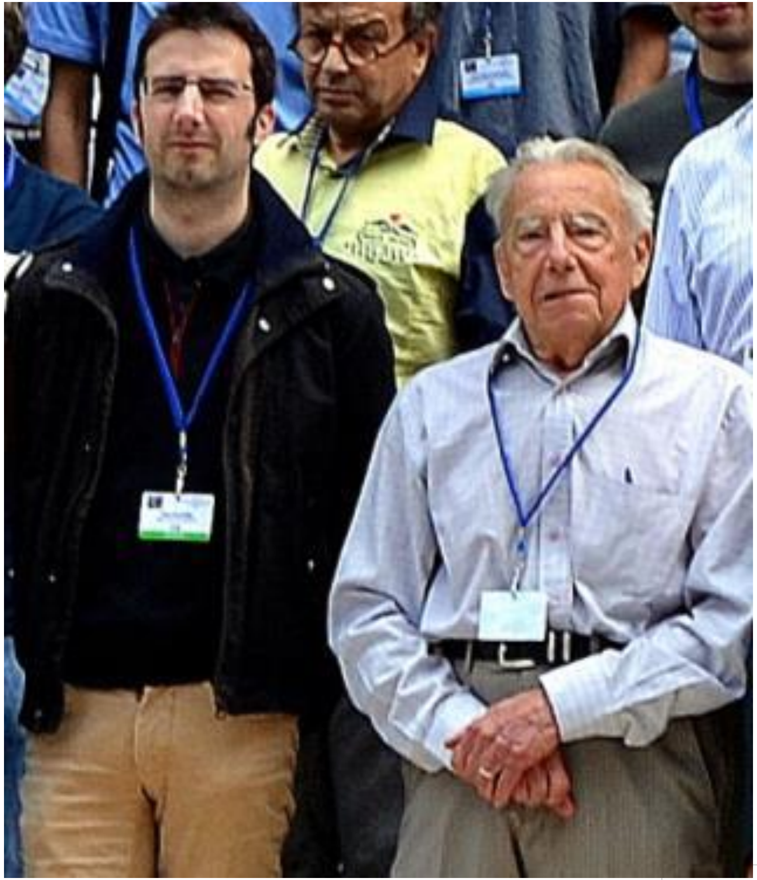
$$\ell(y|x) := -\ln p_\theta(y|x)$$

$$J(\theta) := \mathbb{E}\tilde{g}\tilde{g}^\top = \mathbb{E}_{(x,y)\in\mathcal{D}} \mathbb{E}_{\tilde{y}\sim p_\theta(\tilde{y}|x)} \frac{\partial \ell(\tilde{y}|x)}{\partial \theta}^{\otimes 2}$$

$$\theta^{t+\delta t} = \theta^t - \delta t J(\theta^t)^{-1} \frac{\partial \ell(y_k|x_k)}{\partial \theta^t}$$

$$\frac{\mathrm{d}\theta^t}{\mathrm{d}t} = -J(\theta^t)^{-1} \mathbb{E}_{(x,y)\in D} \frac{\partial \ell(y|x)}{\partial \theta^t}$$

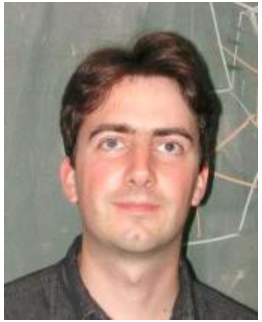
Jean-Louis Koszul & Yann Ollivier



SEE/SMF GSI'13
« Geometric Science of Information »
Ecole des Mines de Paris
August 2013

Koszul @ ENS Ulm 1940

Koszul @ Bourbaki



Ollivier @ ENS

OPEN

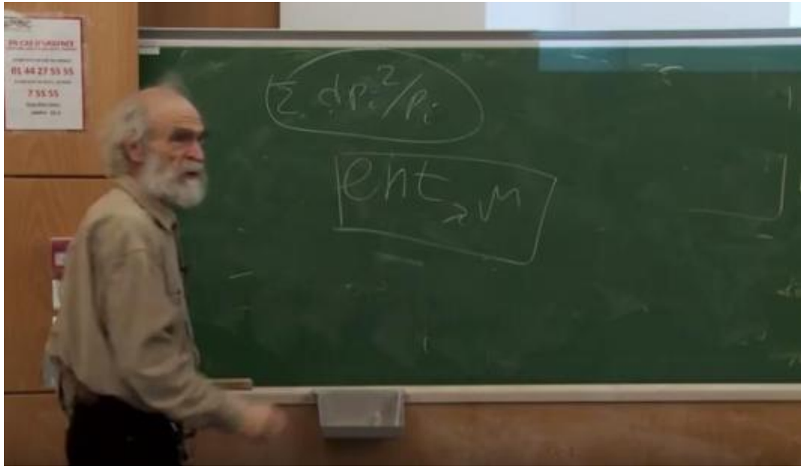
Fisher Metric by Misha Gromov (IHES, Abel Price) (M. Gromov was Yann Ollivier PhD supervisor with Pierre Pansu)

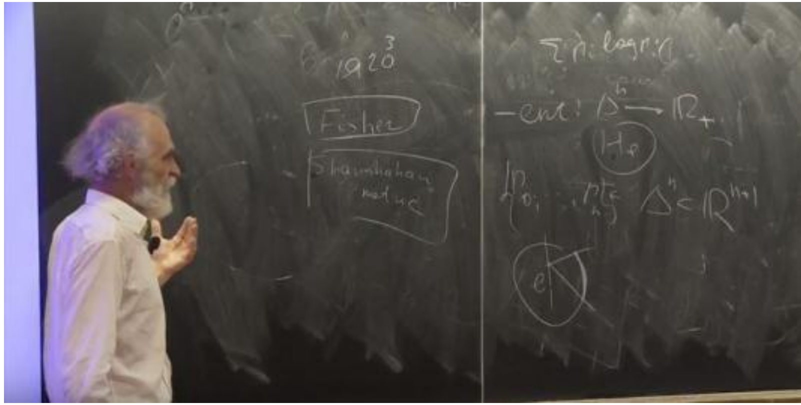
- **M. Gromov, In a Search for a Structure, Part 1: On Entropy. July 6, 2012**
  - http://www.ihes.fr/~gromov/PDF/structre-serch-entropy-july5-2012.pdf
- **Gromov Six Lectures on Probability, Symmetry, Linearity. October 2014, Jussieu, November 6th , 2014**
  - Lecture Slides & video:

  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)



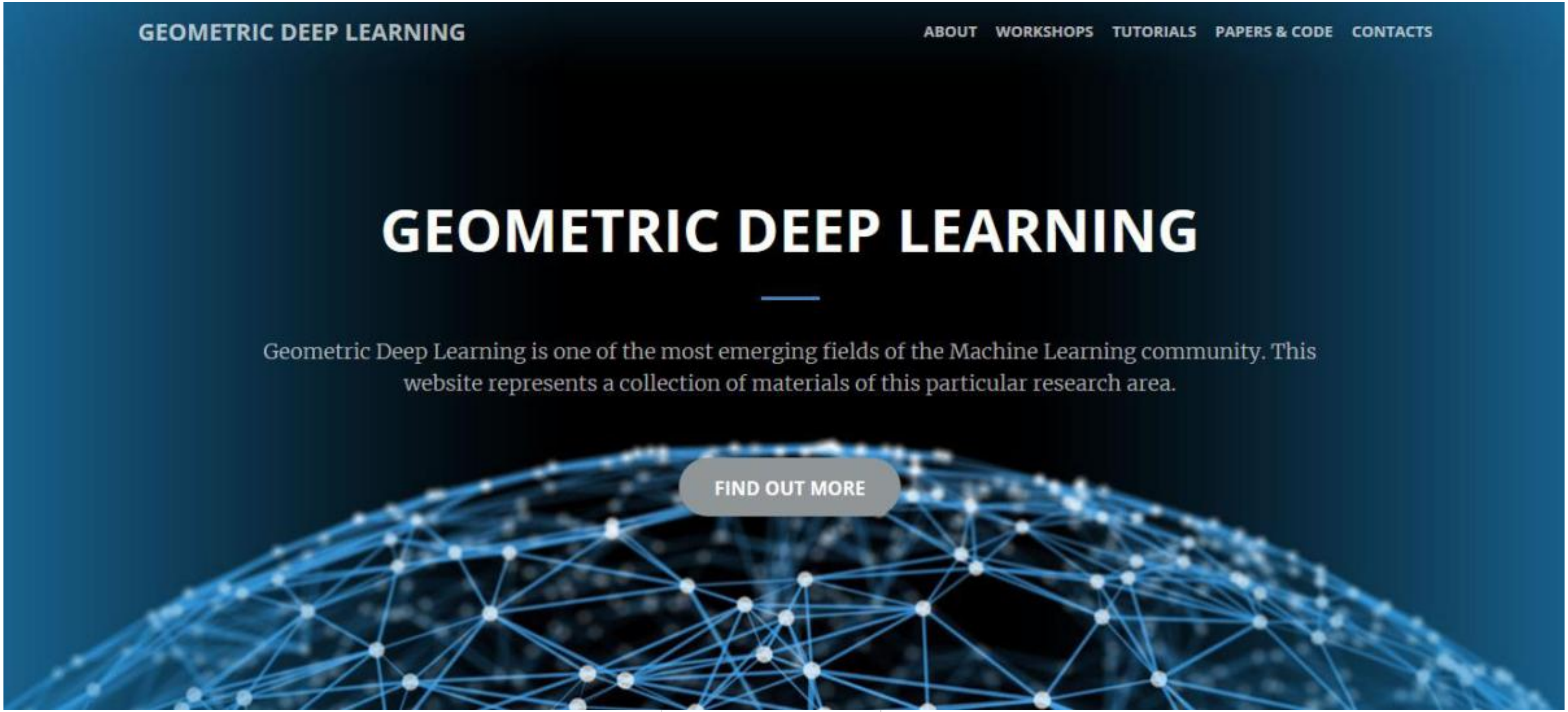


OPEN

THALES

Geometric Deep Learning

Website: http://geometricdeeplearning.com/



THALES

What to remember of my talk in 4 slides

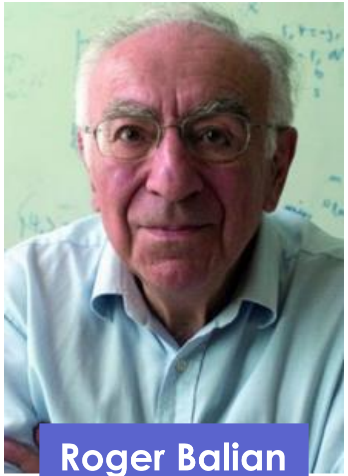
OPEN

www.thalesgroup.com

This presentation is a synthesis of 4 avant-gardistes works on Geometric structures of Information (& Statistical Physics)

Jean-Louis Koszul

Fundation of Hessian Structures Of Information Geometry (Koszul-Vinberg Characteristic function, Koszul 2-form)



Roger Balian

Fundation of Quantum Information Geometry Structures (Fisher-Balian Quantum Metric)

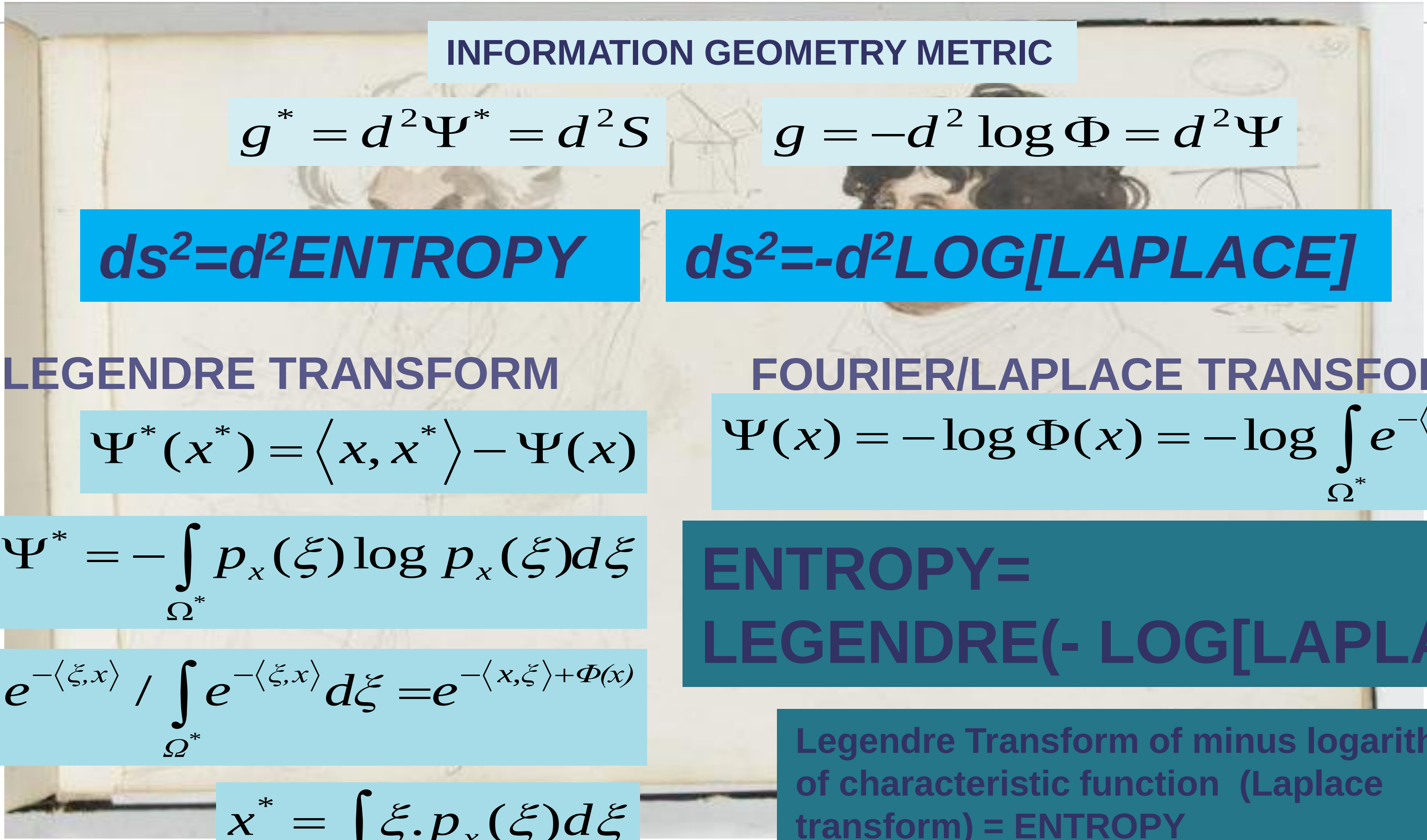
Jean-Marie Souriau

Fundation of Structures of Lie Group Thermodynamics and Statistical Geometric Mechanics (geometric temperature and entropy, Fisher-Souriau metric)

Maurice Fréchet

Fundation of Structures for probability extension in metric spaces and Fréchet Bound (Clairaut-Legendre equation of Information Geometry)

Projective Legendre Duality and Koszul Characteristic Function



INFORMATION GEOMETRY METRIC

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

ds²=d²ENTROPY

ds²=-d²LOG[LAPLACE]

LEGENDRE TRANSFORM

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

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

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

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

FOURIER/LAPLACE TRANSFORM

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

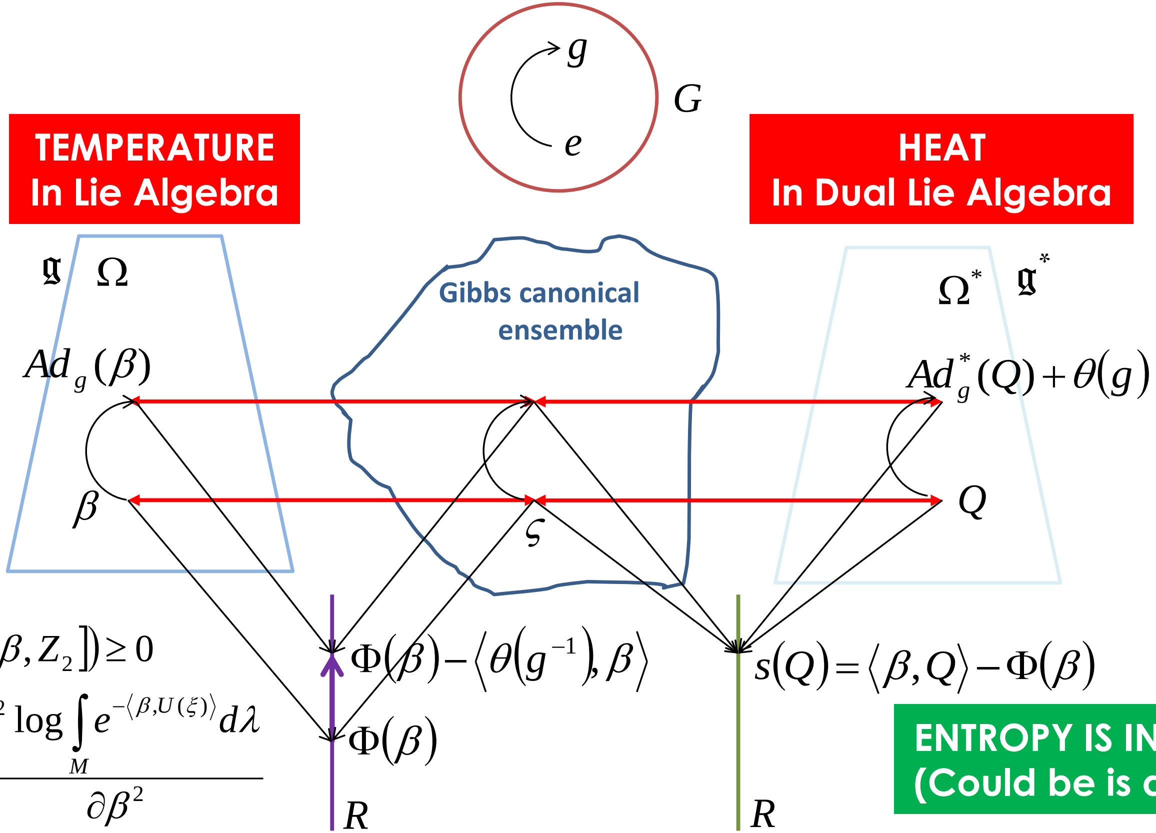
ENTROPY=
LEGENDRE(- LOG[LAPLACE])

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

OPEN

THALES

Souriau Model of Lie Group Thermodynamics



FISHER METRIC

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

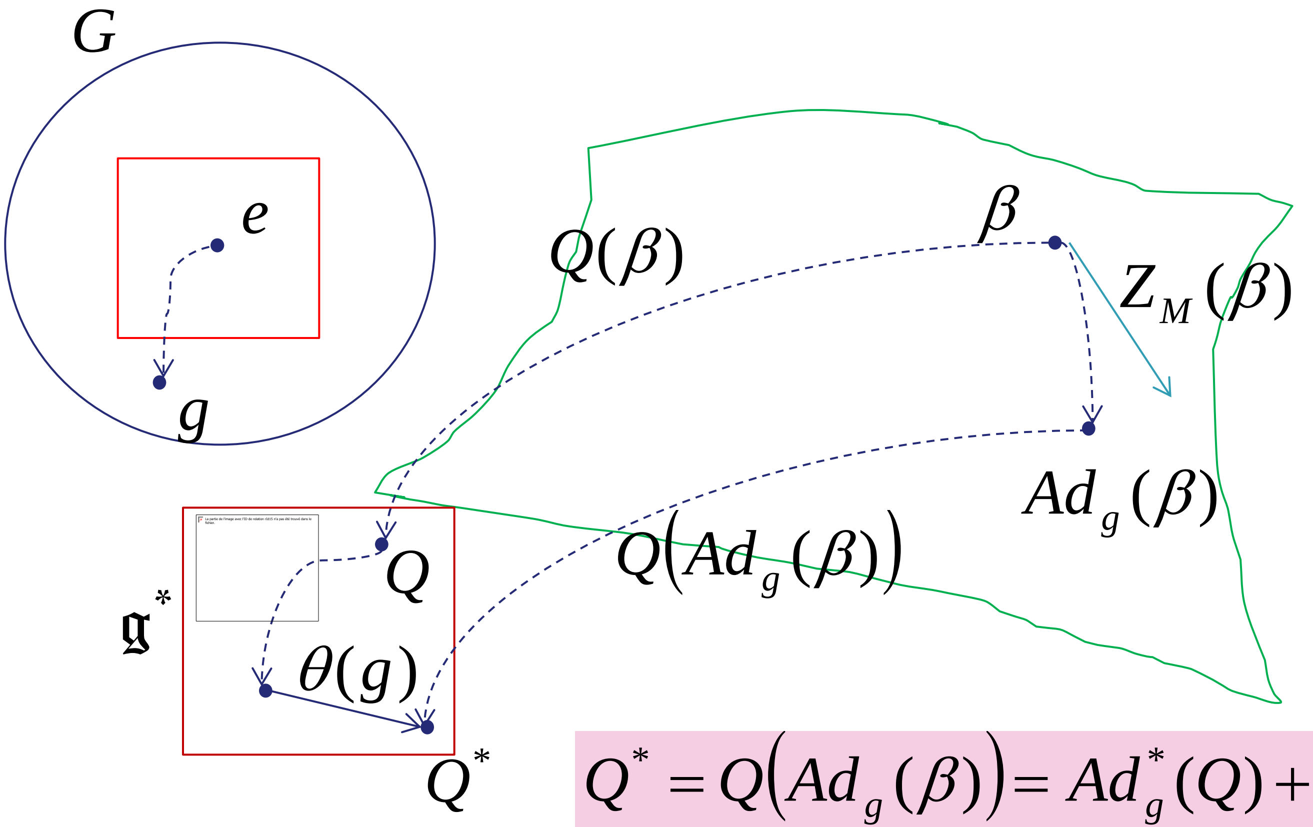
$$I(\beta)=I(Ad_g(\beta))=-\frac{\partial^2\Phi}{\partial\beta^2}=\frac{\partial^2\log\int_M e^{-\langle\beta,U(\xi)\rangle}d\lambda}{\partial\beta^2}$$

ENTROPY IS INVARIANT (Could be is defintion)

LOG OF CHARACTERISIC FUNCTION

ENTROPY

Lie Group Action on Symplectic Manifold



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

THALES

THALES

SEMINAL STRUCTURE: DUAL POTENTIAL FUNCTIONS (characteristic function & Entropy) LINKED BY LEGENDRE TRANSFORM

CONTRIBUTION OF “CORPS DES MINES” (Massieu, Poincaré, Levy, Balian)

Seminal Work of François Jacques Dominique Massieu

- **Before introducing, Information Geometry with Koszul model, we have to explain the history of « caracteristic 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

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 Chracteristic function in Probability

1986: Roger Balian

- Balian introduces the Fisher Quantum Metric as hessian of von Neumann Entropy

$$S=\phi-\frac{1}{T}\cdot\frac{\partial\phi}{\partial\left(1/T\right)}=\phi-\beta.U\quad(\text{Tr. Legendre})$$

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 : *characteristicfunction (from Poincaré)*

$$\left.\begin{aligned}S(\hat{D})&=F(\hat{X})-\langle\hat{D},\hat{X}\rangle\\F(\hat{X})&=\log Tr\exp(\hat{X})\end{aligned}\right\}$$

$$\Rightarrow ds^2=-d^2S=Tr\left[d\hat{D}d\log\hat{D}\right]$$

François Massieu: Seminal Idea of Characteristic Function

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.l. - 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

$$S = \phi - \frac{1}{T} \cdot \frac{\partial \phi}{\partial \left(1/T \right)} = \phi - \beta . U \quad (\text{Tr. Legendre})$$

S : Entropy, ϕ : Characteritic Function

MÉMOIRES

PRÉSENTÉS PAR DIVERS SAVANTS

A L'ACADÉMIE DES SCIENCES

DE L'INSTITUT NATIONAL DE FRANCE.

TOME XXII. — N° 2.

THERMODYNAMIQUE.

MÉMOIRE

SUR

LES FONCTIONS CARACTÉRISTIQUES DES DIVERS FLUIDES

ET SUR LA THÉORIE DES VAPEURS,

PAR F. MASSIEU,

INGÉNIEUR DES MINES, PROFESSEUR À LA FACULTÉ DES SCIENCES DE RENNES.

MÉMOIRES PRÉSENTÉS.

THERMODYNAMIQUE. — *Addition au précédent Mémoire sur les fonctions caractéristiques.* Note de M. F. Massieu, présentée par M. Combes.

» Cette conclusion résultait *à 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.

$$\boxed{\psi = S - \frac{U}{T}.}$$

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 publitions, 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 Funcion of Massieu and its good parameterization were discovered again by Max Planck (1897) and developed by Herbert Callen (1960) and Roger Balian.

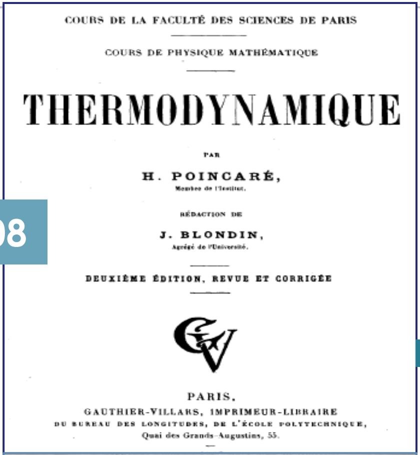
M. Planck

H. Callen

R. Balian

Henri Poincaré Re-Use for Thermodynamics and Probability

1908


COURS DE LA FACULTÉ DES SCIENCES DE PARIS

COURS DE PHYSIQUE MATHÉMATIQUE

THERMODYNAMIQUE

PAR

H. POINCARÉ,

Membre de l'Institut.

RÉDACTION DE

J. BLONDIN,

Agrégé de l'Université.

DEUXIÈME ÉDITION, REVUE ET CORRIGÉE

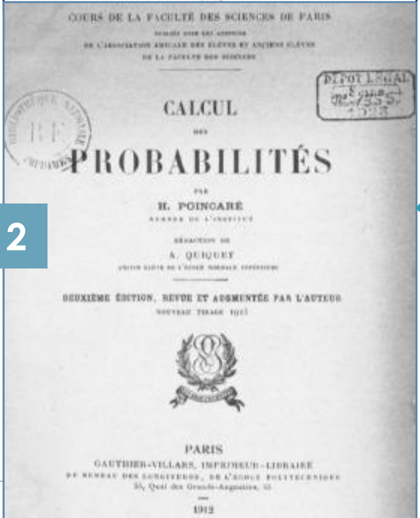
PARIS,

GAUTHIER-VILLARS, IMPRIMEUR-LIBRAIRE

DU BUREAU DES LONGITUDES, DE L'ÉCOLE POLYTECHNIQUE,

Quai des Grands-Augustins, 55.

1912


COURS DE LA FACULTÉ DES SCIENCES DE PARIS

CALCUL DES PROBABILITÉS

PAR

H. POINCARÉ

RÉDACTION DE

A. QUIQUET

DEUXIÈME ÉDITION, REVUE ET AUGMENTÉE PAR L'AUTEUR

PARIS

GAUTHIER-VILLARS, IMPRIMEUR-LIBRAIRE

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 » i probability in his Lecture of1912 (inspired by Massieu; both related by logarithm)
- It is ***introduced with Laplace Transform***
- Characteristic function of a real random variable ***defines completly its density of probability***.
- Moments of the random variable could be deduced from successive derivatives at zero of the caracteristic function.
- The 2nd characteristic function is given by the logarithm,: ***generating function of cumulants***.
- **Cumulants** have been introduced in 1889 by danish astronome, 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

$$\boxed{f(\alpha)=\int \varphi(x)\,e^{\alpha x}\,dx,}$$

si x varie d'une manière continue et si $\varphi(x)$ représente la loi de probabilité. Il est clair que

$$f(\alpha)=1+\frac{\alpha}{1}(x)+\frac{\alpha^2}{1.2}(x^2)+\frac{\alpha^3}{1.2.3}(x^3)+\ldots,$$

(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 $e^{\alpha y}$.

THALES

Quantum Information Geometry of Roger Balian (1/3)

- $Tr\left[\hat{D}\,\hat{O}\right]$ mean values through two dual spaces of observables $\hat{O}$ and of the states $\hat{D}$
- $S = -Tr\left[\hat{D}\log(\hat{D})\right]$ Entropy in space of states
- Entropy S cold be written as a scalar product $S = -\left\langle \hat{D}, \log(\hat{D}) \right\rangle$ where $\log(\hat{D})$ is an element of space of observables, allowing a physical geometrical structure in these spaces.
- The 2nd differential d^2S 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 btween state $\hat{D}$ and its neiborhood $\hat{D} + d\hat{D}$: $ds^2 = -d^2S = Tr\left[d\hat{D}.d\log\hat{D}\right]$
- Where the Riemannian metric tensor is $-S(\hat{D})$ as function of a set of independant coordinates of $\hat{D}$.

OPEN

THALES

Quantum Information Geometry of Roger Balian (2/3)

- It is possible to introduce the logarithm of a quantum characteristic function $F(\hat{X})$:

$$F(\hat{X}) = \log Tr \exp \hat{X}$$

- Von Neumann Entropy S appears as Legendre transorm of $F(\hat{X})$:

$$S(\hat{D}) = F(\hat{X}) - \left\langle \hat{D}, \hat{X} \right\rangle$$

- with $S(\hat{D}) = -Tr\hat{D} \log \hat{D} = -\left\langle \hat{D}, \log \hat{D} \right\rangle$

- 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 $\hat{X} = \beta.\hat{H}$ and where hamiltonian is $\hat{H}$.

OPEN

THALES

Quantum Information Geometry of Roger Balian (3/3)

- $dF = Tr\hat{D}d\hat{X}$ with Maximum Entropy Gibbs Density: $\hat{D} = \dfrac{\exp \hat{X}}{Tr \exp \hat{X}}$

- dF are partial derivative of $F(\hat{X})$ with respect to coordinates of $\hat{X}$. $\hat{D}$ Is hermitian, normalised and positifve and can be interpreted as a density matrix .

- Legendre Transform appears with the following development:

$$S(\hat{D}) = -Tr\hat{D}\log\hat{D} = -Tr\left(\hat{D}\left(\hat{X} - \log Tr \exp \hat{X}\right)\right) = -Tr\hat{D}\hat{X} + Tr\left(\hat{D}\right)\log Tr \exp \hat{X}$$

$$Tr\left(\hat{D}\right) = 1 \Rightarrow S(\hat{D}) = F\left(\hat{X}\right) - \left\langle \hat{D}, \hat{X} \right\rangle$$

- Roger Balian has defined the dual Riemanian metric from F, $ds^2 = d^2F$ in conjugate space $\hat{X}$:

$$ds^2 = -dS^2 = Trd\hat{D}d\hat{X} = d^2F$$

- Normalisation of $\hat{D}$ implies $Trd\hat{D} = 0$ and $Trd^2\hat{D} = 0$

Fundamental strcture of Roger Balian Quantum Information Geometry

Legendre Transform

$$S(\hat{D}) = F(\hat{X}) - \left\langle \hat{D}, \hat{X} \right\rangle$$

Von Neumann Entropy

$$S = -Tr\left[\hat{D}\log(\hat{D})\right]$$

Characteristic Function

$$F(\hat{X}) = \log Tr \exp \hat{X}$$

Maximum Entropy Density

$$\hat{D} = \frac{\exp \hat{X}}{Tr \exp \hat{X}}$$

Balian Metric of Quantum Information Geometry (1986)

$$ds^2 = -d^2S = Tr\left[d\hat{D}.d\log\hat{D}\right]$$

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 previoulsy Clairaut Equation
- Maurice Fréchet introduced Clairaut equation associated to « dstinguished densities » (densities with parameters achieving the Fréchet-Cramer-Rao Bound)

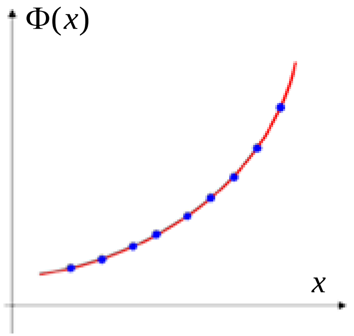
Legendre Transform interpretation

Legendre Transform

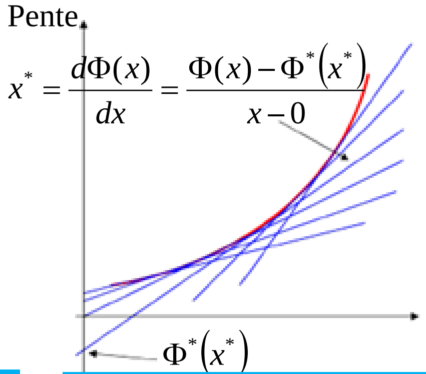
- Legendre Transform transformes one fonction defined by its value in one point in a fonction defined by its tangent.
- Used in thermodynamics and Lagrangian Mechanics

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

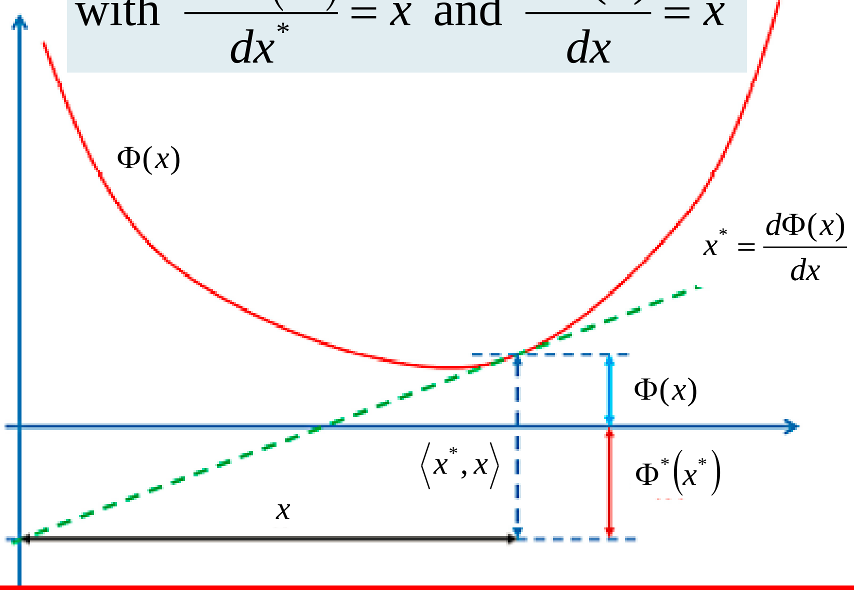
$$\text{with } \frac{d\Phi^*(x^*)}{dx^*} = x \text{ and } \frac{d\Phi(x)}{dx} = x^*$$



Classical Geometry
(curve is given by a continuum of points)



Plücker Geometry
(curve is given by the envelop of its tangents)



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.

MÉMOIRE SUR L'INTÉGRATION DE QUELQUES ÉQUATIONS AUX DIFFÉRENCES PARTIELLES.

Par M. LE GENDRE.

(I.)

De l'Équation de la moindre Surface.

ON ſait, d'après M. de la Grange, que la ſurface la moindre entre des limites données, a pour équation différentielle

$$(1+q^2)\frac{ddz}{dx^2} - 2pq\frac{ddz}{dx\,dy} + (1+p^2)\frac{ddz}{dy^2} = 0,$$

en faiſant, pour abréger, $\frac{dz}{dx} = p$, $\frac{dz}{dy} = q$. M. Monge a tenté d'intégrer cette équation dans les Mémoires de l'Académie de 1784; mais l'intégrale qu'il a donnée *(page 149)* n'étant pas à l'abri de toute objection, attendu que les ſignes

Lorſque la valeur de ω ſera connue, il eſt clair qu'on aura celles de x, y, z, exprimées en p & q; ſavoir,

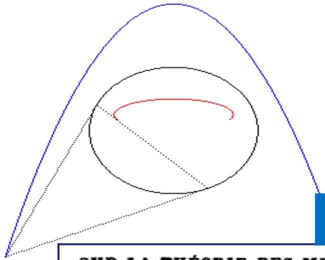
$$x = \frac{d\omega}{dp}, \quad y = \frac{d\omega}{dq},$$

$$z = px + qy - \omega.$$

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 paraboloïde*** »
- 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

SUR LA THÉORIE DES MULTIPLICITÉS ET LE CALCUL DES VARIATIONS

PAR M. E. VESSIOT.


LEÇONS

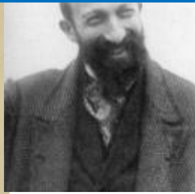
CALCUL DES VARIATIONS

J. HADAMARD

TOME PREMIER

LA VARIATION PREMIÈRE ET LES CONDITIONS DU PREMIER ORDRE

J. Hadamard



Partons d'abord du problème de Lagrange, et posons comme précédemment

(200) $$q_i = f_{y'_i} \qquad (i = 1, \dots n)$$

(201) $$\Pi = \sum_{i=1}^{n} y'_i f_{y'_i} - f.$$

(202) $$\Pi = \Pi(q_1, \dots q_n, y_1, \dots y_n, x)$$

La différentielle totale de Π sera dès lors la même qu'au n° **140**

(204) $$d\Pi = \sum_i y'_i dq_i - f_{y_i} dy_i.$$

La transformation de Legendre, définie par les équations (200), (201) reviendra encore à prendre la polaire réciproque de cette figurative par rapport au paraboloïde

(205) $$y'^2_1 + y'^2_2 - 2f_0 = 0.$$

Geometric Interpretation of Legendre Transform by Reciprocal Polar with respect to a paraboloid

- First, Let's consider the surface: $z = f(x,y)$ with $p = \dfrac{\partial z}{\partial x}$ and $q = \dfrac{\partial z}{\partial y}$
- 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) :

$$z - z_0 = p_0(x - x_0) + q_0(y - y_0) \Rightarrow p_0 x + q_0 y - z - (p_0 x_0 + q_0 y_0 - z_0) = 0$$

- This equality provides:

$$X = p_0 \;,\; Y = q_0 \;,\; Z = p_0 x_0 + q_0 y_0 - z_0$$

This is the **Legendre Transform**

(1) GOURSAT, *Traité d'Analyse*, tome I, p. 89; HUMBERT, *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 $Q = \frac{\partial Z}{\partial Y}$

$$Z = px + qy - z \Rightarrow dZ = pdx + qdy - dz + xdp + ydq \underset{\substack{dz = pdx + qdy \\ X = p, Y = q}}{\Rightarrow} dZ = xdX + ydQ \Rightarrow P = x, Q = y$$

- We compute 2nd derivative of Z:

$$R = \frac{\partial^2 Z}{\partial X^2} = \frac{\partial P}{\partial X} = \frac{\partial x}{\partial X}\ ,\ T = \frac{\partial^2 Z}{\partial X \partial Y} = \frac{\partial P}{\partial Y} = \frac{\partial Q}{\partial X} = \frac{\partial x}{\partial Y} = \frac{\partial y}{\partial X}\ ,\ S = \frac{\partial^2 Z}{\partial Y^2} = \frac{\partial Q}{\partial Y} = \frac{\partial y}{\partial Y}$$

$$\begin{cases} dX = rdx + sdy \\ dY = sdx + tdy \end{cases} \Rightarrow \begin{cases} dx = \frac{t}{rt - s^2} dX - \frac{s}{rt - s^2} dY \\ dy = \frac{-s}{rt - s^2} dX + \frac{r}{rt - s^2} dY \end{cases} \Rightarrow \begin{cases} R = \frac{\partial x}{\partial X} = \frac{t}{rt - s^2} \\ S = \frac{\partial x}{\partial Y} = \frac{-s}{rt - s^2} \\ T = \frac{\partial y}{\partial Y} = \frac{r}{rt - s^2} \end{cases} \Rightarrow \begin{cases} r = \frac{T}{RT - S^2} \\ s = \frac{-S}{RT - S^2} \\ t = \frac{R}{RT - S^2} \end{cases}$$

$$r = \frac{\partial^2 z}{\partial x^2}, t = \frac{\partial^2 z}{\partial y^2}, s = \frac{\partial^2 z}{\partial x \partial y}$$

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:

$$dZ - PdX - QdY = \rho\left(dz - pdx - qdy\right)$$

where ρ is a function of *x,y,z,p* and *q*

Proof:

$$\begin{cases} p = \dfrac{\partial z}{\partial x} \\ q = \dfrac{\partial z}{\partial y} \end{cases} \Rightarrow dz - pdx - qdy = 0 \Rightarrow dZ = PdX + QdY \Rightarrow \begin{cases} P = \dfrac{\partial Z}{\partial X} \\ Q = \dfrac{\partial Z}{\partial Y} \end{cases}$$

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:

$$dz - pdx - qdy = d(z - qy) - pdx + ydq$$

$$\text{Set} \begin{cases} Z = z\text{-}qy, X = x, Y = q \\ P = p, Q = \text{-}y \end{cases} \Rightarrow dZ - pdX - QdY = dz - pdx - qdy$$

- Then $\rho = 1$, and we have a contact transformation, also valid when Legendre transform is no longervalide (when $rt - s^2 = 0$, *p* and *q* are not independant)

Legendre transformation and Ampere transformation

- Legendre transform is obtained by same equality

$$dz - pdx - qdy = d(z - px - qy) - xdp - ydq$$

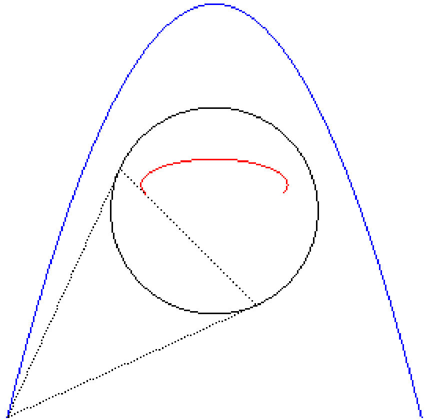
- We can set $X = p, Y = q, Z = z - px - qy$

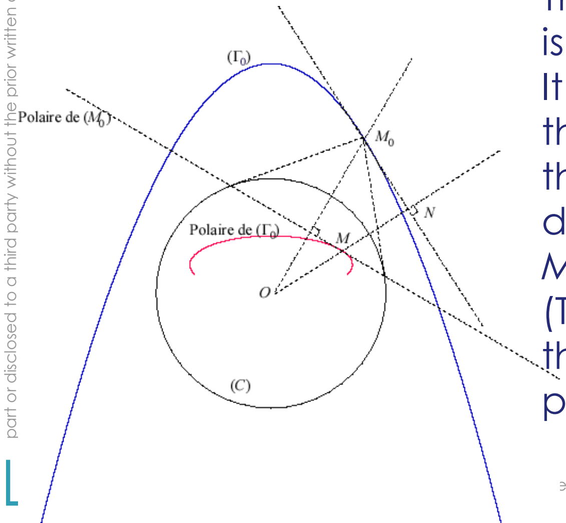
$$P = x, Q = y$$

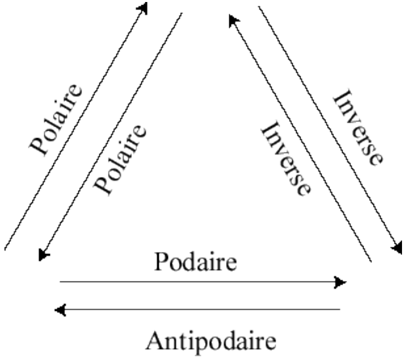
OPEN

THALES

Duality Principle & Reciprocal Polar





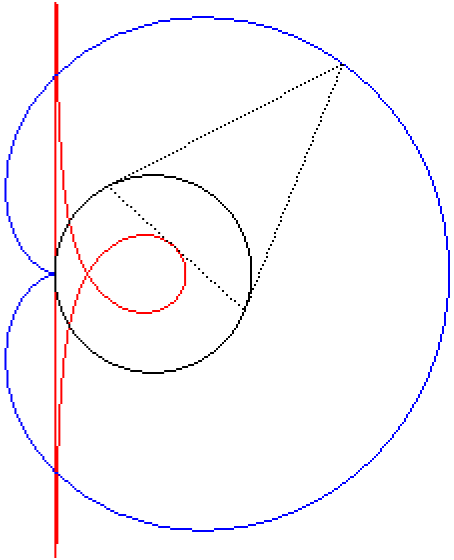


Reciprocal Polar with respect to a circle

The characteristic point of the polar of M0 is also the pole of the tangent to Γ_0 at M_0; It is therefore the point of intersection of the perpendicular to this tangent passing through O and the polar. The point *N* thus describes the podar of Γ_0, and the point *M*, the inverse of this podar.

(The inverse of the polar with respect to the same circle is none other than the podaire)

Maclaurin trisectrix curve is reciprocal polar of cardioïd with respect of its conchoïdal circle



THALES

Project of 2nd Edition of his Book (Send by Claude Vallée)

J. M. SOURIAU

DYNAMIC SYSTEMS STRUCTURE

§ 16	convexité	231 -
§ 17	Mesures	248
§ 18	Etats statistiques	305
§ 19	Thermodynamique	345

Souriau Definition of Entropy by Legendre Transform

- Let E a vector space of finite size, μ a measure of its dual 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 is definition set $def(F)$. Its p-th derivative is given by the following convergent integral for all point inside :

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

OPEN

THALES

Souriau Definition of Entropy by Legendre Transform

Souriau Theorem:

- Let E a vector space of finite size, μ a non-zero positive measure of dual space E^*, F its Laplace transform, then:
  - F is semi-definite convex function, $F(\alpha) > 0, \forall \alpha \in def(F)$
  - $f = \log(F)$ is convex and semi-continuous
  - Let α an interior point of $def(F)$ 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 envelop(support(μ)) = E^*

Souriau Definition of Entropy by Legendre Transform

Lemma:

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

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

such that the integral is converging.

Demo:

- The integral is strictly positive when it converges, and then insures existence of its logarithm.
- The epigraph of 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 shows that this epigraph is convex.

Souriau Definition of Entropy by Legendre Transform

Souriau Entropy Definition:

- (Neg)Entropie is given by the legendre transform of :

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

- We call "Boltzmann Law" (relative to λ) all measure μ of X such that the set of real values $\mu(h) - \Phi(h)$, $h \in def(\Phi)$ and h is μ -integrable.

Project of 2nd Edition of his Book (Send by Claude Vallée)

ENTROPIE

Lemme :

(17.190) Soit X un espace localement compact; soit λ une mesure positive de X ayant X comme support.

Alors la fonction Φ :

$$\Phi(h) = \operatorname{Log} \int_X e^{h(x)} \lambda(x)\, dx \quad \left[\forall h \in \mathcal{C}(X) \text{ tel que l'intégrale converge} \right.$$

est convexe.

D'après (17.141), l'intégrale est strictement positive lorsqu'elle est convergente, ce qui assure l'existence de son logarithme. L'épigraphe de Φ (16.46) est l'ensemble des $\begin{pmatrix} h \\ y \end{pmatrix}$ tels que $\int_X e^{h(x)-y} \lambda(x)\, dx \leqslant 1$; la convexité de l'exponentielle montre que cet épigraphe est convexe.

C.Q.F.D.

Nous allons définir – sous le nom de négentropie – une transformée de Legendre formelle de cette fonction Φ :

Définition (suite de (17.190))

(17.191) Nous appellerons loi de Boltzmann (relative à λ) toute mesure μ de X telle que l'ensemble de réels

$$\mu(h) - \Phi(h) \quad \left[\begin{array}{l} h \in \operatorname{def}(\Phi) \\ \quad \text{et} \\ h \quad \mu\text{-intégrable} \end{array} \right]$$

OPEN

THALES

THALES

The geometry of Hessian Structures in Information Geometry: Maurice Fréchet & Jean-Louis Koszul

OPEN

Hessian Structure of Information Geometry

OPEN

Fisher Matrix and Fréchet(-Cramer-Rao) Bound

In 1943, Maurice Fréchet published a seminal paper where he introduced first lower bound for all estimator, given by the inverse of Fisher Matrix

- **Fréchet M., 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.**

$\hat{\theta}$ estimator of θ, Fréchet bound :

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

$$I(\theta)\ \text{Fisher Matrix : } \left[I(\theta)\right]_{i,j} = -E\left[\frac{\partial^2 \log P(Z/\theta)}{\partial\theta_i \partial\theta_j}\right] \quad \text{with} \quad \theta = \begin{bmatrix}\theta_1 \\ \vdots \\ \theta_n\end{bmatrix}$$

- Fréchet informed us that the content of this paper is extracted from his IHP Lecture of Winter 1939 « *Le contenu de ce mémoire a formé une partie de notre cours de statistique mathématique à l'Institut Henri Poincaré pendant l'hiver 1939-1940* ».

Fisher Matrix, Fisher Metric and Information Geometry

- Classically, Fisher Metric of Information Geometry is introduced through Kullback Divergence and 2nd Taylor of ordre 2 expansion:

$$ds^2 = K\left[P(Z/\theta), P(Z/\theta + d\theta)\right] \underset{\substack{Taylor\\ordre\ 2}}{=} d\theta^T I(\theta) d\theta = \sum_{i,j=1}^{n} g_{ij} d\theta_i d\theta_j$$

K : Divergence de Kullback

$$K[P,Q] = \int P(Z/\theta) \log\left(\frac{P(Z/\theta)}{Q(Z/\theta)}\right) dZ$$

- Even if we can introduce Kullback divergence (combinatoricstools and stirling formula), this apprach is not fully satisfactory. We prefer to introduce Information Geometry from « **characteristic function** » (introduved by **François Massieu** in Thermodynalics and reused by **Henri Poincaré** in probability) and its developements by **Jean-Louis Koszul**.

Seminal Papers on Information Geometry

Information Geometry:

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

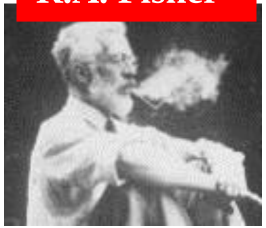
CRFD Bound

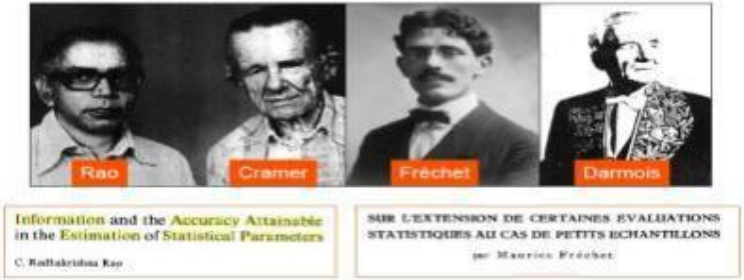
$$E\left[\left(\theta-\hat{\theta}\right)\left(\theta-\hat{\theta}\right)^{+}\right]\geq I(\theta)^{-1}$$

Fisher Information Matrix

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

R.A. Fisher





1945

1943 (IHP Lecture 1939)

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

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

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

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

N. N. Chentsov

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

Fisher Metric and Information Geometry (IG)

IG could be introduced with Koszul-Vinberg Characteristic Function:

$$\psi_\Omega(x) = \int_{\Omega^*} e^{-\langle x,\xi\rangle} d\xi, \ \forall x \in \Omega$$ with Ω and Ω^* are dual convex cones

$$\psi_\Omega(x+\lambda u) = \psi_\Omega(x) - \lambda\langle x^*, u\rangle + \frac{\lambda^2}{2}\langle K(x)u, u\rangle + \ldots$$

Density is given by Solution of Maximum Entropy: $\Phi^*(\bar{\xi}) = -\int_{\Omega^*} p_{\bar{\xi}}(\xi)\log p_{\bar{\xi}}(\xi) d\xi$

$$\underset{p}{Max}\left[-\int_{\Omega^*} p_{\bar{\xi}}(\xi)\log p_{\bar{\xi}}(\xi)d\xi\right] \text{ such that } \int_{\Omega^*} p_{\bar{\xi}}(\xi)d\xi = 1 \text{ and } \int_{\Omega^*} \xi . p_{\bar{\xi}}(\xi)d\xi = \bar{\xi}$$

$$p_{\bar{\xi}}(\xi) = \frac{e^{-\langle \xi, \Theta^{-1}(\bar{\xi})\rangle}}{\int_{\Omega^*} e^{-\langle \xi, \Theta^{-1}(\bar{\xi})\rangle} d\xi} \text{ with } \bar{\xi} = \Theta(x) = \frac{\partial \Phi(x)}{\partial x} \text{ where } \Phi(x) = -\log \int_{\Omega^*} e^{-\langle x,\xi\rangle} d\xi = -\log \psi_\Omega(x)$$

The inversion $\Theta^{-1}(\bar{\xi})$ is given by Legendre transform based on :

$$\Phi^*(x^*) = \langle x, x^*\rangle - \Phi(x) \text{ with } x^* = \frac{d\Phi(x)}{dx} \quad \Phi(x) = -\log \int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi \quad \forall x \in \Omega \text{ and } \forall x^* \in \Omega^*$$

Fisher Metric and Information Geometry (IG)

- Maurice Fréchet, studying "distinguished functions" (densities with estimator reaching the Fréchet-Darmois bound), have also observed that solution should verify the Alexis **Clairaut Equation**:

$$\Phi^*(x^*) = \left\langle \Theta^{-1}(x^*), x^* \right\rangle - \Phi\left[\Theta^{-1}(x^*)\right] \quad \forall x^* \in \left\{\Theta(x) / x \in \Omega\right\}$$

- Fisher Metric appears as hessian of characteristic function logarithm:

$$\log p_x(\xi) = -\left\langle x, \xi \right\rangle + \Phi(x) \Rightarrow \frac{\partial^2 \log p_x(\xi)}{\partial x^2} = \frac{\partial^2 \Phi(x)}{\partial x^2}$$

$$I(x) = -E_\xi\left[\frac{\partial^2 \log p_x(\xi)}{\partial x^2}\right] = -\frac{\partial^2 \Phi(x)}{\partial x^2}$$

$$\frac{\partial^2 \Phi}{\partial x^2} = \left[\frac{\partial^2 \Phi^*}{\partial x^{*2}}\right]^{-1}$$

$$I(x) = \frac{\partial^2 \log \psi_\Omega(x)}{\partial x^2} = E_\xi\left[\xi^2\right] - E_\xi\left[\xi\right]^2 = Var(\xi)$$

THALES

Fisher Matrix, Cramer-Rao-Fréchet-Darmois Bound & Information Geometry

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

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

- Rao has proposed to introduce a differential metric in space of parameters of probability density (axiomatized by N. Chentsov):

$$ds_\theta^2 = Kullback\_Divergence\left(p_\theta(z), p_{\theta+d\theta}(z)\right)$$

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

$$w = W(\theta) \Rightarrow ds_w^2 = ds_\theta^2$$

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

Gibbs density (Maximum Entropy) and Legendre Transform

Maximum Entropy Principle for Density Estimation: Gibbs-Duhem Density

$$\underset{p}{Max}\left[-\int_{\Omega^*} p_{\hat{\xi}}(\xi)\log p_{\hat{\xi}}(\xi).d\xi\right] \text{ such that } \int_{\Omega^*} p_{\hat{\xi}}(\xi)d\xi = 1 \text{ and } \int_{\Omega^*} \xi.p_{\hat{\xi}}(\xi)d\xi = \hat{\xi}$$

$$p_{\hat{\xi}}(\xi) = \frac{e^{-\left\langle \Theta^{-1}(\hat{\xi}),\xi \right\rangle}}{\int_{\Omega^*} e^{-\left\langle \Theta^{-1}(\hat{\xi}),\xi \right\rangle}.d\xi} \quad \hat{\xi} = \Theta(\beta) = \frac{\partial \Phi(\beta)}{\partial \beta} \quad \text{where } \Phi(\beta) = -\log \psi_\Omega(\beta)$$

$$\psi_\Omega(\beta) = \int_{\Omega^*} e^{-\langle \beta,\xi \rangle} d\xi \quad , \quad S(\hat{\xi}) = -\int_{\Omega^*} p_{\hat{\xi}}(\xi)\log p_{\hat{\xi}}(\xi).d\xi \text{ and } \beta = \Theta^{-1}(\hat{\xi})$$

$$S(\hat{\xi}) = \left\langle \hat{\xi}, \beta \right\rangle - \Phi(\beta)$$

LEGENDRE TRANSFORM

Fisher Metric and Information Geometry (IG)

Fisher Metric appears as hessian of characteristic function logarithm:

$$\log p_{\hat{\xi}}(\xi) = -\langle \xi, \beta \rangle + \Phi(\beta)$$

$$S(\hat{\xi}) = -\int_{\Omega^*} p_{\hat{\xi}}(\xi).\log p_{\hat{\xi}}(\xi).d\xi = -E\left[\log p_{\hat{\xi}}(\xi)\right]$$

$$S(\hat{\xi}) = \langle E[\xi], \beta \rangle - \Phi(\beta) = \langle \hat{\xi}, \beta \rangle - \Phi(\beta)$$ **LEGENDRE TRANSFORM**

$$I(\beta) = -E\left[\frac{\partial^2 \log p_\beta(\xi)}{\partial \beta^2}\right] = -E\left[\frac{\partial^2\left(-\langle \xi, \beta \rangle + \Phi(\beta)\right)}{\partial \beta^2}\right] = -\frac{\partial^2 \Phi(\beta)}{\partial \beta^2}$$

$$\hat{\xi} = \frac{\partial \Phi(\beta)}{\partial \beta}$$

$$I(\beta) = E\left[\frac{\partial \log p_\beta(\xi)}{\partial \beta} \frac{\partial \log p_\beta(\xi)}{\partial \beta}^T\right] = E\left[\left(\xi - \hat{\xi}\right)\left(\xi - \hat{\xi}\right)^T\right] = E\left[\xi^2\right] - E[\xi]^2 = Var(\xi)$$

2 metrics in dual coordinates systems for dual potential functions

1st Metric of Information Geometry: Fisher Metric = hessian of logarithm characteristic function

$$I(\beta) = -E\left[\frac{\partial^2 \log p_\beta(\xi)}{\partial \beta^2}\right] = -\frac{\partial^2 \Phi(\beta)}{\partial \beta^2}$$

$$ds_g^2 = d\beta^T I(\beta) d\beta = \sum_{ij} g_{ij} d\beta_i d\beta_j \quad \text{with} \quad g_{ij} = [I(\beta)]_{ij}$$

2nd Metric of Information Geometry: hessian of Shannon Entropy

$$\frac{\partial^2 S(\hat{\xi})}{\partial \hat{\xi}^2} = \left[\frac{\partial^2 \Phi(\beta)}{\partial \beta^2}\right]^{-1} \qquad S(\hat{\xi}) = \langle \hat{\xi}, \beta \rangle - \Phi(\beta)$$

$$ds_h^2 = d\hat{\xi}^T \left[\frac{\partial^2 S(\hat{\xi})}{\partial \hat{\xi}^2}\right] d\hat{\xi} = \sum_{ij} h_{ij} d\hat{\xi}_i d\hat{\xi}_j \quad \text{with} \quad h_{ij} = \left[\frac{\partial^2 S(\hat{\xi})}{\partial \hat{\xi}^2}\right]_{ij}$$

Same Distance for Dual metrics

$$ds_g^2 = ds_h^2$$

OPEN

THALES

Example : Gaussian scalar distribution

For Gaussian Law, Fisher Information matrix is given by :

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

- Fisher matrix induced the following differential metric :

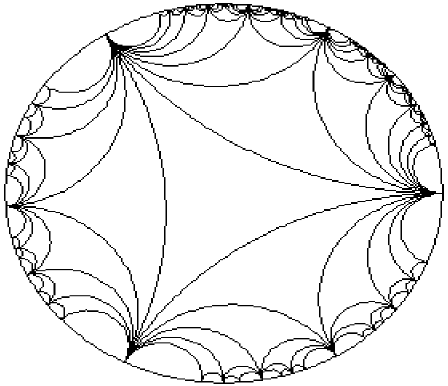
$$\boxed{ds^2=d\theta^T.I(\theta).d\theta=\frac{dm^2}{\sigma^2}+2.\frac{d\sigma^2}{\sigma^2}=2.\sigma^{-2}\left[\left(\frac{dm}{\sqrt{2}}\right)^2+(d\sigma)^2\right]}$$

- Poincaré model of hyperbolic Space :

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

$$\Rightarrow ds^2=8.\frac{|d\omega|^2}{\left(1-|\omega|^2\right)^2}$$

Geometry of Gaussian Law Is Geometry of Hyperbolic Poincare Space



OPEN

THALES

Example : Gaussian scalar distribution

Gaussian Law Metric :

- If we set $r = |\omega|$, we can integrate along one radial :

$$ds^2 = 8.\left(\frac{dr}{1-r^2}\right)^2 = 2.d\ln\frac{1+r}{1-r}$$

- Homeomorphisme is used then to compute distance between two arbitrary points in the unit disk :

$$v = \phi_\tau(\omega) = \frac{\omega-\tau}{\bar{\tau}\omega-1}.e^{j.\varphi} \quad \text{and} \quad 0 = \phi_\tau(\tau)$$

- Distance between two Gaussian Law is then given by :

$$D^2(\{m_1,\sigma_1\},\{m_2,\sigma_2\}) = 2.\left(\ln\frac{1+\delta(\omega,\tau)}{1-\delta(\omega,\tau)}\right)^2 \quad \text{with} \quad \delta(\omega,\tau) = \left|\frac{\omega-\tau}{1-\omega\bar{\tau}}\right|$$

$$z = \frac{m}{\sqrt{2}} + i.\sigma \quad \text{and} \quad \omega = \frac{z-i}{z+i}$$

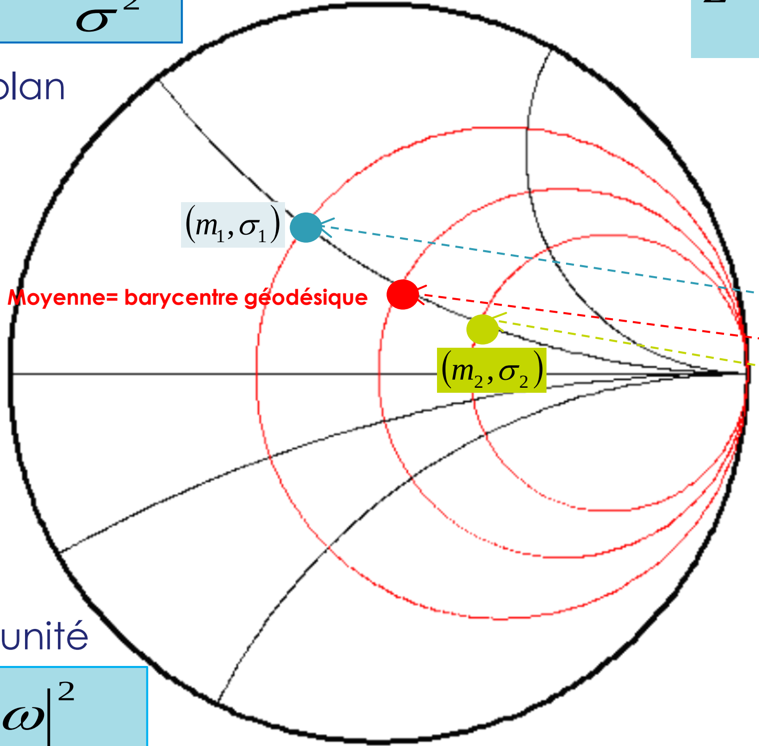
Monovariate Gaussian = 1 point in Poincaré Unit Disk

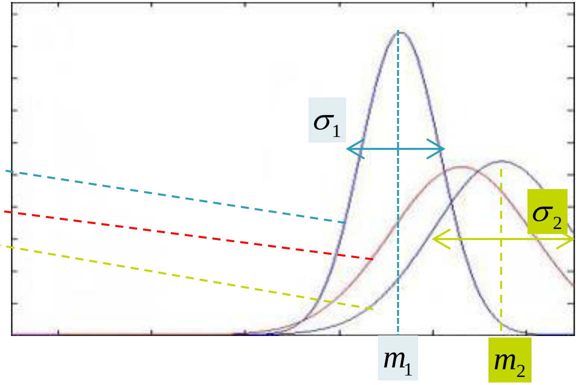
$$ds^2 = \frac{dm^2}{\sigma^2} + 2.\frac{d\sigma^2}{\sigma^2}$$

Métrique demi-plan

$$z = \frac{m}{\sqrt{2}} + i.\sigma$$

$$\omega = \frac{z-i}{z+i} \quad (|\omega| < 1)$$





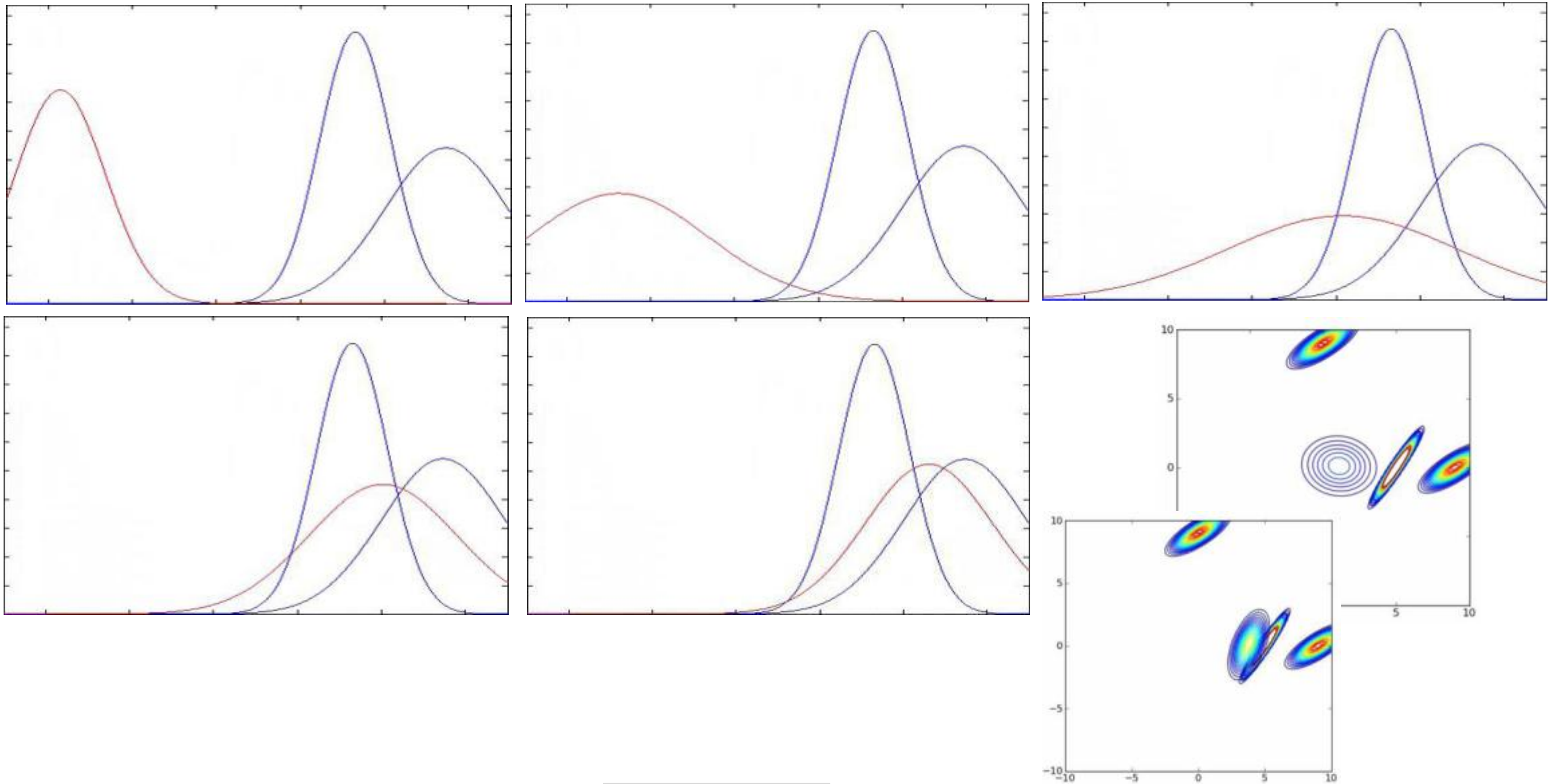
Métrique disque unité

$$ds^2 = 8.\frac{|d\omega|^2}{\left(1-|\omega|^2\right)^2}$$

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

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

Mean/Mediane of Gaussian densities



Example of Multivariate Gaussian Law (real case)

Multivariate Gaussian law parameterized by moments

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

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

$$= \frac{1}{2} z^T R^{-1} z - m^T R^{-1} z + \frac{1}{2} m^T R^{-1} m$$

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

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

Example of Multivariate Gaussian Law (real case)

Multivariate Gaussian Density given by their moments (and not cumulants)

$$p_{\hat{\xi}}(\xi) = \frac{1}{\int\limits_{\Omega^*} e^{-\langle \xi,\beta \rangle}.d\xi} e^{-\langle \xi,\beta \rangle} = \frac{1}{Z} e^{-\langle \xi,\beta \rangle} \quad \text{with } \log(Z) = n\log(2\pi) + \frac{1}{2}\log\det(R) + \frac{1}{2} m^T R^{-1} m$$

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

$$R = E\left[(z-m)(z-m)^T\right] = E\left[zz^T - mz^T - zm^T + mm^T\right] = E\left[zz^T\right] - mm^T$$

1st Potential function (Free Energy / logarithm of characteristic function)

$$\psi_\Omega(\beta) = \int\limits_{\Omega^*} e^{-\langle \xi,\beta \rangle}.d\xi \quad \text{and} \quad \Phi(\beta) = -\log\psi_\Omega(\beta) = \frac{1}{2}\left[-Tr\left[H^{-1}aa^T\right] + \log\left[(2)^n \det H\right] - n\log(2\pi)\right]$$

Relation between 1st Potential function and moment

$$\frac{\partial\Phi(\beta)}{\partial\beta} = \frac{\partial\left[-\log\psi_\Omega(\beta)\right]}{\partial\beta} = \int\limits_{\Omega^*} \xi \frac{e^{-\langle \xi,\beta \rangle}}{\int\limits_{\Omega^*} e^{-\langle \xi,\beta \rangle}.d\xi}.d\xi = \int\limits_{\Omega^*} \xi.p_{\hat{\xi}}(\xi)d\xi = \hat{\xi}$$

Example of Multivariate Gaussian Law (real case)

2nd Potential function (Shannon Entropy) as Legendre Transform of 1st one:

$$S(\hat{\xi}) = \left\langle \hat{\xi}, \beta \right\rangle - \Phi(\beta) \quad \text{with} \quad \frac{\partial \Phi(\beta)}{\partial \beta} = \hat{\xi} \quad \text{and} \quad \frac{\partial S(\hat{\xi})}{\partial \hat{\xi}} = \beta$$

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

How to make Density dependent on moments only:

$$\hat{\xi} = \frac{\partial \Phi(\beta)}{\partial \beta} = \Theta(\beta) \Rightarrow \beta = \Theta^{-1}(\hat{\xi}) \quad \text{or} \quad \beta = \frac{\partial S(\hat{\xi})}{\partial \hat{\xi}} \Rightarrow \beta = \begin{bmatrix} a \\ H \end{bmatrix} = \begin{bmatrix} -R^{-1}m \\ \frac{1}{2}R^{-1} \end{bmatrix}$$

$$p_{\hat{\xi}}(\xi) = \frac{1}{\int_{\Omega^*} e^{-\langle \xi, \beta \rangle}.d\xi} e^{-\langle \xi, \beta \rangle} \quad \text{with} \quad \left\langle \xi, \beta \right\rangle = a^T z + z^T H z = Tr\left[z a^T + H^T z z^T \right] = -m^T R^{-1} z + \frac{1}{2} z^T R^{-1} z$$

$$S(\hat{\xi}) = -\int_{\Omega^*} p_{\hat{\xi}}(\xi) \log p_{\hat{\xi}}(\xi).d\xi = \frac{1}{2}\left[\log(2)^n \det\left[H^{-1}\right] + n \log(2\pi.e) \right] = \frac{1}{2}\left[\log \det[R] + n \log(2\pi.e) \right]$$

OPEN

THALES

General Scheme based on Cartan-Killing form

$\langle .,. \rangle$ inner product from Cartan - Killing Form :

$$\langle \hat{\xi}, \beta \rangle = -B(\hat{\xi}, \theta(\beta)) \quad \text{with} \quad B(\hat{\xi}, \theta(\beta)) = Tr\left(Ad_{\hat{\xi}} Ad_{\theta(\beta)}\right)$$

$$S(\hat{\xi}) = \langle \hat{\xi}, \beta \rangle - \Phi(\beta)$$

Legendre Transform

$$\Phi(\beta) = -\log \psi_{\Omega}(\beta)$$

$$S(\hat{\xi}) = -\int_{\Omega^*} p_{\hat{\xi}}(\xi) \log p_{\hat{\xi}}(\xi) . d\xi$$

$$\text{with} \quad \psi_{\Omega}(\beta) = \int_{\Omega^*} e^{-\langle \beta, \xi \rangle} d\xi$$

$$p_{\hat{\xi}}(\xi) = \frac{e^{-\langle \Theta^{-1}(\hat{\xi}), \xi \rangle}}{\int_{\Omega^*} e^{-\langle \Theta^{-1}(\hat{\xi}), \xi \rangle} . d\xi} \qquad \hat{\xi} = \Theta(\beta) = \frac{\partial \Phi(\beta)}{\partial \beta}$$

$$\beta = \frac{\partial S(\hat{\xi})}{\partial \hat{\xi}}$$

$$I(\beta) = -E\left[\frac{\partial^2 \log p_{\beta}(\xi)}{\partial \beta^2}\right]$$

$$ds_g^2 = \sum_{ij} g_{ij} d\beta_i d\beta_j$$

$$ds_h^2 = \sum_{ij} h_{ij} d\hat{\xi}_i d\hat{\xi}_j$$

$$ds_g^2 = ds_h^2$$

$$I(\beta) = -\frac{\partial^2 \Phi(\beta)}{\partial \beta^2}$$

$$\text{with} \quad g_{ij} = \left[\frac{\partial^2 \Phi(\beta)}{\partial \beta^2}\right]_{ij}$$

$$\text{with} \quad h_{ij} = \left[\frac{\partial^2 S(\hat{\xi})}{\partial \hat{\xi}^2}\right]_{ij}$$

Application for Density of Symmetric Positive Definite Matrices

If we apply previous equation for Symmetric Positive Definite Matrices:

$$p_{\hat{\xi}}(\xi) = \frac{e^{-\left\langle \Theta^{-1}(\hat{\xi}),\xi \right\rangle}}{\int\limits_{\Omega^*} e^{-\left\langle \Theta^{-1}(\hat{\xi}),\xi\xi \right\rangle}.d\xi} \qquad \hat{\xi} = \Theta(\beta) = \frac{\partial \Phi(\beta)}{\partial \beta} \qquad \Phi(\beta) = -\log \psi_\Omega(\beta)$$

$$\text{with} \quad \psi_\Omega(\beta) = \int\limits_{\Omega^*} e^{-\langle \beta,\xi \rangle} d\xi$$

$$\langle \eta,\xi \rangle = Tr\left(\eta^T \xi\right), \quad \forall \eta,\xi \in Sym(n)$$

Application:

$$\psi_\Omega(\beta) = \int\limits_{\Omega^*} e^{-\langle \beta,\xi \rangle} d\xi = \det(\beta)^{-\frac{n+1}{2}} \psi_\Omega(I_d)$$

$$\hat{\xi} = \frac{\partial \Phi(\beta)}{\partial \beta} = \frac{\partial(-\log \psi_\Omega(\beta))}{\partial \beta} = \frac{n+1}{2}\beta^{-1}$$

$$p_{\hat{\xi}}(\xi) = e^{-\left\langle \Theta^{-1}(\hat{\xi}),\xi \right\rangle + \Phi\left(\Theta^{-1}(\hat{\xi})\right)} = \psi_\Omega(I_d).\left[\det\left(\alpha\hat{\xi}^{-1}\right)\right] e^{-Tr\left(\alpha\hat{\xi}^{-1}\xi\right)} \quad \text{with} \quad \alpha = \frac{n+1}{2}$$

Fisher Metric and Euler-Lagrange Equation

Fisher Metric for Multivariate Gaussian Law

$$ds^2 = \sum_{ij} g_{ij} d\theta_i d\theta_j = dm^T R^{-1} dm + \frac{1}{2} Tr\left[\left(R^{-1}dR\right)^2\right]$$

Classical Euler-Lagrange equation

$$\sum_{i=1}^{n} g_{ik}\ddot{\theta}_i + \sum_{i,j=1}^{n} \Gamma_{ijk}\dot{\theta}_i\dot{\theta}_j = 0 \quad , \quad k = 1,\ldots,n$$

$$\text{with} \quad \Gamma_{ijk} = \frac{1}{2}\left[\frac{\partial g_{jk}}{\partial \theta_i} + \frac{\partial g_{jk}}{\partial \theta_j} + \frac{\partial g_{ij}}{\partial \theta_k}\right]$$

$$\Rightarrow \begin{cases} \ddot{R} + \dot{m}\dot{m}^T - \dot{R}R^{-1}\dot{R} = 0 \\ \ddot{m} - \dot{R}R^{-1}\dot{m} = 0 \end{cases}$$

We cannot easily integrate this Euler-Lagrange Equation (we will see that Lie group Theory will provide new equation: Euler-Poincaré equation)

Souriau Chapter on Statistical Physics: Multivariate Gaussian Law

Exemple : (*loi normale*) :

Prenons le cas $V = R^n$, λ = 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

$$Z(\Psi(x)) \equiv \overline{a}.x + \tfrac{1}{2}\,\overline{x}.H.x$$

[$a \in R^n$; 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

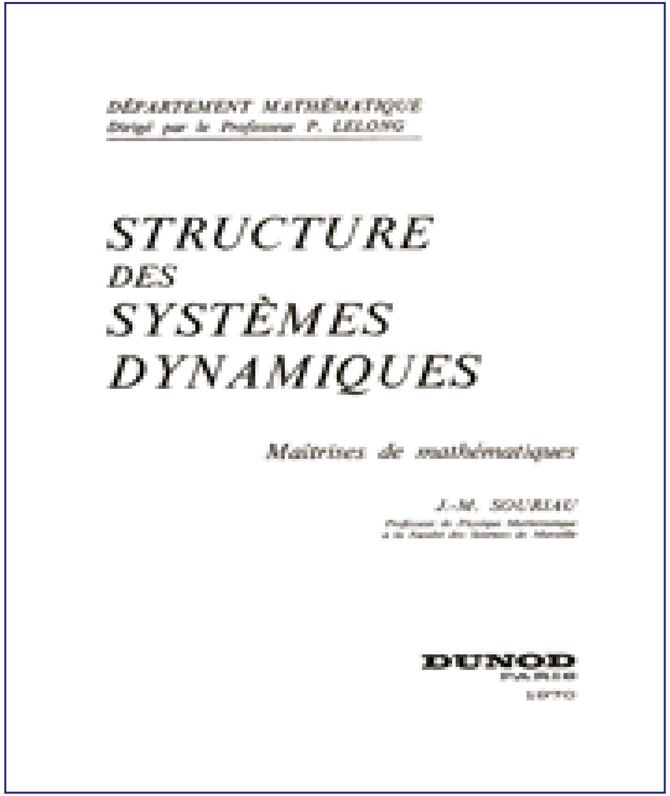
$$\boxed{z = \tfrac{1}{2}\,[\overline{a}.H^{-1}.a - \log(\det(H)) + n\log(2\pi)]}$$

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 $-H^{-1}.a$ 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* :

$$\boxed{s = \frac{n}{2}\log(2\pi e) - \frac{1}{2}\log(\det(H))} \;;$$

(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. LELONG

STRUCTURE DES SYSTÈMES DYNAMIQUES

Maîtrises de mathématiques

J.-M. SOURIAU

DUNOD
PARIS

http://www.jmsouriau.com/structure_des_systemes_dynamiques.htm

Maximum Entropy / Gibbs Density for Multivariate Gaussian Law

- if we take vector with tensor components $\xi = \begin{pmatrix} z \\ z \otimes z \end{pmatrix}$, components of $\bar{\xi}$ will provide moments of 1st and 2nd order of the density of probability $p_{\bar{\xi}}(\xi)$, that is defined by Gaussian law. In this particular case, we can write:

$$\langle \xi, x \rangle = a^T z + \frac{1}{2} z^T H z \quad \text{with} \quad a \in R^n \quad \text{and} \quad H \in Sym(n)$$

- By change of variable given by $z' = H^{1/2} z + H^{-1/2} a$, we can then compute the logarithm of the Koszul characteristic function:

$$\Phi(x) = -\frac{1}{2}\left[a^T H^{-1} a + \log\det\left[H^{-1}\right] + n\log(2\pi)\right]$$

- We can prove that 1st moment is equal to $-H^{-1}a$ and that components of variance tensor are equal to elements of matrix H^{-1}, that induces the 2nd moment. The Koszul Entropy, its Legendre transform, is then given by:

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

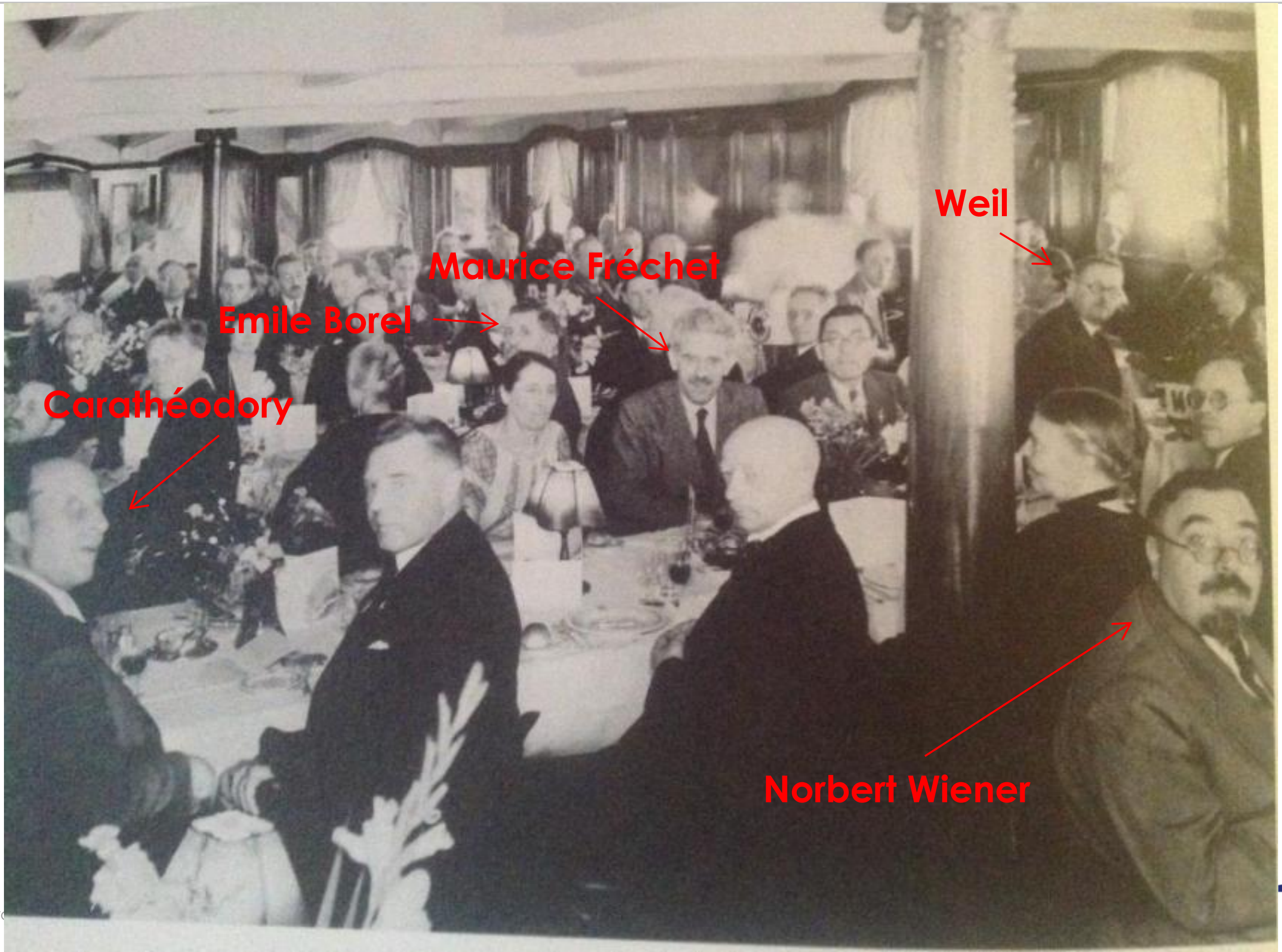
Fréchet seminal work on Cramer-Rao Bound 6 years before Rao and discovery of Clairaut(-Legendre) Equation of Information Geometry

OPEN

Page Facebook de Maurice Fréchet (création personnelle)

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

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



THALES

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.

THALES

Maurice Fréchet et André Blanc-Lapierre
Colloque international de Calcul des Probabilités - Lyon 1948

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)

Thèse de 1945 d' **André Blanc-Lapierre**

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

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.

THALES

Maurice Fréchet : IHP Lecture 1939, Paper 1943

- 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 <u>published first by Maurice Fréchet in 1939 in his winter "Mathematical Statistics" Lecture at the Institut Henri Poincaré during winter 1939–1940</u>. 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 <u>our lecture of mathematical statistics at the Henri Poincaré institute during winter 1939–1940</u>. 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, Tournois et Constans. Paris".

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:

$$(\sigma_T)^2 \geqslant \frac{1}{n(\sigma_A)^2} \text{ avec } T = H(X_1, \ldots, X_n), \quad A = \frac{1}{f(X,\theta)} \frac{\partial f(X,\theta)}{\partial \theta}$$

$$\hat{\theta} \text{ estimateur de } \theta \text{, borne de Fréchet} : R_\theta = E\left[\left(\theta - \hat{\theta}\right)\left(\theta - \hat{\theta}\right)^T\right] \geq I(\theta)^{-1}$$

$$[I(\theta)]_{i,j} = -E\left[\frac{\partial^2 \log P(Z/\theta)}{\partial \theta_i \partial \theta_j}\right] = E\left[\frac{\partial \log P(Z/\theta)}{\partial \theta_i} \frac{\partial \log P(Z/\theta)}{\partial \theta_j}\right]$$

- In his1943 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**

$$\Phi^*(x^*) = \left\langle x^*, \frac{d\Phi^*(x^*)}{dx^*} \right\rangle - \Phi\left(\frac{d\Phi^*(x^*)}{dx^*}\right) \text{ et } x^* = \frac{d\Phi(x)}{dx}$$

(55) $$\boxed{\mu = \theta\, \mu' - \psi(\mu')}$$

c'est-à-dire une équation de Clairaut. La solution $\mu' =$ 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

THALES

Clairaut Equation, 1734

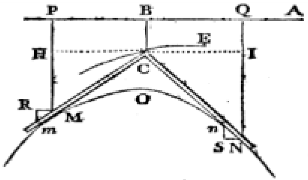
1724, Alexis Claude Clairaut

- Clairaut introduced before Legendre an equation related to Legendre transform.
- Legendre transform

$$\Phi^*(x^*) = \langle x^*, x \rangle - \Phi(x) \quad \text{avec} \quad \frac{d\Phi^*(x^*)}{dx^*} = x \ \text{ et } \ \frac{d\Phi(x)}{dx} = x^*$$

- **Clairaut Equation**

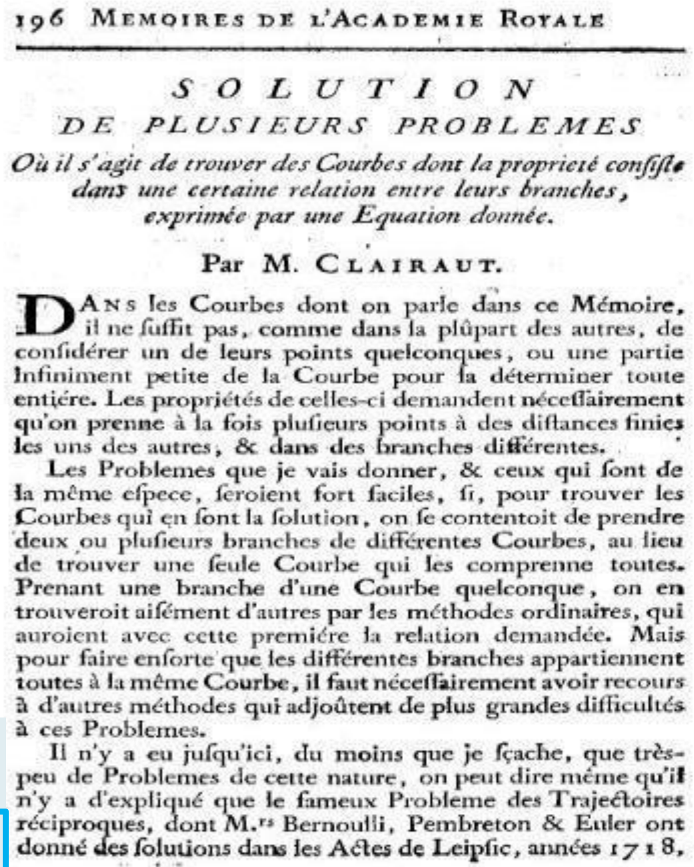
$$\Phi^*(x^*) = \left\langle x^*, \frac{d\Phi^*(x^*)}{dx^*} \right\rangle - \Phi\left(\frac{d\Phi^*(x^*)}{dx^*} \right)$$



- Singular solution: envelolp of straigth lines solutions

$$\frac{d\Phi^*(x^*)}{dx^*} = \frac{d\Phi^*(x^*)}{dx^*} + \left\langle x^*, \frac{d^2\Phi^*(x^*)}{dx^{*2}} \right\rangle - \Phi'\left(\frac{d\Phi^*(x^*)}{dx^*} \right) \frac{d^2\Phi^*(x^*)}{dx^{*2}}$$

$$\Rightarrow 0 = \left\langle x^* - \Phi'\left(\frac{d\Phi^*(x^*)}{dx^*} \right), \frac{d^2\Phi^*(x^*)}{dx^{*2}} \right\rangle \Rightarrow \boxed{x^* = \Phi'\left(\frac{d\Phi^*(x^*)}{dx^*} \right) \ \text{ et } \ \Phi(x^*) = \langle C, x^* \rangle + \Phi(C)}$$



196 MEMOIRES DE L'ACADEMIE ROYALE

SOLUTION DE PLUSIEURS PROBLEMES

Où il s'agit de trouver des Courbes dont la proprieté 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 plûpart des autres, de conſidérer un de leurs points quelconques, ou une partie Infiniment petite de la Courbe pour la déterminer toute entiére. 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 eſpece, feroient fort faciles, ſi, pour trouver les Courbes qui en ſont 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 premiére la relation demandée. Mais pour faire enſorte que les différentes branches appartiennent toutes à la même Courbe, il faut néceſſairement avoir recours à d'autres méthodes qui adjoûtent 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, Pembreton & Euler ont donné des ſolutions dans les Actes de Leipſic, années 1718,

THALES

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) $$\theta + \frac{\dfrac{\partial\, L f(x,\theta)}{\partial\theta}}{\displaystyle\int_{-\infty}^{+\infty}\left[\frac{\partial}{\partial\theta} f(x,\theta)\right]^2 \frac{dx}{f(x,\theta)}}$$

soit indépendante de θ.

Pour ces densités distinguées, on va pouvoir déterminer la fonction minimisante $H'(X_1,\dots X_n)$ et étendre au cas des petits échantillons la comparaison des méthodes 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) $$\lambda(\theta)\frac{\partial}{\partial\theta} L f(x,\theta) = h(x) - \theta$$

où $\lambda(\theta) > 0$. On peut considérer $\dfrac{1}{\lambda(\theta)}$ comme la dérivée seconde d'une fonction $\mu(\theta)$; d'où $\dfrac{\partial}{\partial\theta} L f(x,\theta) = \mu_\Theta''(\theta)\,[h(x)-\theta]$.

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) $$f(x,\theta) = e^{\mu'_\Theta[h(x)-\Theta]+\mu(\Theta)+l(x)}$$

(52bis) $$\lambda\,\mu'' = 1.$$

Incidemment, puisque, d'après (52), $\lambda(\theta)$ est positif, il en résulte aussi que $\mu''\left(=\dfrac{1}{\lambda(\theta)}\right)$ 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* $l(x)$ [1]) *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

$$\int_{-\infty}^{+\infty} e^{s h(x)+l(x)}\,dx \qquad ^{1})$$

sera une fonction positive connue que nous pourrons représenter par $e^{\Psi(s)}$. On voit alors que $\mu(\theta)$ sera défini par

$$\theta\,\mu' - \mu = \psi(\mu') \quad \text{ou}$$

(55) $$\mu = \theta\,\mu' - \psi(\mu')$$ **Legendre-Clairaut**

c'est-à-dire une équation de Clairaut. La solution $\mu' =$ 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

$$e^{\Theta s-\mu} = \int_{-\infty}^{+\infty} e^{s h(x)+l(x)}\,dx \text{ et}$$

(55bis) $$\int_{-\infty}^{+\infty} e^{s h(x)+l(x)}\,[h(x)-\theta]\,dx = 0.$$

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: $T = H(X_1, ..., X_n)$
- And the random variable $A(X) = \dfrac{\partial \log p_\theta(X)}{\partial \theta}$ with associated value $U = \sum_i A(X_i)$
- The constraint $\int_{-\infty}^{+\infty} p_\theta(x)dx = 1$ induces:

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \prod_i p_\theta(x_i)dx_i = 1 \underset{\substack{\textit{dérivée} \\ \textit{par } \theta}}{\Rightarrow} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \left[\sum_i A(x_i) \right] \prod_i p_\theta(x_i)dx_i = 0 \Rightarrow E_\theta[U] = 0$$

- If we assume $E_\theta[T] = \theta \Rightarrow \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} H(x_1, ..., x_n) \prod_i p_\theta(x_i)dx_i = \theta \underset{\substack{\textit{dérivée} \\ \textit{par } \theta}}{\Rightarrow} E[(T-\theta)U] = 1$
- Then as $E[T] = \theta$ and $E[U] = 0$, we have : $E[(T - E[T])(U - E[U])] = 1$
- According to Schwarz inequality: $[E(ZT)]^2 \leq E[Z^2]E[T^2]$

$$1 \leq E\left[(T - E[T])^2\right] E\left[(U - E[U])^2\right] = (\sigma_T \sigma_U)^2$$

- *U* as sum of independant variables, Bienaymé equality gives :

$$(\sigma_U)^2 = \sum_i \left[\sigma_{A(X_i)}\right]^2 = n(\sigma_A)^2 \Rightarrow \boxed{(\sigma_T)^2 \geq \frac{1}{n(\sigma_A)^2}}$$

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 λ' 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 independant of θ. Indeed, if we consider $E_{\theta_0}\left[H'\right]=\theta_0$, $E\left[\left(H'-\theta_0\right)^2\right]\leq E_{\theta_0}\left[\left(H-\theta_0\right)^2\right]\ \forall H\ \text{tq}\ E_{\theta_0}\left[H\right]=\theta_0$
- However $H=\theta_0$ verifyes 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 tis step, Fréchet studied « **densités distinguées** » (distinguished density), all density of probability $p_\theta(x)$ such that the following expression is independant of : θ

$$h(x)=\theta+\frac{\dfrac{\partial \log p_\theta(x)}{\partial\theta}}{\displaystyle\int_{-\infty}^{+\infty}\left[\frac{\partial p_\theta(x)}{\partial\theta}\right]^2\frac{dx}{p_\theta(x)}}$$

- Fréchet objective is to determine the minimizing function $T=H'\left(X_1,...,X_n\right)$ 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 $\dfrac{1}{\lambda(\theta)}$ as secund derivative of the function $\Phi(\theta)$ such that:

$$\frac{\partial \log p_\theta(x)}{\partial \theta} = \frac{\partial^2 \Phi(\theta)}{\partial \theta^2}\left[h(x)-\theta\right] \qquad \frac{1}{\lambda(x)} = \int_{-\infty}^{+\infty} p_\theta(x)\left[\frac{\partial \ln p_\theta(x)}{\partial \theta}\right]^2 dx = E\left[\left[\frac{\partial \ln p_\theta(x)}{\partial \theta}\right]^2\right] = \frac{\partial^2 \Phi(\theta)}{\partial \theta^2}$$

- we deduce that the following function is independant of θ :

$$\ell(x) = \log p_\theta(x) - \frac{\partial \Phi(\theta)}{\partial \theta}\left[h(x)-\theta\right] - \Phi(\theta)$$

- A **distinguished density** could be written:

$$p_\theta(x) = e^{\frac{\partial \Phi(\theta)}{\partial \theta}\left[h(x)-\theta\right] + \Phi(\theta) + \ell(x)} \quad \text{with} \quad \int_{-\infty}^{+\infty} p_\theta(x)dx = 1$$

- These 2 conditions are enough

Fréchet proved that « distinguished densities » are in the familly 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:

$$\theta + \frac{\dfrac{\partial \log p_\theta(x)}{\partial \theta}}{\int_{-\infty}^{+\infty}\left[\dfrac{\partial p_\theta(x)}{\partial \theta}\right]^2 \dfrac{dx}{p_\theta(x)}} = \theta + \lambda(x)\frac{\partial^2 \Phi(\theta)}{\partial \theta^2}[h(x)-\theta]$$

- if $\lambda(x)\dfrac{\partial^2 \Phi(\theta)}{\partial \theta^2} = 1$ when $\dfrac{1}{\lambda(x)} = \int_{-\infty}^{+\infty}\left[\dfrac{\partial \log p_\theta(x)}{\partial \theta}\right]^2 p_\theta(x)dx = (\sigma_A)^2$ 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 divise 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 : $\lambda(x).\frac{\partial^2\Phi(\theta)}{\partial\theta^2}=1$
- 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:

$$\int_{-\infty}^{+\infty} e^{\frac{\partial \Phi(\theta)}{\partial \theta}[h(x)-\theta]+\Phi(\theta)+\ell(x)} dx = 1$$

- That we can written: $e^{\theta.\frac{\partial \Phi(\theta)}{\partial \theta}-\Phi(\theta)} = \int_{-\infty}^{+\infty} e^{\frac{\partial \Phi(\theta)}{\partial \theta} h(x)+\ell(x)} dx$

- If we slect arbitrary $h(x)$ and $l(x)$, and let s an arbitrary variable, the following function will be a positive function known, given by : $e^{\Psi(s)}$

$$\int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)} dx = e^{\Psi(s)}$$

- We obtain $\Phi(\theta)$ by Clairaut Equation: $\Phi(\theta) = \theta.\frac{\partial \Phi(\theta)}{\partial \theta} - \Psi\left(\frac{\partial \Phi(\theta)}{\partial \theta}\right)$

- 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: $\Phi = \theta.s - \Psi(s)$ 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:

$$e^{\theta.s-\Phi(\theta)} = \int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)} dx \qquad \int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)} \left[h(x)-\theta\right] dx = 0$$

- We obtain $\Phi(\theta) = -\log \int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)} dx + \theta.s$ where s is given by implicite equation:

$$\int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)} \left[h(x)-\theta\right] dx = 0$$

- Fréchet also obersved that:

$$\left(\sigma_{T_n}\right)^2 = \frac{1}{n\left(\sigma_A\right)^2} = \frac{1}{n\int_{-\infty}^{+\infty}\left[\frac{\partial p_\theta(x)}{\partial \theta}\right]^2 \frac{dx}{p_\theta(x)}} = \frac{1}{n\frac{\partial^2 \Phi(\theta)}{\partial \theta^2}}$$

- T_n follows:

$$p_\theta(t) = \sqrt{n}\frac{1}{\sigma_A\sqrt{2\pi}} e^{-\frac{n(t-\theta)^2}{2.\sigma_A^2}} \text{ with } \left(\sigma_A\right)^2 = \frac{\partial^2 \Phi(\theta)}{\partial \theta^2}$$

OPEN

THALES

Fréchet 1943 paper: Clairaut Equation

- We can write the estimator as : $H'(X_1,\dots,X_n)=\frac{1}{n}\left[h(X_1)+\dots+h(X_n)\right]$
- 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 t of θ corresponds at a sample $x_1,\dots,x_n$ that is a root of previous equation in t. This euqtaion 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: $t=\frac{1}{n}\sum_i h(x_i)$
- $T\equiv H'(X_1,\dots,X_n)=\frac{1}{n}\sum_i h(X_i)$ cannot have an arbitrary expression, but it is given by a sum.

THALES

Jean-Louis Koszul work on homogeneous bounded domains: Koszul-Vinberg Characteristic Function & Koszul Forms

OPEN

Fondation of Information Geometry by Jean-Louis Koszul

Fisher Metric = Hassian Metric

- In case of Maximum Entropy Density (Gibbs density), Riemannian metric associated to hessian of logarithm of the partition function is equal to Fisher Metric.
- These Hessian Geometrical Structures have been studied in parallel par the mathemacian **Jean-Louis Koszul** and his PhD student **Jacques Vey** in more general framework of sharp convex cones:
  - Koszul Forms
  - Koszul-Vinberg Characteristic Function
- In Jean-Louis Koszul model , fondamental structures are deduced of affine represnetation of Liegroup and Lie Algebra (this affine representation is also at the heat of Jean-Marie Souriau model).

ENS Promo 40

Jean-Louis Koszul and Hirohiko Shima

Jean-Louis Koszul, Correspondant of the French Academy of Sciences, PhD student of Henri Cartan, Bourbaki member

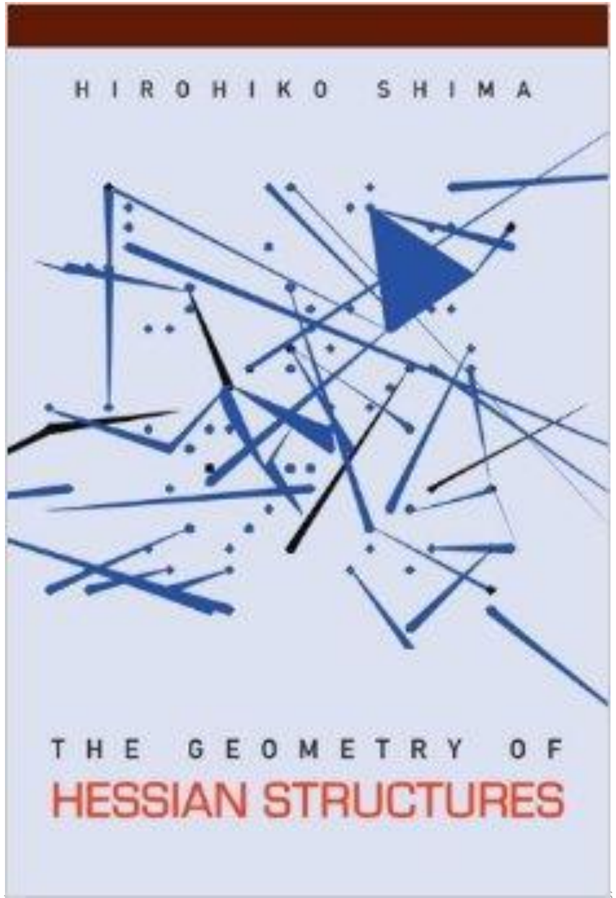
- Koszul, J.L. Sur la forme hermitienne canonique des espaces homogènes complexes. Can. J. Math. 1955, 7, 562–576.
- Koszul, J.L. Exposés sur les Espaces Homogènes Symétriques; Publicação da Sociedade de Matematica de São Paulo: São Paulo, Brazil, 1959.
- Koszul, J.L. Domaines bornées homogènes et orbites de groupes de transformations affines. Bull. Soc. Math. Fr. 1961, 89, 515–533.
- Koszul, J.L. Ouverts convexes homogènes des espaces affines. Math. Z. 1962, 79, 254–259.
- Koszul, J.L. Variétés localement plates et convexité. Osaka. J. Math. 1965, 2, 285–290.
- Koszul, J.L. Déformations des variétés localement plates. Ann. Inst. Fourier 1968, 18, 103–114.
- Koszul, J.L. Trajectoires Convexes de Groupes Affines Unimodulaires. In Essays on Topology and Related Topics; Springer: Berlin, Germany, 1970; pp. 105–110.
- Vey, J. Sur les Automorphismes Affines des Ouverts Convexes Saillants. Annali della Scuola Normale Superiore di Pisa, Classe di Science 1970, 24, 641–665.

Hirohiko Shima, Emeritus Professor of Yamaguchi university, Phd Student at Osaka University

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

Prof. Michel Nguiffo Boyom tutorial : Géométrie de l'Information
http://repmus.ircam.fr/_media/brillouin/ressources/une-source-de-nouveaux-invariants-de-la-geometrie-de-l-information.pdf

Jean-Louis Koszul and Hirohiko Shima



Jean-Louis Koszul and Hirohiko Shima
2013, Ecole des Mines de Paris
GSI'13 « Geometric Science of Information »

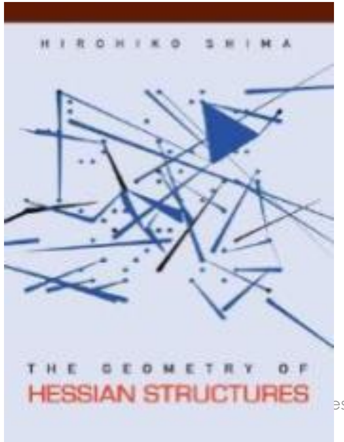
THALES

Jean-Louis Koszul and hessian structures

Hessian Geometry of Jean-Louis Koszul

- Hirohiko Shima Book, « Geometry of Hessian Structures », world Scientific Publishing 2007, dedicated to Jean-Louis Koszul
- Hirohiko Shima Keynote Talk at GSI'13
  - http://www.see.asso.fr/file/5104/download/9914
- Prof. M. Boyom tutorial :
  - http://repmus.ircam.fr/_media/brillouin/ressources/une-source-de-nouveaux-invariants-de-la-geometrie-de-l-information.pdf

Jean-Louis Koszul



J.L. Koszul, « Sur la forme hermitienne canonique des espaces homogènes complexes », Canad. J. Math. 7, pp. 562-576., 1955

J.L. Koszul, « Domaines bornées homogènes et orbites de groupes de transformations affines », Bull. Soc. Math. France 89, pp. 515-533., 1961

J.L. Koszul, « Ouverts convexes homogènes des espaces affines », Math. Z. 79, pp. 254-259., 1962

J.L. Koszul, « Variétés localement plates et convexité », Osaka J. Maht. 2, pp. 285-290., 1965

J.L. Koszul, « Déformations des variétés localement plates », .Ann Inst Fourier, 18 , 103-114., 1968

Study of Koszul Works by Michel Nguiffo Boyom

Michel Nguiffo Boyom (Institut Montpelierain Alexander Grothendieck)

- Bridges between differential topology (Feuilletages), and Fisher metric of Information Geometry have their pillars in **KV (Koszul-Vinberg) cohomology of locally flat manifolds and in Koszul convexity**
- Connexions are objects of Information Geometry. They are **deformations in space of linear connexions of Koszul connexion that control the geometry of statistical model**. These connexions are parameterized by **2-cochaines of KV(Koszul-Vinberg) complexe**. Fisher connexion is one of them. Curvature of Linear Connexion is a value of Maurer-Cartan Polynomial. One raison in favor of this polynomial is that it **controles all locally flat structures** via anomalies theory. Nijenhuis, Koszul, Gerstenhaber, Vinberg & soviet school have initiated **study of the connexion locally flat**.
- Notion of **convexity of locally flat manifolds** is introduced by Jean-Louis Koszul. Considerations that drive this approach find their roots in **Geometry of bounded homogeneous manafolds**. Convexity notion introduced by Koszul is the real analogue of holomorph convexity of Kaup
- References:

  - M. Nguiffo Boyom and R. Wolak. ransversely Hessian Foliations and Information Geometry: International Journal of Mathematics, Vol 27, (11)(2017)
  - M. Nguiffo Boyom. Foliations-Webs-Hessian-Geometry- Information Geometry -Entropy and Cohomology . Entropy 12 No 18, 433 (2016).
  - M. Nguiffo Boyom. Numerical properties of Koszul connections , arXiv.1708.01108. 3 august 2017

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**

THALES

Koszul-Vinberg Characteristic Function/Metric of convex cone

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

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

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

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

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

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

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

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

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

Koszul Entropy via Legendre Transform

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

$$\Phi^*(x^*) = \left\langle x, x^* \right\rangle - \Phi(x)$$

with $x^* = D_x \Phi$ and $x = D_{x^*} \Phi^*$ where $\psi_\Omega(x) = \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \quad \forall x \in \Omega$

- **Demonstration:** we set $x^* = \int_{\Omega^*} \xi . e^{-\langle \xi, x \rangle} d\xi \, / \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$

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

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

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

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

Koszul-Vinberg Characteristic Function Legendre Transform

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

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

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

$$\Phi^*(x^*)=\left[\log\int_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi.\left(\int_{\Omega^*}\frac{e^{-\langle\xi,x\rangle}}{\int_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi}d\xi\right)-\int_{\Omega^*}\log e^{-\langle\xi,x\rangle}.\frac{e^{-\langle\xi,x\rangle}}{\int_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi}d\xi\right]\text{ with }\int_{\Omega^*}\frac{e^{-\langle\xi,x\rangle}}{\int_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi}d\xi=1$$

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

SHANNON ENTROPY

Koszul metric & Fisher Metric

- 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:

$$\log p_x(\xi) = -\Phi^*(\xi) = \Phi(x) - \langle x, \xi \rangle$$

$$\frac{\partial^2 \log p_x(\xi)}{\partial x^2} = \frac{\partial^2 \left[\Phi(x) - \langle x, \xi \rangle\right]}{\partial x^2} = \frac{\partial^2 \Phi(x)}{\partial x^2}$$

$$\Rightarrow I(x) = -E_\xi\left[\frac{\partial^2 \log p_x(\xi)}{\partial x^2}\right] = -\frac{\partial^2 \Phi(x)}{\partial x^2} = \frac{\partial^2 \log \psi_\Omega(x)}{\partial x^2}$$

- We could then deduce the close interrelation between Fisher metric and hessian of Koszul-Vinberg characteristic logarithm.

$$\boxed{I(x) = -E_\xi\left[\frac{\partial^2 \log p_x(\xi)}{\partial x^2}\right] = \frac{\partial^2 \log \psi_\Omega(x)}{\partial x^2}}$$

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 :

$$\log \Psi_{\Omega}(x) = \log \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \Rightarrow \frac{\partial \log \Psi_{\Omega}(x)}{\partial x} = -\frac{1}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} \int_{\Omega^*} \xi . e^{-\langle \xi, x \rangle} d\xi$$

$$\frac{\partial^2 \log \Psi_{\Omega}(x)}{\partial x^2} = -\frac{1}{\left(\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \right)^2} \left[- \int_{\Omega^*} \xi^2 . e^{-\langle \xi, x \rangle} d\xi . \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi + \left(\int_{\Omega^*} \xi . e^{-\langle \xi, x \rangle} d\xi \right)^2 \right]$$

$$\frac{\partial^2 \log \Psi_{\Omega}(x)}{\partial x^2} = \int_{\Omega^*} \xi^2 . \frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} d\xi - \left(\int_{\Omega^*} \xi . \frac{e^{-\langle \xi, x \rangle}}{\int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} d\xi \right)^2 = \int_{\Omega^*} \xi^2 . p_x(\xi) d\xi - \left(\int_{\Omega^*} \xi . p_x(\xi) d\xi \right)^2$$

$$I(x) = -E_{\xi}\left[\frac{\partial^2 \log p_x(\xi)}{\partial x^2} \right] = \frac{\partial^2 \log \psi_{\Omega}(x)}{\partial x^2} = E_{\xi}\left[\xi^2\right] - E_{\xi}\left[\xi\right]^2 = Var(\xi)$$

Definition of Maximum Entropy Density

- How to replace x by mean value of ξ , $\bar{\xi}(=x^*)$ in :

$$p_x(\xi)=\frac{e^{-\langle \xi,x\rangle}}{\int\limits_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi} \quad \text{with} \quad \bar{\xi}=\int\limits_{\Omega^*}\xi.p_x(\xi)d\xi$$

- Legendre Transform will do this inversion by inversing $\bar{\xi}=\frac{d\Phi(x)}{dx}$

- We then observe that Koszul Entropy provides density of Maximum Entropy with this general definition of density:

$$p_{\bar{\xi}}(\xi)=\frac{e^{-\left\langle \xi,\Theta^{-1}(\bar{\xi})\right\rangle}}{\int\limits_{\Omega^*} e^{-\left\langle \xi,\Theta^{-1}(\bar{\xi})\right\rangle}d\xi}$$

$$x=\Theta^{-1}(\bar{\xi})$$

$$\bar{\xi}=\Theta(x)=\frac{d\Phi(x)}{dx}$$

where $\bar{\xi}=\int\limits_{\Omega^*}\xi.p_{\bar{\xi}}(\xi)d\xi$ and $\Phi(x)=-\log\int\limits_{\Omega^*} e^{-\langle x,\xi\rangle}d\xi$

THALES

Cartan-Killing Form and Invariant Inner Product

- It is not possible to define an ad(g)-invariant inner product for any two elements of a Lie Algebra, but a symmetric bilinear form, called "**Cartan-Killing form**", could be introduced (Elie Cartan PhD 1894)

- This form is defined according to the adjoint endomorphism ad_x of $\mathfrak{g}$ that is defined for every element x of $\mathfrak{g}$ with the help of the Lie bracket: $ad_x(y) = [x, y]$

- The trace of the composition of two such endomorphisms defines a bilinear form, the **Cartan-Killing form**: $B(x, y) = Tr(ad_x ad_y)$

- The Cartan-Killing form is symmetric: $B(x, y) = B(y, x)$

- and has the associativity property: $B([x, y], z) = B(x, [y, z])$

- given by:
$$B([x, y], z) = Tr(ad_{[x,y]} ad_z) = Tr([ad_x, ad_y] ad_z)$$
$$B([x, y], z) = Tr(ad_x [ad_y, ad_z]) = B(x, [y, z])$$

Cartan-Killing Form and Invariant Inner Product

- Elie Cartan has proved that if $\mathfrak{g}$ is a simple Lie algebra (the Killing form is non-degenerate) then **any invariant symmetric bilinear form on $\mathfrak{g}$ is a scalar multiple of the Cartan-Killing form.**

- The Cartan-Killing form is invariant under automorphisms $\sigma \in Aut(\mathfrak{g})$ of the algebra $\mathfrak{g}$: $\boxed{B(\sigma(x),\sigma(y)) = B(x,y)}$

- To prove this invariance, we have to consider:

$$\begin{cases} \sigma[x,y] = [\sigma(x),\sigma(y)] \\ z = \sigma(y) \end{cases} \Rightarrow \sigma[x,\sigma^{-1}(z)] = [\sigma(x),z]$$

$$\text{rewritten} \quad ad_{\sigma(x)} = \sigma \circ ad_x \circ \sigma^{-1}$$

$$B(\sigma(x),\sigma(y)) = Tr\left(ad_{\sigma(x)} ad_{\sigma(y)}\right) = Tr\left(\sigma \circ ad_x ad_y \circ \sigma^{-1}\right)$$

$$B(\sigma(x),\sigma(y)) = Tr\left(ad_x ad_y\right) = B(x,y)$$

Cartan-Killing Form and Invariant Inner Product

A natural G-invariant inner product could be introduced by Cartan-Killing form:

- Cartan Generating Inner Product: The following Inner product defined by Cartan-Killing form is invariant by automorphisms of the algebra

$$\langle x, y \rangle = -B(x, \theta(y))$$

where $\theta \in \mathfrak{g}$ is a Cartan involution (An involution on $\mathfrak{g}$ is a Lie algebra automorphism θ of $\mathfrak{g}$ whose square is equal to the identity).

From Cartan-Killing Form to Koszul Information Metric

$$B(x,y)=Tr\left(ad_x ad_y\right)$$

$$\text{Cartan}-\text{Killing Form}$$

$$\langle x,y\rangle=-B(x,\theta(y))$$

$$\text{with } \theta\in g \text{ , Cartan Involution}$$

$$\text{Koszul Characteristic Function}$$

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

$$\text{Koszul Entropy}$$

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

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

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

$$\text{Koszul Density}$$

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

$$\text{Koszul Metric}$$

$$I(x)=-E_\xi\left[\frac{\partial^2\log p_x(\xi)}{\partial x^2}\right]$$

$$I(x)==-\frac{\partial^2\Phi(x)}{\partial x^2}=\frac{\partial^2\log\int_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi}{\partial x^2}$$

OPEN

Relation of Koszul density with Maximum Entropy Principle

- The density from Maximum Entropy Principle is given by:

$$\underset{p_x(.)}{Max}\left[-\int_{\Omega^*} p_x(\xi)\log p_x(\xi)d\xi\right] \text{ such } \begin{cases} \int_{\Omega^*} p_x(\xi)d\xi = 1 \\ \int_{\Omega^*} \xi.p_x(\xi)d\xi = x^* \end{cases}$$

- If we take $q_x(\xi) = e^{-\langle \xi,x\rangle} / \int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi = e^{-\langle x,\xi\rangle - \log \int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi}$ such that

$$\begin{cases} \int_{\Omega^*} q_x(\xi).d\xi = \int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi / \int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi = 1 \\ \log q_x(\xi) = \log e^{-\langle x,\xi\rangle - \log \int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi} = -\langle x,\xi\rangle - \log \int_{\Omega^*} e^{-\langle x,\xi\rangle}d\xi \end{cases}$$

- Then by using the fact that $\log x \geq \left(1 - x^{-1}\right)$ with equality if and only if , $x = 1$, we find the following:

$$-\int_{\Omega^*} p_x(\xi)\log\frac{p_x(\xi)}{q_x(\xi)}d\xi \leq -\int_{\Omega^*} p_x(\xi)\left(1 - \frac{q_x(\xi)}{p_x(\xi)}\right)d\xi$$

THALES

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 $q_x(\xi)$:

$$-\int_{\Omega^*} p_x(\xi)\log p_x(\xi)d\xi \leq -\int_{\Omega^*} p_x(\xi)\left[-\langle x,\xi\rangle - \log\int_{\Omega^*} e^{-\langle x,\xi\rangle}d\xi\right]d\xi$$

$$-\int_{\Omega^*} p_x(\xi)\log p_x(\xi)d\xi \leq \left\langle x, \int_{\Omega^*} \xi.p_x(\xi)d\xi\right\rangle + \log\int_{\Omega^*} e^{-\langle x,\xi\rangle}d\xi$$

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

$$-\int_{\Omega^*} p_x(\xi)\log p_x(\xi)d\xi \leq \Phi^*(x^*)$$

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

Koszul Forms for Homogeneous Bounded domains

- Koszul has developed his previously described theory for Homogenous 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 :

$$\Psi(X) = \langle X, \xi\rangle = Tr[ad(JX) - J.ad(X)] \quad \forall X \in \mathfrak{g}$$

- 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:

$$\alpha = -\frac{1}{4} d\Psi(X) \qquad \beta = D\alpha$$

Koszul Forms/Metric for Homogenous Siegel Domains SD

Koszul Forms

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

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

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

Application for Poincaré Upper-Hal Plane:

$$V = \{z = x + iy \ / \ y > 0\} \qquad Y = y\frac{d}{dy} \quad \Rightarrow \begin{cases} ad(Y).Z = [Y, Z] \\ [X, Y] = -Y \\ JX = Y \end{cases}$$

With $X = y\frac{d}{dx}$ and $\begin{cases} Tr\left[ad(JX) - Jad(X)\right] = 2 \\ Tr\left[ad(JY) - Jad(Y)\right] = 0 \end{cases}$

We can deduce that

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

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

Jean-Louis Koszul Forms for Siegel Upper-Half Plane

Koszul form for Siegel Upper-Half Plane: $V=\{Z=X+iY \,/\, Y>0\}$

- Symplectic Group :

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

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

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

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

Carl Ludwig Siegel
(Siegel space of 1st king 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 homogenesous space with positive definite hermitian canonical form is isomorphe to a bounded domain, study of affine transformation groups of locally flat manifolds)

Henri Poincaré
(half-plane) n=1

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

Study of Symmetric Spaces and their classification

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

Euclidean Space: isometry

X

$B = G_{(A,B)}A$

$A \bullet B = \dfrac{(A+B)}{2}$

$A = G_{(A,B)}B$

$G_{(A,B)}X = (A \bullet B) - X + (A \bullet B)$

Courbure nulle

Metric space: isometry (e.g. space of Hermitian Positive Definite matrix)

$B = G_{(A,B)}A$

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

X

$G_{(A,B)}X = (A \circ B) X^{-1} (A \circ B)$

$A = G_{(A,B)}B$

Courbure négative

Euclidean space : geodesic

$\gamma(1) = B$

$\gamma(t) = A + t(B - A)$

$t \in [0,1]$

$\gamma(0) = A$

Metric Space: geodesic

$\gamma(1) = B$

$\gamma(t) = A^{1/2}\left(A^{-1/2} B A^{-1/2}\right)^{t} A^{1/2}$

$t \in [0,1]$

$\gamma(0) = A$

E. J. Cartan

M. Berger

THALES

SOURIAU MODEL OF INFORMATION GEOMETRY: Lie Group Thermodynamics

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 aicraft with respect to all positions of engine: Caravelle), supervised by André Lichnerowicz (Collège de France) & Joseph Pérès

ONERA
THE FRENCH AEROSPACE LAB

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

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

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

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

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

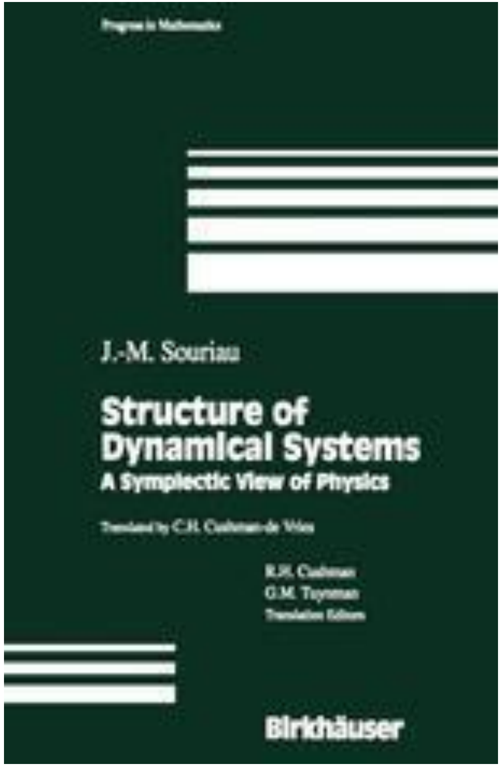
$$A_n = B_{n-1}A \quad \text{et} \quad k_n = -\frac{1}{n}\mathrm{tr}(A_n)$$

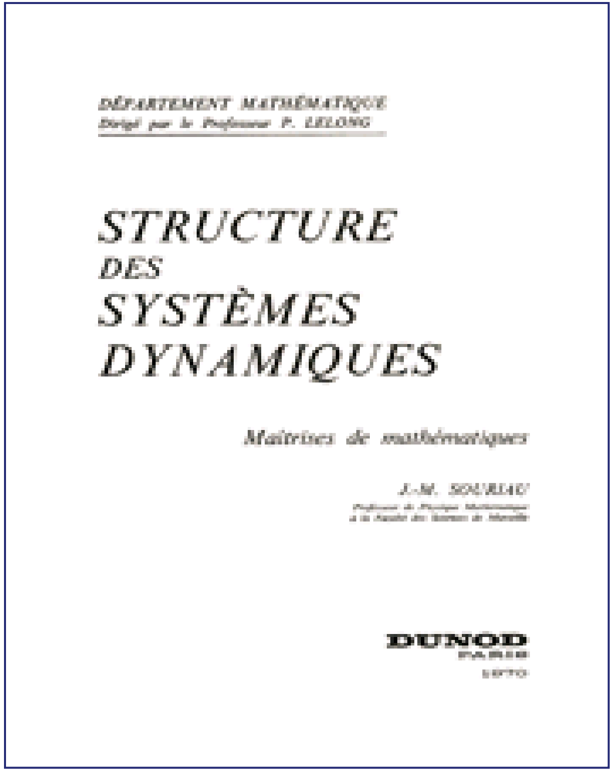
- Introduit de la géométrie Symplectique en Mécanique (relecture de Lagrange): **J.M. Souriau, Structure des systèmes dynamiques, Dunod, Paris, 1970**

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

THALES

Fundamental Book of Jean-Marie Souriau



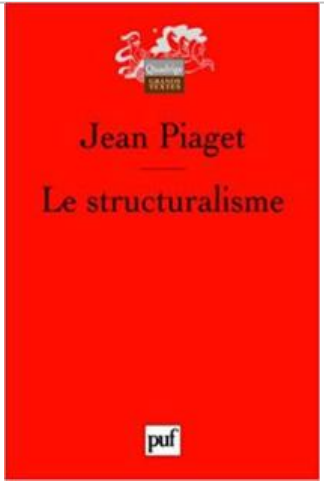


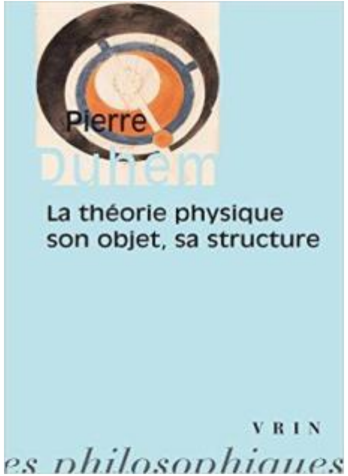
http://www.jmsouriau.com/structure_des_systemes_dynamiques.htm
http://www.springer.com/us/book/9780817636951

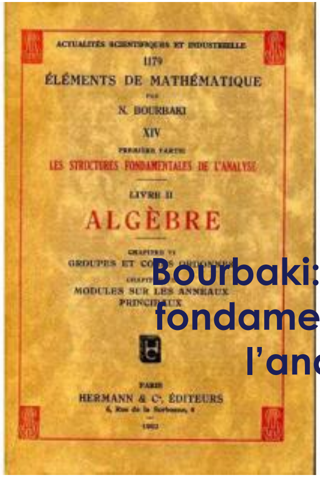
- Introduction of Symplectic Geometry in Mechanics
- Invention of Moment(um) application
- Geometrization of Noether theorem
- Barycentric Decomposition Theorem
- Total mass of an isolated dynamic system is the class of cohomology of the equivariance default of momentum application (for Gallilee group).
- Lie Group Thermodynamics (Chapter IV on Statistical Mechanics)

WHERE IS THE STRUCTURE ?

STRUCTURE:
from latin structura, de struere (to assemble).

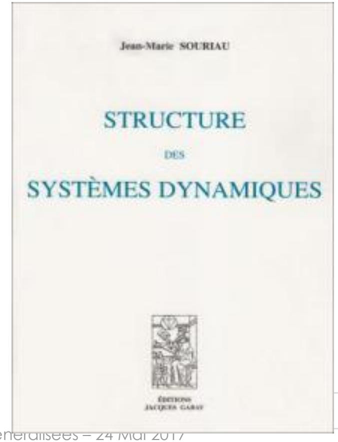


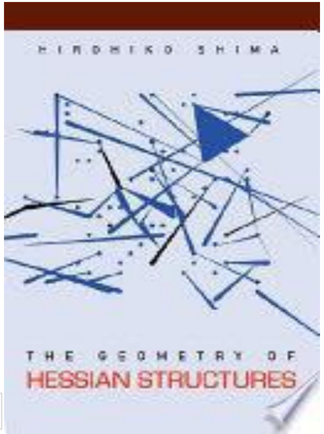




Bourbaki: structures fondamentales de l'analyse

J.M. Souriau





J.L. Koszul

Balian, R., Valentin P., Hamiltonian structure of thermodynamics with gauge, Eur. Phys. J. B 21, 2001, pp. 269-282.

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

THALES

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

THALES

Genesis of the Souriau book « Structures of Dynamical Systems »

- Jean-Marie Souriau, graduated from ENS ULM was the nephew of Etienne Souriau (Philosopher, collaborator of Gaston Bachelard in Paris Sorbonne University) and grandson of Paul Souriau (Philosopher) who both have worked on « aesthetic ».

- SSD book was elaborated in Carthage & Marseille, where Souriau was installed with his wife Christiane Souriau-Hoebrecht. In 1952 Souriau found a position at Institut des Hautes Études de Tunis (8 rue de Rome, Tunis, Tunisie.: http://www.persee.fr/doc/remmm_0035-1474_1985_num_39_1_2076) and was back in Marseille in a position in 1958 at Faculté des Sciences.

- The manuscript is given to the editor Dunod in 1969, but only edited in 1970

- About the title, we are at the apogee (« acmé » from greek ἀκμή) of the STRUCTURALISM in antropology / sociology / linguistic / philosophy / epistemology in France (Levi-Strauss , Barthes, Foucault, Althusser, Lacan,…). The word "structure" was in the air of time, fashionable at the moment, circulating on all the lips (François Dosse, "Histoire du structuralisme I & II").

- After his ONERA PhD Defence in 1953, his PhD supervisor André Lichnerowicz made one comment *« vous avez de nombreuses formes anti-symétriques dans vos calculs, vous devriez vous intéresser aux structures symplectiques »* (source: Sur la Symplectisation de la Physique; séminaire histoire des géométries 2013, https://www.youtube.com/watch?v=hdmBLD44G_Y)

THALES

Institut des Hautes Etudes de Tunis (J.M. Souriau: 1952 – 1958)

http://www.ina.fr/video/AFE01000164

Reference to François Gallisot in Souriau book

The Lagrange 2-form

Let us return to the evolution space V and let us define a priori

$$\sigma(\delta y)(\delta' y) = \sum_j \Big(\langle m_j\,\delta \mathbf{v}_j - \mathbf{F}_j\,\delta t, \delta' \mathbf{r}_j - \mathbf{v}_j\,\delta' t\rangle - \langle m_j\,\delta' \mathbf{v}_j - \mathbf{F}_j\,\delta' t, \delta \mathbf{r}_j - \mathbf{v}_j \delta t\rangle \Big)\,. \tag{12.40}$$

It is clear that σ is a *2-form* on V (definition (4.26)), and that the differential equation of motion $\delta y \in \mathcal{E}$ (notation (12.27)) implies

$$\sigma(\delta y)(\delta' y) = 0 \qquad \forall\, \delta' y\,, \tag{12.41}$$

which, using (4.2), can also be written as

$$\sigma(\delta y) = 0\,, \tag{12.42}$$

or as

$$\delta y \in \ker(\sigma)\,. \tag{12.43}$$

The 2-form σ was introduced by Lagrange in his study of celestial mechanics in 1808,[198] in a different language of course. Its explicit expression (12.40) can be found in Gallissot.[199]

[15] F. Gallissot, *Les formes extérieures en mécanique*, Ann. Inst. Fourier **4**, (1952), 145–297.

Souriau was influenced by François Gallissot work:

- François Gallissot, Les formes extérieures en mécanique, Annales de l'Institut Fourier, tome 4, p.145-297, 1952
- François Gallissot, les formes extérieures et la mécaniques des milieux continus, Annales de l'Institut Fourier, tome 8, p. 291-335, 1958

Souriau and Galissot both attended ICM'54 in Moscow ?
Did they discuss about 1952 paper ?

THALES

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

1952

LES FORMES EXTÉRIEURES EN MÉCANIQUE

par F. GALLISSOT.

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 Élie Cartan a montré que toutes les propriétés des équations différentielles de la dynamique des sytèmes 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 I l'étude des bases logiques sur lesquelles est édifiée la mécanique galiléenne. Je montre ainsi dans le § I que lorsqu'on se propose de trouver des formes génératrices des équations du mouvement d'un point matériel invariantes dans les transformations du groupe galiléen, la forme la plus intéressante est une forme extérieure de degré deux définie sur une variété $V_7 = E_3 \otimes E_3 \otimes T$ (E_3 espace euclidien, T droite numérique temporelle)(1). Dans le § II 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 $i(\ )$ antidérivation de M. H. Cartan (3), le champ caractéristique E de la forme Ω étant défini par la relation $i(E)\Omega = 0$.

(1) M. Kravtchenko 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.

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

Théorème I. — *Il existe trois types de formes différentielles génératrices des équations du mouvement d'un point matériel invariantes dans les transformations du groupe galiléen*

$$\text{A} \quad \begin{cases} s = \dfrac{1}{2m} \displaystyle\sum_1^3 (m dv^i - X^i dt)^2 \\ e = \dfrac{m}{2} \displaystyle\sum_1^3 (dx^j - v^j dt)^2 \end{cases}$$

B $f = \sum_1^3 \delta_{ij}(dx^i - v^i dt)(m dv^j - X^j dt)$ δ_{ij} symboles de Krönecker,

C $\omega = \sum_1^3 k_{ij}(m dv^i - X^i dt) \wedge (dx^j - v^j dt)$ k_{ij} symbole de Krönecker.

Si l'on revient à l'expression de ω.

$$\omega = \sum_1^3 k_{ij}(m\, dv^i - x^i\, dt) \wedge (dx^j - v^j\, dt)$$

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

$d\omega = 0$ impose à la forme de Pfaff $k_{ij} X^i dx^j$ d'être fermée, par suite de se réduire à la différentielle d'une fonction U, d'où

(I, 6) $$\omega = mk_{ij}\, dv^i \wedge dx^j - dH \wedge dt.$$

Avec $H = T - U$, $T = \frac{1}{2} \sum_1^3 m(v^i)^2$ demi-force vive, U fonction de force. (I, 6) montre que ω est la dérivée extérieure de

$$\omega^1 = \sum_{i=1}^3 mv^i\, dx^i - H\, dt.$$

La forme ω^1 engendre l'invariant intégral d'Elie Cartan (10); elle diffère de la forme $\overline{\omega}^1$ (I, 4) d'une forme fermée.

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 (Gallilee Group in classical Mechanic 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 Dynamic 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**

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

Homogeneous Gallileo group

$$\begin{cases} \vec{x}' = R.\vec{x} + \vec{u}.t + \vec{w} \\ t' = t + e \end{cases} \qquad \begin{bmatrix} \vec{x}' \\ t' \\ 1 \end{bmatrix} = \begin{bmatrix} \boxed{\begin{matrix} R & \vec{u} \\ 0 & 1 \end{matrix}} & \begin{matrix} \vec{w} \\ e \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 1 \end{bmatrix} \begin{bmatrix} \vec{x} \\ t \\ 1 \end{bmatrix}$$

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

$$R \in SO(3)$$

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

- Bargmann Central extension:

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

OPEN

THALES

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.

Gibbs Canonical Ensemble on Symplectic Manifold

- In classical statistical mechanics, a state is given by the solution of **Liouville equation** on the phase space, the partition function.

- As symplectic manifolds have a completely continuous measure, invariant by diffeomorphisms, the **Liouville measure** λ , all statistical states will be the product of Liouville measure by the scalar function given by the generalized density function $e^{\Phi(\beta)-\langle\beta,U(\xi)\rangle}$ defined by:
  - the energy U (defined in dual of Lie Algebra of the dynamic group)
  - the geometric temperature β
  - $\Phi(\beta)$ a normalizing constant such the mass of probability is equal to 1

- The Gibbs equilibrium state is extended to all Symplectic manifolds with a dynamic group. To ensure that all integrals could converge, the canonical Gibbs ensemble is the largest open proper subset (in Lie algebra) where these integrals are convergent. This canonical Gibbs ensemble is convex.

Souriau Theorem of Lie Group Thermodynamics

- Let Ω be the largest open proper subset of $\mathfrak{g}$, Lie algebra of G, such that $\int_M e^{-\langle \beta, U(\xi)\rangle} d\lambda$ and $\int_M \xi . e^{-\langle \beta, U(\xi)\rangle} d\lambda$ are convergent integrals, this set Ω is convex and is invariant under every transformation $Ad_g(.)$, where $g \mapsto Ad_g(.)$ is the adjoint representation of G, such that $Ad_g = T_e i_g$ with $i_g : h \mapsto ghg^{-1}$.

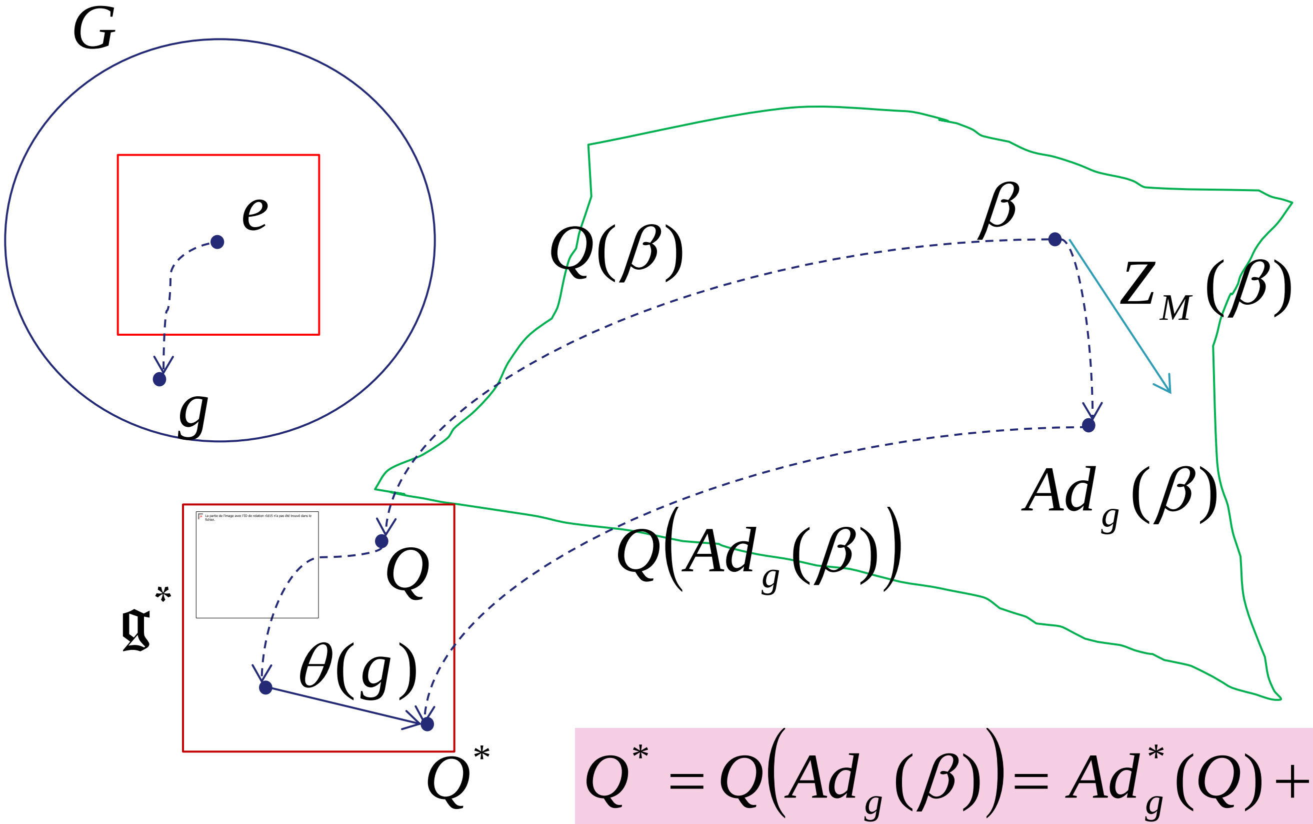
  Let $a : G \times \mathfrak{g}^* \to \mathfrak{g}^*$ a unique affine action a such that linear part is coadjoint representation of G, that is the contragradient of the adjoint representation. It associates to each $g \in G$ the linear isomorphism, $Ad_g^* \in GL(\mathfrak{g}^*)$ satisfying, for each $\xi \in \mathfrak{g}^*$, $X \in \mathfrak{g}$: $\left\langle Ad_g^*(\xi), X \right\rangle = \left\langle \xi, Ad_{g^{-1}}(X) \right\rangle$

$\beta \to Ad_g(\beta)$ $\quad$ $Q \to a(g,Q) = Ad_g^*(Q) + \theta(g)$ $\quad$ $\Phi \to \Phi - \theta(g^{-1})\beta$

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

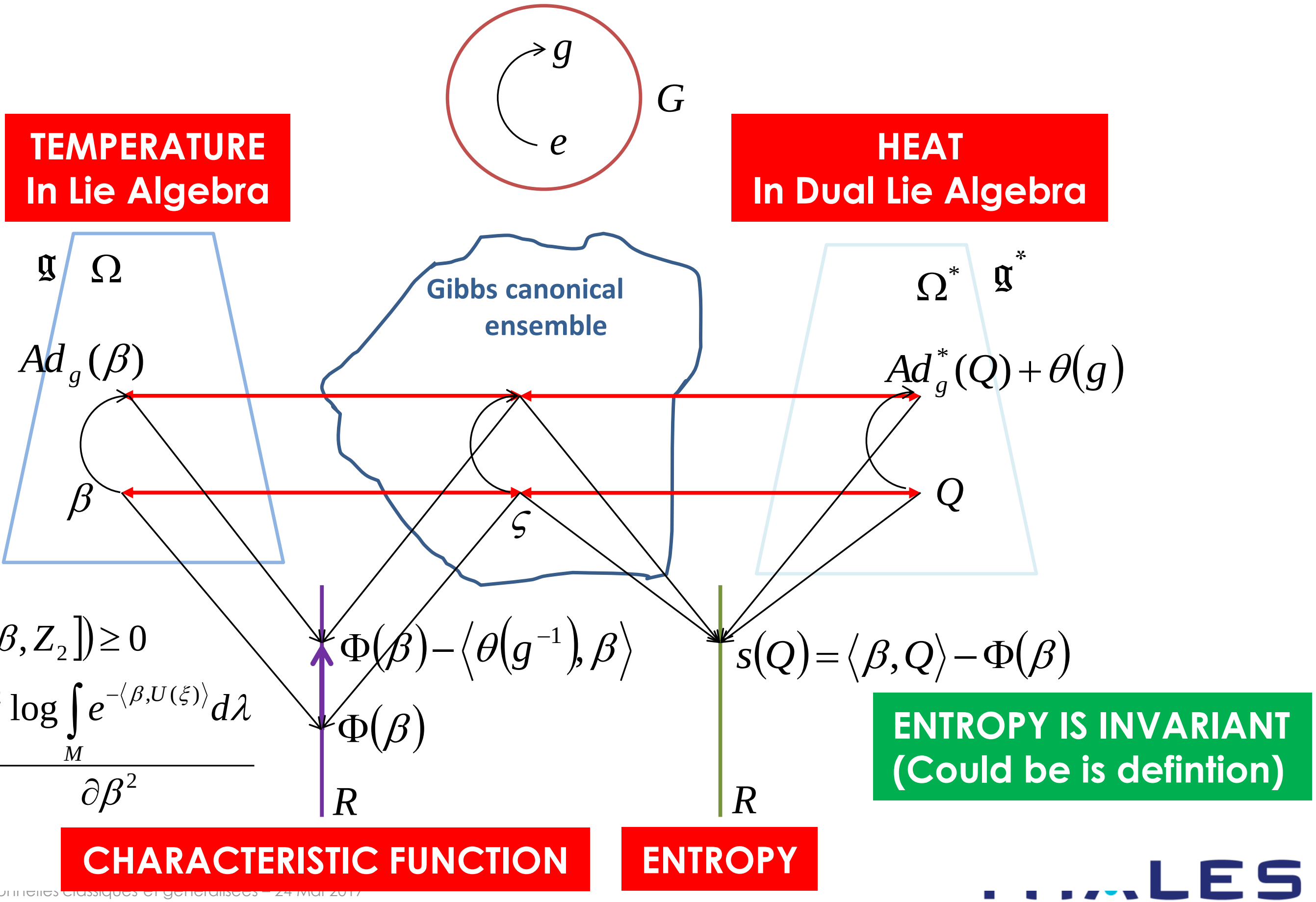
THALES

Lie Group Action on Symplectic Manifold



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

Souriau Model of Lie Group Thermodynamics



FISHER METRIC

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

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

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

Link with Classical Thermodynamics

We have the reciprocal formula:

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

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

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

Geometric (Planck) Temperature in the Lie Algrbra

- Let a Group G of a Manifold M with a moment map E, the **Geometric (Planck) Temperature** β is all elements of Lie Agebra $\mathfrak{g}$ 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 infinitly 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):

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

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

THALES

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: $Q = \dfrac{\partial \Phi}{\partial \beta}$

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

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

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

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

- This quatratic 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:

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

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

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

(**)

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

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

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

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

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

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

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

$$\tilde{\Theta}(X, Y) : \mathfrak{g} \times \mathfrak{g} \to \Re$$
$$\mathrm{X}, \mathrm{Y} \mapsto \left\langle \Theta(X), Y \right\rangle$$

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

OPEN

THALES

Geometric (Planck) Temperature

- We have previously observed that: $\tilde{\Theta}(\beta,Z)+\langle Q,[\beta,Z]\rangle=0$

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

$$\tilde{\Theta}(X,Y)=J_{[X,Y]}-\{J_X,J_Y\} \text{ with } \{.,.\} \text{Poisson Bracket and } J \text{ the Moment Map}$$

$$\mathfrak{g}\to C^\infty(M,R)$$

$$X\to J_X$$

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

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

- $\tilde{\Theta}(X,Y)$ 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 : $\beta\in Ker\,\tilde{\Theta}_\beta$ $\quad \tilde{\Theta}_\beta(\beta,\beta)=0\ ,\ \forall\beta\in\mathfrak{g}$

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 orbite with respect to $Z \in \mathfrak{g}$ and by using positivity of $-\dfrac{\partial Q}{\partial \beta} \geq 0$, 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:

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

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

Fisher Metric of Souriau Lie Group Thermodynamics

Souriau has introduced the Riemannian metric

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

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

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

$$\frac{\partial Q}{\partial\beta}(-[Z_1,\beta],.)=\tilde{\Theta}(Z_1,[\beta,.])+\left\langle Q,Ad_{.Z_1}([\beta,.])\right\rangle=\tilde{\Theta}_\beta(Z_1,[\beta,.])$$

$$-\frac{\partial Q}{\partial\beta}([Z_1,\beta],Z_2.)=\tilde{\Theta}(Z_1,[\beta,Z_2])+\left\langle Q,Ad_{.Z_1}([\beta,Z_2])\right\rangle=\tilde{\Theta}_\beta(Z_1,[\beta,Z_2])$$

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

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

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

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

Galilean Transformation:

- Galilean Lie Group:

$$\begin{pmatrix} A & \vec{b} & \vec{d} \\ 0 & 1 & e \\ 0 & 0 & 1 \end{pmatrix} \text{ with } \begin{cases} A \in SO(3): \text{rotation} \\ \vec{b} \in R^3: \text{boost} \\ \vec{d} \in R^3: \text{space translation} \\ e: \text{time translation} \end{cases}$$

- Galilean Lie Algebra:

$$\begin{pmatrix} j(\vec{\omega}) & \vec{\alpha} & \vec{\delta} \\ 0 & 1 & \varepsilon \\ 0 & 0 & 0 \end{pmatrix} \text{ with } \begin{cases} \vec{\omega} = \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix}, \ \vec{\alpha} \text{ and } \vec{\delta} \in R^3, \ \varepsilon \in R \\ j(\vec{\omega}) = \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix} \in so(3), \ j(\vec{\omega})\vec{r} = \vec{\omega} \times \vec{r} \end{cases}$$

- Action of Lie Group:

$$\begin{pmatrix} A & \vec{b} & \vec{d} \\ 0 & 1 & e \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \vec{r} \\ t \\ 1 \end{pmatrix} = \begin{pmatrix} A\vec{r} + t\vec{b} + \vec{d} \\ t + e \\ 1 \end{pmatrix} \text{ with } \vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

OPEN

THALES

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 **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:

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

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

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

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

Souriau Demo for Galilean moment map for a free particule

- Definition of moment map: $\sigma(dp)(\delta p) = -d\langle J, Z\rangle \ , \ \forall dp$
- Definition of tangent vector field: $Z_V(p) = \delta[a_V(p)]$

$$Z = \begin{pmatrix} j(\vec{\omega}) & \vec{\alpha} & \vec{\delta} \\ 0 & 1 & \varepsilon \\ 0 & 0 & 0 \end{pmatrix} \in \mathfrak{g} \underset{Z_V(p)=\delta[a_V(p)]}{\Rightarrow} \begin{cases} \delta t = \varepsilon \\ \delta r_j = \vec{\omega} \times r_j + \vec{\alpha} t + \vec{\delta} \\ \delta v_j = \vec{\omega} \times v_j + \vec{\alpha} \end{cases}$$

- Lagrange 2 Form:

$$\sigma(dp)(\delta p) = \sum_j \langle m dv, \vec{\omega} \times r_j + \vec{\alpha} t + \vec{\delta} - v\varepsilon \rangle - \langle m(\vec{\omega} \times v_j + \vec{\alpha}), dr - v dt \rangle = -d\langle J, Z\rangle = -dJ_Z = -dH$$

F=0

$$dp = \begin{pmatrix} dt \\ dr \\ dv \end{pmatrix} \text{ and } \delta p = \begin{pmatrix} \delta t \\ \delta r \\ \delta v \end{pmatrix} \Rightarrow \sigma(dp)(\delta p) = \langle m dv - F dt, \delta r - v \delta t \rangle - \langle m \delta v - F \delta t, dr - v dt \rangle$$

- Cocycle: $\theta(g) = J(Ad_g Z) - Ad_g^*(J(Z)) = \left\{ \vec{d} \times \vec{b}, \vec{d} - \vec{b} e, \vec{b}, \frac{1}{2} \left\| \vec{b} \right\|^2 \right\}$

OPEN

THALES

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 Gibbs states for Hamiltonian actions of subgroups of the Galilean group

A gas in a moving box

- Fixed the affine Euclidean reference frame of space $\left(0,\vec{e}_x,\vec{e}_y,\vec{e}_z\right)$ at $t=0$, if we set the value $\tau=t/\varepsilon$, moving frame $\left(0,\vec{e}_x(t),\vec{e}_y(t),\vec{e}_z(t)\right)$ velocity and acceleration is given by the vector field associated to β element of the Lie algebra. Each point has a rotation speed $\left\|\vec{\omega}\right\|/\varepsilon$, speed $\vec{\delta}/\varepsilon$ and aceleration $\vec{\alpha}/\varepsilon$.
- Consider a gas of N point particles, indexed by $i\in\{1,2,...,N\}$, contained in a box with rigid, undeformable walls, whose motion in space is given by the action of the 1-parameter subgroup of the Galilean group, made by the $A(t/\varepsilon)$ with t ∈ R.
- $m_i, r_i(t), v_i(t)$ the mass, position vector and velocity vector, respectively, of the i^th particle at time t.
- Assumption: free particle with neglection of contributions of the collisions of the particles between themselves and with the walls:

$$\langle J,\beta\rangle=\sum_{i=1}^{N}\langle J_i,\beta\rangle \text{ with } \langle J_i(\vec{r}_i,t,\vec{v}_i,m_i),\beta\rangle=m_i\left(\vec{\omega}.(\vec{r}_i\times\vec{v}_i)-(\vec{r}_i-t\vec{v}_i).\vec{\alpha}+\vec{v}_i.\vec{\delta}-\frac{1}{2}\left\|\vec{v}_i\right\|^2\varepsilon\right)$$

- Invariance: $\langle J_i,\beta\rangle$ is s invariant by the action of 1-parameter subgroup

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

Invariance of $\langle J_i, \beta \rangle$

- If the action of the 1-parameter subgroup is $\exp\left(\frac{t}{\varepsilon}\beta\right)$
- According to Souriau equation:

$$a(g,J) = Ad_g^*(J) + \theta(g)$$

We obtain:

$$\langle J_i(p), \beta \rangle = \left\langle Ad_g^*(J_i(p_0)), \beta \right\rangle + \langle \theta(g), \beta \rangle$$

$$\langle J_i(p), \beta \rangle = \left\langle J_i(p_0), Ad_{g^{-1}}\beta \right\rangle + \langle \theta(g), \beta \rangle$$

$$\begin{cases} Ad_{g^{-1}}\beta = \beta \\ \langle \theta(g), \beta \rangle = 0 \end{cases} \Rightarrow \langle J_i(p), \beta \rangle = \langle J_i(p_0), \beta \rangle$$

$$\text{at t} = 0 \text{ then } \left\langle J_i(\vec{r}_i, t, \vec{v}_i, m_i), \beta \right\rangle = m_i \left(\vec{\omega}.(\vec{r}_{i0} \times \vec{v}_{i0}) - \vec{r}_{i0}.\vec{\alpha} + \vec{v}_{i0}.\vec{\delta} - \frac{1}{2}\|\vec{v}_i\|^2 \varepsilon \right)$$

$$= m_i \left(\vec{v}_{i0}.(\vec{\omega} \times \vec{v}_{i0} + \vec{\delta}) - \vec{r}_{i0}.\vec{\alpha} - \frac{1}{2}\|\vec{v}_i\|^2 \varepsilon \right)$$

THALES

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

- By change of variable: $\vec{U}^* = \frac{1}{\varepsilon}\left(\vec{\omega}\times\vec{v}_{i0}+\vec{\delta}\right)$

$$\left\langle J_i(\vec{r}_i,t,\vec{v}_i,m_i),\beta\right\rangle = m_i\varepsilon\left(-\frac{1}{2}\left\|\vec{v}_{i0}-\vec{U}^*\right\|^2-\vec{r}_{i0}.\frac{\vec{\alpha}}{\varepsilon}+\frac{1}{2}\left\|\vec{U}^*\right\|^2\right)$$

- We can then write:

$$\left\langle J_i(\vec{r}_{i0},\vec{p}_{i0}),\beta\right\rangle = -\varepsilon\left(-\frac{1}{2m_i}\left\|\vec{p}_{i0}\right\|^2+m_i f_i(\vec{r}_{i0})\right) \text{ with } \varepsilon=-\frac{1}{\kappa T}$$

$$\text{with}\begin{cases}\vec{p}_{i0}=m_i\vec{w}_{i0}=m_i\left(\vec{v}_{i0}-\vec{U}^*\right)\\ f_i(\vec{r}_{i0})=\vec{r}_{i0}.\frac{\vec{\alpha}}{\varepsilon}-\frac{1}{2\varepsilon^2}\left\|\vec{\omega}\times\vec{r}_{i0}\right\|^2-\frac{\vec{\delta}}{\varepsilon}.\left(\frac{\vec{\omega}}{\varepsilon}\times\vec{r}_{i0}\right)-\frac{1}{2\varepsilon^2}\left\|\vec{\delta}\right\|^2\end{cases}$$

- Gibbs density is given by:

$$\rho(\beta)=\prod_{i=1}^{N}\rho_i(\beta) \text{ with } \rho_i(\beta)=\frac{1}{P_i(\beta)}\exp\left(-\langle J_i,\beta\rangle\right)$$

$$P_i(\beta)=\int_{M_i}\exp\left(-\langle J_i,\beta\rangle\right)d\lambda_{\omega_i}\ ,\ Q_i(\beta)=\int_{M_i}J_i\exp\left(-\langle J_i,\beta\rangle\right)d\lambda_{\omega_i} \text{ et } P(\beta)=\prod_{i=1}^{N}P_i(\beta)$$

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

If we consider the case of the centrifuge

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

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

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

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

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

- 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)

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

OPEN

THALES

Higher Order Maximum Entropy Density

OPEN

THALES

Poly-Symplectic Model of Higher Order Souriau Lie Groups Thermodynamics

Lagrange-Souriau 2 form
& Poincaré-Cartan 1 form

OPEN

Rationale

Objective:

- We will generalize Souriau theory of “Lie Groups Thermodynamics” in the framework of higher order thermodynamics as introduced by R.S. Ingarden for mesoscopic systems

Motivation:

- Souriau Geometric Theory of Heat is well adapted to describe density of probability (Maximum Entropy Gibbs density) of data living on groups or on homogeneous manifolds.
- For Small Data Analytics (Rarified Gases , sparse statistical survey,…), density of maximum entropy should consider Higher Order Moments constraints (Gibbs density is not only defined by first moment but fluctuations request 2nd order and higher moments) as introduced by R.S. Ingarden.

Solution:

- Use of Poly-sympletic model introduced by Christian Günther, replacing the symplectic form by a vector -valued form.
- The polysymplectic approach generalizes the Noether theorem, the existence of momentum mappings, the Lie algebra structure of the space of currents, and the classification of G-homogeneous systems.
- The formalism is covariant, i.e. no special coordinates or coordinate systems on the parameter space are used to construct the Hamiltonian equations.

Higher Order Thermodynamics & Gibbs Density

Polish School works by Ingarden and Jaworski

- Boltzmann ideal gas model can fail if the number of particles is not large enough (rnesoscopic systems), and if the interactions between particles are not weak enough. Gibbs hypothesis can also fail if stochastic interactions with the environment are not sufficiently weak.
- Ingarden in 1992 and Jaworski in 1981 have introduced the concept of second and higher-order temperatures, by assuming a distribution function which includes information not only on the average of the energy but also on higher-order moments, in particular 2nd moment related to fluctuations.
- Ingarden proposed that if we can measure the second cumulant of the energy (the fluctuation of the energy), the equilibrium state is not the canonical state, but would need a second higher order temperature.
- Jaworski showed that the maximum entropy inference has a certain stability property with respect to information corresponding to higher order moments of extensive quantities. It can serve as an argument in favor of the maximum entropy method

$$P_\beta(x) = \frac{1}{Z(\beta)} e^{-\beta . H(x)} \qquad \beta = \frac{1}{k_\beta T} \quad \Rightarrow \quad P_{(\beta_1,\ldots,\beta_n)}(x) = \frac{1}{Z(\beta_1,\ldots,\beta_n)} e^{-\beta_1 . H(x) - \beta_2 (H(x)-U)^2 - \ldots - \beta_n (H(x)-U)^n}$$

Ingarden Higher Order Model of Gibbs Density

- Entropy: $$S = -\int P_{(\beta_1,\ldots,\beta_n)}(x) \log P_{(\beta_1,\ldots,\beta_n)}(x) dx$$
- Massieu Potential: $$\beta_0 = -\log Z(\beta_1,\ldots,\beta_n)$$ with $$Z = \int e^{-\sum_{k=1}^{n} \beta_k x^k} dx$$
- Legendre Transform Preservation: $$S = \sum_{k=1}^{n} \beta_k E(x^k) + \log Z = \sum_{k=1}^{n} \beta_k \frac{\partial \beta_0}{\partial \beta_k} - \beta_0$$
- Higher Order Moments (Heat): $$Q_k = E(x^k) = \frac{\partial \beta_0}{\partial \beta_k} = -\frac{\partial \log Z}{\partial \beta_k}$$

$$Q_k = E(x^k) = Z^{-1} \int x^k e^{-\sum_{k=1}^{n} \beta_k x^k} dx = \int x^k P_{(\beta_1,\ldots,\beta_n)}(x) dx$$

- Higher-order Temperatures and Capacities:

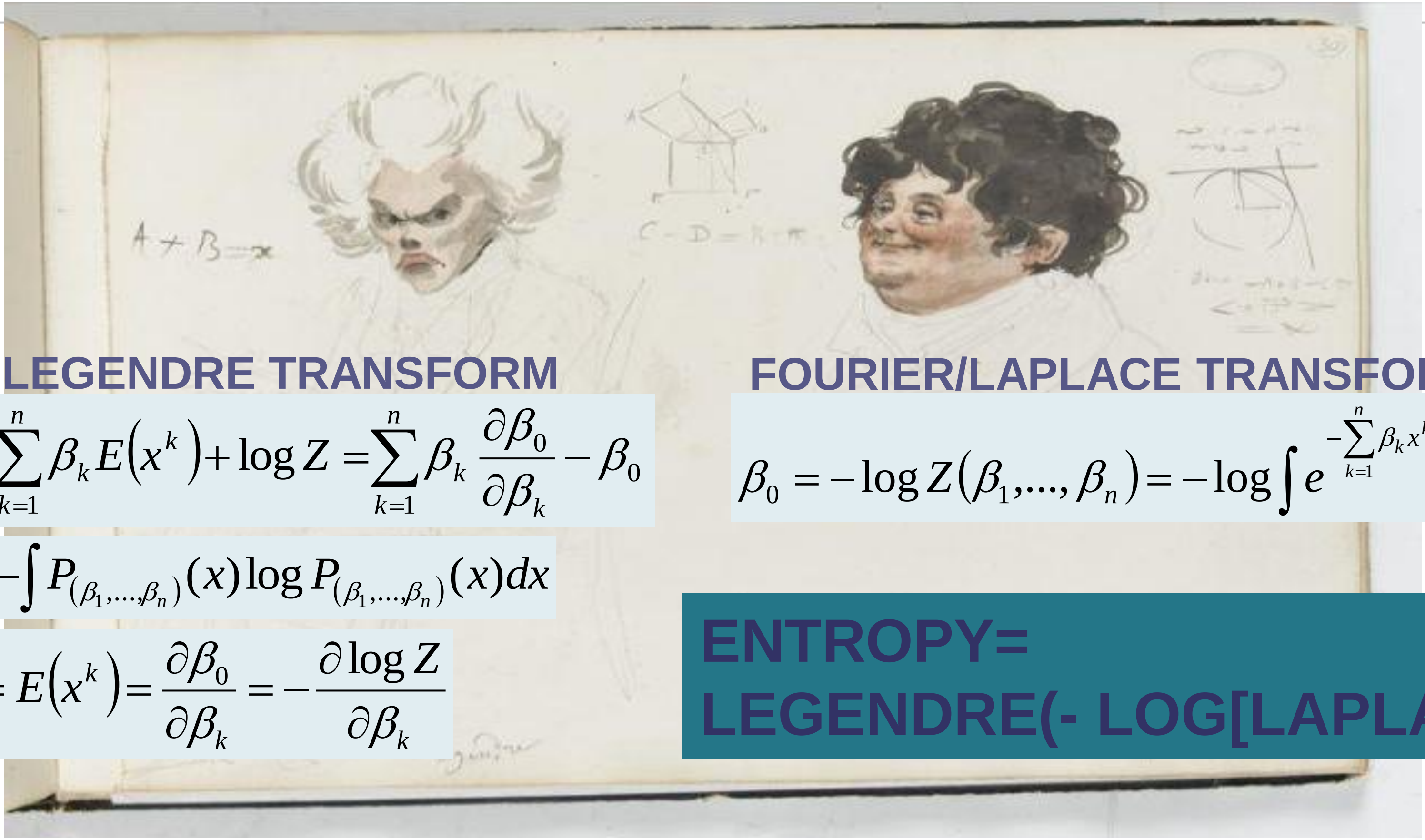
$$\beta_k = \frac{\partial S(Q_1,\ldots,Q_n)}{\partial Q_k}$$ and $$K_k = -\frac{\partial Q_k}{\partial \beta_k}$$

High Order Temperature Model by R.S. Ingarden

High order thermodynamics

- High order moments: $$Q_k = \frac{\partial \Phi(\beta_1,\dots,\beta_n)}{\partial \beta_k} = \frac{\int_M U^k(\xi).e^{-\sum_{k=1}^{n}\langle \beta_k, U^k(\xi)\rangle} d\omega}{\int_M e^{-\sum_{k=1}^{n}\langle \beta_k, U^k(\xi)\rangle} d\omega}$$
- High order 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: $$\beta_k = \frac{\partial S(Q_1,\dots,Q_n)}{\partial Q_k} \qquad K_k = -\frac{\partial Q_k}{\partial \beta_k}$$
- 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: $$p_{Gibbs}(\xi) = e^{\sum_{k=1}^{n}\langle \beta_k, U^k(\xi)\rangle - \Phi(\beta_1,\dots,\beta_n)} = \frac{e^{-\sum_{k=1}^{n}\langle \beta_k, U^k(\xi)\rangle}}{\int_M e^{-\sum_{k=1}^{n}\langle \beta_k, U^k(\xi)\rangle} d\omega}$$

Preservation of Legendre Structure by Higher Order Thermodynamics



LEGENDRE TRANSFORM

$$S=\sum_{k=1}^{n}\beta_k E\left(x^k\right)+\log Z=\sum_{k=1}^{n}\beta_k\frac{\partial\beta_0}{\partial\beta_k}-\beta_0$$

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

$$Q_k=E\left(x^k\right)=\frac{\partial\beta_0}{\partial\beta_k}=-\frac{\partial\log Z}{\partial\beta_k}$$

FOURIER/LAPLACE TRANSFORM

$$\beta_0=-\log Z(\beta_1,\ldots,\beta_n)=-\log\int e^{-\sum_{k=1}^{n}\beta_k x^k}dx$$

ENTROPY= LEGENDRE(- LOG[LAPLACE])

THALES

Higher Order Maximum Entropy Density

Simple and explicit example given by Ingarden

- Maximum Entropy:

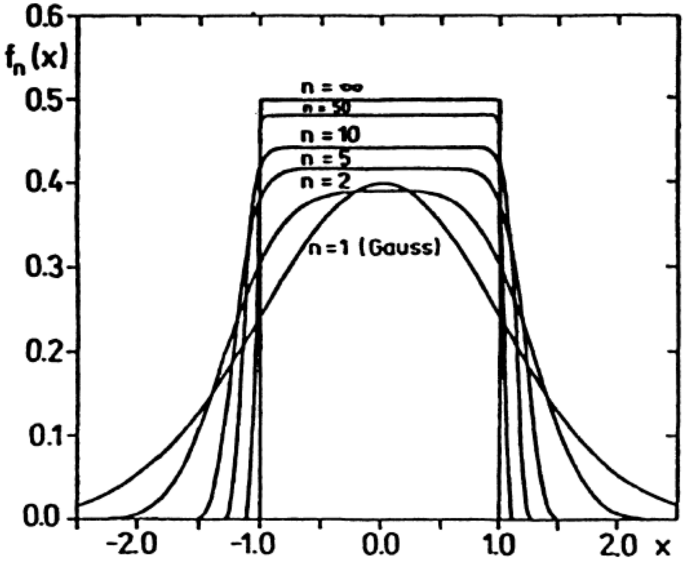
$$S(P) = -\int_{-\infty}^{+\infty} P(x)\log P(x)dx$$

- Under the constraints:

$$P(x) \geq 0 \qquad \int_{-\infty}^{+\infty} P(x)dx = 1 \qquad E\left(x^{2n}\right) = \int_{-\infty}^{+\infty} x^{2n} P(x)dx = \sigma^{2n}$$

- Maximum Entropy Density:

$$P(x) = \frac{1}{2(2n)^{\frac{1}{2n}}\sigma.\Gamma\left(1+1/2n\right)} \exp\left(-\frac{x^{2n}}{2n\sigma^{2n}}\right) = f_n(x)$$



- High Order temperatures : $\beta_n = \dfrac{1}{2n\sigma^{2n}}$

$$Z\left(\beta_n\right) = \frac{2\Gamma\left(1+1/2n\right)}{\beta_n^{1/2n}}$$

$$S(P) = \log Z\left(\beta_n\right) + \frac{1}{2n}$$

$$-\frac{\partial \log Z\left(\beta_k\right)}{\partial \beta_k} = \sigma^{2k} = E\left(x^{2k}\right) = \frac{(2n)^{k/n}\sigma^{2k}\Gamma\left(1+(2k+1)/2n\right)}{(2k+1)\Gamma\left(1+1/2n\right)}$$

$$E\left(x^{2k-1}\right) = 0$$

Higher Order Maximum Entropy Density

Simple and explicit example given by Ingarden

- Maximum Entropy:

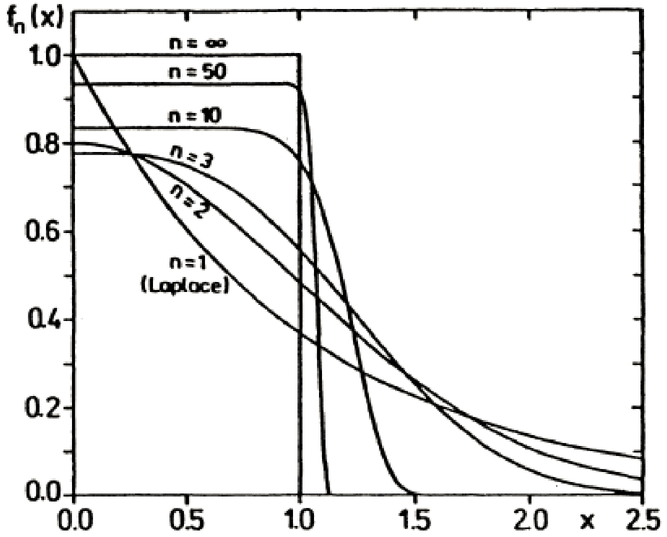
$$S(P) = -\int_{0}^{+\infty} P(x)\log P(x)dx$$

- Under the constraints:

$$P(x) \geq 0 \qquad \int_{0}^{+\infty} P(x)dx = 1 \qquad E\left(x^{2n}\right) = \int_{0}^{+\infty} x^n P(x)dx = \sigma^n$$

- Maximum Entropy Density:

$$P(x) = \frac{1}{n^{\frac{1}{n}}\sigma.\Gamma\left(1+1/n\right)} \exp\left(-\frac{x^n}{n\sigma^n}\right) = f_n(x)$$



- High Order temperatures : $\beta_n = \dfrac{1}{n\sigma^n}$ $\qquad Z(\beta_n) = \dfrac{\Gamma\left(1+1/n\right)}{\beta_n^{1/n}}$

$$-\frac{\partial \log Z(\beta_k)}{\partial \beta_k} = \sigma^k = E\left(x^k\right) = \frac{n^{k/n}\sigma^k\Gamma\left(1+(k+1)/n\right)}{(k+1)\Gamma\left(1+1/n\right)}$$

$$S(P) = \log Z(\beta_n) + \frac{1}{n}$$

Example of Multivariate Gaussian Law (real case)

Multivariate Gaussian law parameterized by moments

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

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

$$= \frac{1}{2} z^T R^{-1} z - m^T R^{-1} z + \frac{1}{2} m^T R^{-1} m$$

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

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

Gaussian Density is a 1st order Maximum Entropy Density !

Poly-symplectic extension of Souriau Lie Groups Thermodynamics

- **We introduce poly-symplectic extension of Souriau Lie group Thermodynamics based on higher-order model of statistical physics introduced by R.S. Ingarden. This extended model could be used for small data analytics.**

- **Initiated by Christian Günther based on n-symplectic model, it has been shown that the symplectic structure on the phase space remains true, if we replace the symplectic form by a vector valued form, that is called polysymplectic:**

  - Gunther C., The polysymplectic Hamiltonian formalism in field theory and calculus of variations I: The local case, J. Differential Geom. n°25, pp. 23-53, 1987
  - Munteanu F., Rey A.M., Salgado M., The Günther's formalism in classical field theory: momentum map and reduction, J. Math. Phys. bf, n°45, vol. 5, pp.1730–1751, 2004
  - Awane A., k-symplectic structures, J. Math. Phys., vol. 33, pp. 4046-4052, 1992
  - A. Awane, M. Goze, Pfaffian systems, k-symplectic systems. Springer, 2000

Günther formalism of Poly-Symplectic Geometry

The Günther formalism is:

- based on the notion of a polysymplectic form, which is a vector valued generalization of symplectic forms.
- Hamiltonian formalism for multiple integral variational problems and field theory presented in a global geometric setting

Günther has introduced in this poly-symplectic formalism:

- Hamiltonian equations
- Canonical transformations
- Lagrange systems,
- Symmetries,
- Field theoretic momentum mappings,
- A classification of G-homogeneous field theoretic systems on a generalization of coadjoint orbits.

Günther formalism of Poly-Symplectic Geometry

6 Günther conditions for a multidimensional Hamiltonian formalism :

- **C0**: For each field system, an evolution space can be constructed, which describes the states of the system completely.
- **C1**: The evolution space carries a geometric structure, which assigns to each function (Hamiltonian density) its Hamiltonian equations.
- **C2**: The geometry of the evolution space gives 'canonical transformations', i.e. the general symmetry group of a system independently of the choice of Hamiltonian density.
- **C3**: The formalism is covariant, i.e. no special coordinates or coordinate systems on the parameter space are used to construct the Hamiltonian equations.
- **C4**: There is an equivalence between regular Lagrange systems and certain (regular) Hamiltonian systems.
- **C5**: For one dimensional parameter space the theory reduces to the ordinary Hamiltonian formalism on symplectic manifolds in classical mechanics.

Hamiltonian field theory by J.E. Marsden is not covariant, because C3 is not verify ! (it causes problems in relativistic theories) : R. Abraham & J. E. Marsden, Foundations of mechanics, 2nd ed., Benjamin and Cummings, New York, 1978

Multisymplectic approach by Tulczyjew, based on general theory by Dedecker,do not satisfy C1 and C2 !

THALES

Günther formalism of Poly-Symplectic Geometry

The key idea for this generalized Hamiltonian formalism:

- to replace the symplectic form in classical mechanics by a vector valued, so called polysymplectic form
- The evolution space of a classical field will appear as the dual of a jet bundle, which carries naturally a polysymplectic structure.
- The polysymplectic form will assign to each function on the evolution space the Hamiltonian equations via the 'musical morphisms.'
- Canonical transformations are bundle isomorphisms leaving this polysymplectic form invariant.

The polysymplectic approach not only recovers all classical results but leads globally and locally to many new results generalizing the Noether theorem based on canonical transformations, <u>the existence of momentum mappings</u>, the Lie algebra structure of the space of currents, the reduction procedure, and the classification of G-homogeneous systems.

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 $\mathfrak{g}^* \times \overset{(n)}{\dots} \times \mathfrak{g}^*$ called n-coadjoint action:

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

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

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

$$\mathrm{Pr}_k : \mathfrak{g}^* \times \overset{(n)}{\dots} \times \mathfrak{g}^* \to \mathfrak{g}^* \quad , \quad (\nu_1, \dots, \nu_n) \mapsto \nu_k$$

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

$$\pi_k : \mathrm{O}_\mu = \mathrm{O}_{(\mu_1, \dots, \mu_n)} \to \mathrm{O}_{\mu_k}$$

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

OPEN

THALES

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(.,.) of the Lie group G on $\mathfrak{g}^* \overset{(n)}{\times \ldots \times} \mathfrak{g}^*$ for which the polymoment map with $x \in M$ and $g \in G$: $J^{(n)} = \left(J^1, \ldots, J^n\right): M \to \mathfrak{g}^* \overset{(n)}{\times \ldots \times} \mathfrak{g}^*$ that verifies: $J^{(n)}\left(\Phi_g(x)\right) = a(g, J^{(n)}(x)) = Ad_g^{*(n)}\left(J^{(n)}(x)\right) + \theta^{(n)}(g)$

  with $Ad_g^{*(n)}\left(J^{(n)}(x)\right) = \left(Ad_g^* J^1, \ldots, Ad_g^* J^n\right)$ and $\theta^{(n)}(g) = \left(\theta^1(g), \ldots, \theta^n(g)\right)$
- $\theta^{(n)}(g)$ is a poly-symplectic one-cocycle
- We can also defined poly-symplectic two-cocycle:

$$\tilde{\Theta}^{(n)} = \left(\tilde{\Theta}^1, \ldots, \tilde{\Theta}^n\right) \text{ with } \tilde{\Theta}^k(X,Y) = \left\langle \Theta^k(X), Y \right\rangle = J^k_{[X,Y]} - \left\{J^k_X, J^k_Y\right\}$$

$$\text{where } \Theta^k(X) = T_e\theta^k\left(X(e)\right)$$

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

$$g_\beta\left([\beta, Z_1], Z_2\right) = diag\left[\tilde{\Theta}_{\beta_k}(Z_1, Z_2)\right]_k \ , \ \forall Z_1 \in \mathfrak{g}, \forall Z_2 \in \mathrm{Im}\left(ad_\beta(.)\right), \beta = \left(\beta_1, \ldots, \beta_n\right)$$

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

Poly-symplectic extension of Souriau Lie Groups Thermodynamics

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

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

- with characteristic function: $\Phi(\beta_1,\ldots,\beta_n) = -\log \int_M e^{-\sum_{k=1}^{n}\left\langle \beta_k, U^{\otimes k}(\xi)\right\rangle} d\omega$

- We extrapolate Souriau results, who proved that $\int_M U^{\otimes k}(\xi).e^{-\left\langle \beta_k, U^{\otimes k}(\xi)\right\rangle} d\omega$ is locally normally convergent using $\left\| U^{\otimes k} \right\| = \underset{U}{Sup}\left\langle E, U \right\rangle^k$, a multi-linear norm and where $U^{\otimes k} = U \otimes \overset{(k)}{U} \ldots \otimes U$ is defined as a tensorial product.

- Entropy is defined by Legendre transform of Souriau-Massieu characteristic function:

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

Poly-symplectic extension of Souriau Lie Groups Thermodynamics

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

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

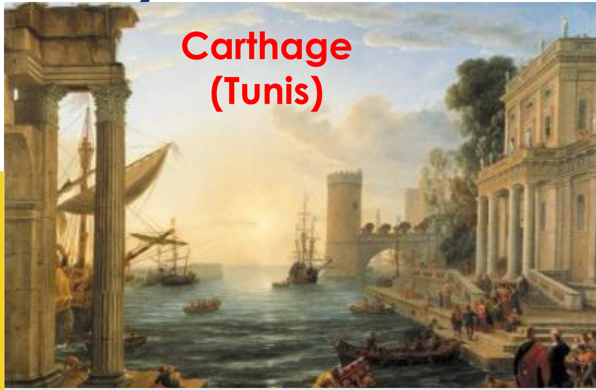
with $U^{\otimes k} = U \otimes \overset{(k)}{U} \dots \otimes U$ and $\Phi(\beta_1,\dots,\beta_n) = -\log \int_M e^{-\sum_{k=1}^{n} \langle \beta_k, U^{\otimes k}(\xi) \rangle} d\omega$

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

$$\begin{cases} Q_k = \dfrac{\partial \Phi(\beta_1,\dots,\beta_n)}{\partial \beta_k} \quad \text{with} \quad Q = (Q_1,\dots,Q_n) \in \mathfrak{g}^* \times \overset{(n)}{\dots} \times \mathfrak{g}^* \\ \beta_k = \dfrac{\partial S(Q_1,\dots,Q_n)}{\partial Q_k} \quad \text{with} \quad \beta = (\beta_1,\dots,\beta_n) \in \mathfrak{g} \times \overset{(n)}{\dots} \times \mathfrak{g} \end{cases}$$

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

1952-1958 : J.M. Souriau Maître de Conférences, puis Professeur titulaire à l'Institut des Hautes Études de 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.

Comparison of Affine Representation of Lie Group and Lie Algebra in Souriau an Koszul works

OPEN

Affine representation of Lie group and Lie algebra by Souriau

- Souriau called the Mechanics deduced from his model: "**Affine Mechanics**"

- Let *G* be a Lie group and *E* a finite-dimensional vector space. A map

$A: G \to Aff(E)$ can always be written as:

$$A(g)(x) = R(g)(x) + \theta(g) \text{ with } g \in G, x \in E$$

where the maps $R: G \to GL(E)$ and $\theta: G \to E$ are determined by *A*. The map *A* is an affine representation of *G* in *E*.

- The map $\theta: G \to E$ is a one-cocycle of *G* with values in *E*, for the linear representation *R*; it means that θ is a smooth map which satisfies, for all $g, h \in G$:

$$\theta(gh) = R(g)(\theta(h)) + \theta(g)$$

OPEN

THALES

Affine representation of Lie group and Lie algebra by Souriau

- Let $\mathfrak{g}$ be a Lie algebra and *E* a finite-dimensional vector space. A linear map $a:\mathfrak{g}\to aff(E)$ always can be written as:

$$a(X)(x)=r(X)(x)+\Theta(X) \text{ with } X\in\mathfrak{g}, x\in E$$

- where the linear maps $r:\mathfrak{g}\to gl(E)$ and $\Theta:\mathfrak{g}\to E$ are determined by *a*. The map *a* is an affine representation of *G* in *E*.

- The linear map $\Theta:\mathfrak{g}\to E$ is a one-cocycle of *G* with values in *E*, for the linear representation *r*; it means that Θ satisfies, for all $X,Y\in\mathfrak{g}$:

$$\Theta([X,Y])=r(X)(\Theta(Y))-r(Y)(\Theta(X))$$

- Θ is called the one-cocycle of $\mathfrak{g}$ associated to the affine representation *a*.

- the associated cocycle $\Theta:\mathfrak{g}\to E$ is related to the one-cocycle $\theta:G\to E$ by:

$$\Theta(X)=T_e\theta(X(e)) \ , \ X\in\mathfrak{g}$$

OPEN

THALES

Equivariance of Souriau Moment Map

- There exists a unique affine action a such that the linear part is a coadjoint representation:

$$a : G \times \mathfrak{g}^* \to \mathfrak{g}^*$$

$$a(g,\xi) = Ad^*_{g^{-1}}\xi + \theta(g)$$

with $\left\langle Ad^*_{g^{-1}}\xi, X \right\rangle = \left\langle \xi, Ad_{g^{-1}}X \right\rangle$

- that induce equivariance of moment J .

Action of Lie Group on a Symplectic Manifold

- Let $\Phi : G \times M \to M$ be an action of Lie Group G on differentiable manifold M, the fundamental field associated to an element X of Lie algebra $\mathfrak{g}$ of group G is the vectors field X_M on M:

$$X_M(x) = \frac{d}{dt}\Phi_{\exp(-tX)}(x)\Big|_{t=0} \quad \text{with} \quad \Phi_{g_1}\left(\Phi_{g_2}(x)\right) = \Phi_{g_1 g_2}(x) \text{ and } \Phi_e(x) = x$$

- Φ is Hamiltonian on a symplectic manifold M, if Φ is symplectic and if for all $X \in \mathfrak{g}$, the fundamental field X_M is globally Hamiltonian.

- There is a unique action *a* of the Lie group G on the dual $\mathfrak{g}^*$ of its Lie algebra for which the moment map J is equivariant, that means satisfies for each $x \in M$:

$$J\left(\Phi_g(x)\right) = a(g, J(x)) = Ad^*_{g^{-1}}\left(J(x)\right) + \theta(g)$$

$$\tilde{\Theta} : \mathfrak{g} \times \mathfrak{g} \to \Re \qquad \left\langle T_e\theta(X), Y\right\rangle = \left\langle \Theta(X), Y\right\rangle = \tilde{\Theta}(X, Y) = J_{[X,Y]} - \left\{J_X, J_Y\right\}$$

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

Affine representation of Lie group and Lie algebra by Koszul

- Let Ω be a convex domain in R^n containing no complete straight lines, and an associated convex cone $V(\Omega)=\{(\lambda x, x)\in R^n\times R / x\in\Omega, \lambda\in R^+\}$
Then there exists an affine embedding:

$$\ell : x\in\Omega \mapsto \begin{bmatrix} x \\ 1 \end{bmatrix} \in V(\Omega)$$

- If we consider η the group of homomorphism of $A(n,R)$ into $GL(n+1,R)$ given by:

$$s\in A(n,R)\mapsto \begin{bmatrix} \mathbf{f}(s) & \mathbf{q}(s) \\ 0 & 1 \end{bmatrix}\in GL(n+1,R)$$

- and associated affine representation of Lie Algebra:

$$\begin{bmatrix} f & q \\ 0 & 0 \end{bmatrix}$$

- with $A(n,R)$ the group of all affine transformations of R^n. We have

$\eta(G(\Omega))\subset G(V(\Omega))$ and the pair (η,ℓ) of the homomorphism

$\eta : G(\Omega)\rightarrow G(V(\Omega))$ and the map $\ell:\Omega\rightarrow V(\Omega)$ is equivariant.

Affine representation of Lie group and Lie algebra by Koszul

- Let G a connex Lie Group and E a real or complex vector space of finite dimension, Koszul has introduced an affine representation of G in E such that the following is an affine transformation: $E \to E$

$$a \mapsto sa \quad \forall s \in G$$

- We set $A(E)$ the set of all affine transformations of a vector space E, a Lie Group called affine transformation group of E. The set $GL(E)$ of all regular linear transformations of E, a subgroup of $A(E)$.

- We define a linear representation from E to $GL(E)$:

$$\mathbf{f} : G \to GL(E)$$

$$s \mapsto \mathbf{f}(s)a = sa - so \quad \forall a \in E$$

- and an application from G to E:

$$\mathbf{q} : G \to E$$

$$s \mapsto \mathbf{q}(s) = so \quad \forall s \in G$$

- Then we have $\forall s,t \in G$: $\mathbf{f}(s)\mathbf{q}(t) + \mathbf{q}(s) = \mathbf{q}(st)$

OPEN

$$\mathbf{f}(s)\mathbf{q}(t) + \mathbf{q}(s) = s\mathbf{q}(t) - so + so = s\mathbf{q}(t) = sto = \mathbf{q}(st)$$

THALES

- On the contrary, if an application $\mathbf{q}$ from G to E and a linear representation $\mathbf{f}$ from G to $GL(E)$ verify previous equation, then we can define an affine representation of G in E, written $(\mathbf{f},\mathbf{q})$:

$$Aff(s): a \mapsto sa = \mathbf{f}(s)a + \mathbf{q}(s) \quad \forall s \in G, \forall a \in E$$

- The condition $\mathbf{f}(s)\mathbf{q}(t) + \mathbf{q}(s) = \mathbf{q}(st)$ is equivalent to requiring the following mapping to be an homomorphism: $Aff : s \in G \mapsto Aff(s) \in A(E)$

- We write f the linear representation of Lie algebra $\mathfrak{g}$ of G, defined by $\mathbf{f}$ and q the restriction to $\mathfrak{g}$ of the differential to $\mathbf{q}$ (f and q the differential of $\mathbf{f}$ and $\mathbf{q}$ respectively), Koszul has proved that:

$$f(X)q(Y) - f(Y)q(X) = q([X,Y]) \quad \forall X,Y \in \mathfrak{g}$$

$$\text{with} \quad f: \mathfrak{g} \to gl(E) \quad \text{and} \quad q: \mathfrak{g} \mapsto E$$

Where $gl(E)$ the set of all linear endomorphisms of E, the Lie algebra of $GL(E)$

Affine representation of Lie group and Lie algebra by Koszul

- Conversely, if we assume that $\mathfrak{g}$ admits an affine representation (f,q) on E, using an affine coordinate system $\{x^1,...,x^n\}$ on E, we can express an affine mapping $v \mapsto f(X)v + q(Y)$ by an $(n+1)\times(n+1)$ matrix representation:

$$aff(X) = \begin{bmatrix} f(X) & q(X) \\ 0 & 0 \end{bmatrix}$$

- where $f(X)$ is a $n \times n$ matrix and $q(X)$ is a n row vector.

- If we denote $\mathfrak{g}_{aff} = aff(\mathfrak{g})$, we write G_{aff} the linear Lie subgroup of $GL(n+1,R)$ generated by $\mathfrak{g}_{aff}$. An element of $s \in G_{aff}$ is expressed by:

$$Aff(s) = \begin{bmatrix} \mathbf{f}(s) & \mathbf{q}(s) \\ 0 & 1 \end{bmatrix}$$

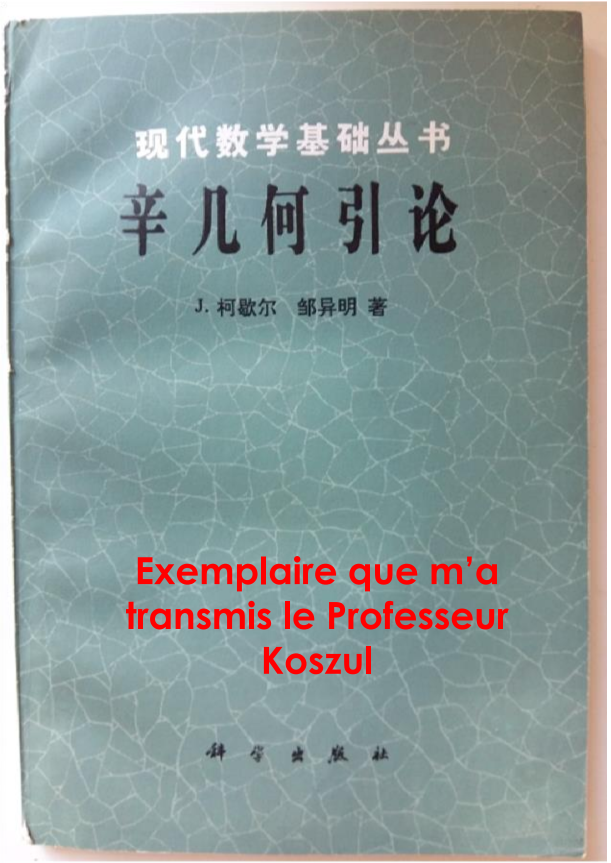
Affine representation of Lie Group and Lie Algebra by Souriau and Koszul

Souriau Model of Affine Representation of Lie Groups and Algebra	Koszul Model of Affine Representation of Lie Groups and Algebra
$A(g)(x) = R(g)(x) + \theta(g)$ with $g \in G, x \in E$ $R : G \to GL(E)$ and $\theta : G \to E$	$Aff(s) : a \mapsto sa = \mathfrak{f}(s)a + \mathfrak{q}(s) \quad \forall s \in G, \forall a \in E$ $\mathfrak{f} : G \to GL(E)$ $s \mapsto \mathfrak{f}(s)a = sa - so \quad \forall a \in E$ $\mathfrak{q} : G \to E$ $s \mapsto \mathfrak{q}(s) = so \quad \forall s \in G$
$\theta(gh) = R(g)(\theta(h)) + \theta(g)$ with $g, h \in G$ $\theta : G \to E$ is a one-cocycle of G with values in E,	$\mathfrak{q}(st) = \mathfrak{f}(s)\mathfrak{q}(t) + \mathfrak{q}(s)$
$a(X)(x) = r(X)(x) + \Theta(X)$ with $X \in \mathfrak{g}, x \in E$ The linear map $\Theta : \mathfrak{g} \to E$ is a one-cocycle of G with values in E: $\Theta(X) = T_e\theta\left(X(e)\right), X \in \mathfrak{g}$	$v \mapsto f(X)v + q(Y)$ f and q the differential of $\mathfrak{f}$ and $\mathfrak{q}$ respectively
$\Theta\left([X,Y]\right) = r(X)\left(\Theta(Y)\right) - r(Y)\left(\Theta(X)\right)$	$q\left([X,Y]\right) = f(X)q(Y) - f(Y)q(X) \; \forall X, Y \in \mathfrak{g}$ with $f : \mathfrak{g} \to gl(E)$ and $q : \mathfrak{g} \mapsto E$
none	$aff(X) = \begin{bmatrix} f(X) & q(X) \\ 0 & 0 \end{bmatrix}$
none	$Aff(s) = \begin{bmatrix} \mathfrak{f}(s) & \mathfrak{q}(s) \\ 0 & 1 \end{bmatrix}$

Link between Jean-Louis Koszul and Jean-Marie Souriau: Small Green Book

Jean-Louis Koszul Lecture in China 1986

- "***Introduction à la géométrie symplectique***", in Chinese
- Chuan Yu Ma has written
  - This beautiful, modern book should not be absent from any institutional library. …. During the past eighteen years there has been considerable growth in the research on symplectic geometry. Recent research in this field has been extensive and varied. This work has coincided with developments in the field of analytic mechanics. Many new ideas have also been derived with the help of a great variety of notions from 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.



THALES

On retrouve dans le livre de Koszul, les équations de Souriau:

17.2. **命题.** 设 (M, ω) 是一连通的 Hamilton G-空间，

$$\boxed{\mu: M \to \mathfrak{g}^*}$$

是 (M, ω) 的一个矩射，则

(i) 对任意的 $s \in G$，

$$\boxed{\varphi_\mu(s) = \mu(sx) - Ad^*(s)\mu(x)}$$

是 $\mathfrak{g}^*$ 中不依赖于点 $x \in M$ 的一个元素.

(ii) 对任意的 $s, t \in G$ 有

$$\varphi_\mu(st) = \varphi_\mu(s) + Ad^*(s)\varphi_\mu(t).$$

(iii) 对任意的 $a, b \in \mathfrak{g}$ 有

$$c_\mu(a, b) = \langle d\varphi_\mu(a), b\rangle,$$

c_μ 的定义见 §16.

推论. 从 $G \times \mathfrak{g}^*$ 到 $\mathfrak{g}^*$ 内的映射

$$\boxed{(s, \xi) \longmapsto s\xi = Ad^*(s)\xi + \varphi_\mu(s), \ s \in G, \ \xi \in \mathfrak{g}^*,}$$

[22] C. L. Siegel, Symplectic geometry, *Amer. J. Math.*, 1—86, 1943.

[23] J. M. Souriau, Structures des systemes dynamiques, Dunod Paris, 1969.

[24] W. W. Symes, Hamiltonian group actions and integrable systems, *Physica*, **1**. D, 339—374, 1980.

[25] E. B. Vinberg, The theory of convex homogeneous cones, *Moscow Math. Soc.*, **12**, 1963.

[26] N. R. Wallach, Symplectic geometry and Fourier analysis, Math. Sci. Press, Brookline, Mass, 1977.

[27] A. Weil, Variétés kaehleriennes, Hermann, Paris, 1958.

[28] A. Weinstein, Symplectic manifolds and their Lagrangian submanifolds, *Adv. Math.*, **6**, 329—346, 1971.

[29] ———, Lectures on symplectic manifolds, C. B. M. S. regional conference series, 29, A. M. S. Rhode Island, 1977.

[30] H. Weyl, Classical groups, Princeton University Press, 1946.

[31] 严志达，半单纯李群李代数表示论，上海科学技术出版社，1963.

Développements de Koszul du modèle de Souriau (1/3)

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

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

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

- Propriété de l'application moment μ: $\mu : M \to \mathfrak{g}^* \qquad x \mapsto sx \ , \quad x \in M$

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

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

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

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

OPEN

Développements de Koszul du modèle de Souriau (2/3)

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

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

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

- Etude de : $\mathrm{c}_\mu(\mathrm{a},\mathrm{b}) = \langle \mathrm{d}\theta_\mu(a), b \rangle \ , \ a,b \in \mathfrak{g}$

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

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

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

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

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

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

Développements de Koszul du modèle de Souriau (3/3)

- Action du groupe sur le dual de l'algèbre de Lie:

$$G \times \mathfrak{g}^* \to \mathfrak{g}^*, (s,\xi) \mapsto s\xi = Ad_s^* \xi + \theta_\mu(s)$$

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

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

- Propriétés:

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

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

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

THALES

Links with Natural Exponential Families Invariant by a Group: Casilis and Letac

OPEN

NEF (Natural Exponential Families): Letac & Casalis

- Let E a vector space of finite size, E^* its dual. $\langle\theta, x\rangle$ Duality braket with $(\theta, x) \in E^* \times E$. μ Positive Radon measure on E, Laplace transform is :

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

- Transformation $k_\mu(\theta)$ defined on $\Theta(u)$ 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 (domian of means): $k'_\mu(\theta) = \int_E xP(\theta,\mu)\mu(dx)$
- And the inverse function: $\psi_\mu : M_F \to \Theta(\mu)$ with $M_F = \mathrm{Im}\left(k'_\mu(\Theta(\mu))\right)$
- Covariance operator: $V_F(m) = k''_\mu\left(\psi_\mu(m)\right) = \left(\psi'_\mu(m)\right)^{-1}, m \in M_F$

NEF (Natural Exponential Families): Letac & Casalis

- Measure generetad by a familly 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 familly $\varphi(F) = \{\varphi(P(\theta,\mu)), \theta \in \Theta(\mu)\}$ is an exponential familly of E generated by $\varphi(\mu)$

- <u>Definition:</u> An exponential familly 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 familly of E and G affine group of E, then F is invariant by G if and only:

$$\exists a : G \to E^*, \exists b : G \to R, \text{such that}:$$

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

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

- 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^*, $\Im(G^n, E^*)$ called inhomogenesous n-cochains, then we can define the operators: $d^n : \Im(G^n, E^*) \rightarrow \Im(G^{n+1}, E^*)$

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

$$+ (-1)^{n+1} F(g_1, g_2, \cdots, g_n)$$

NEF (Natural Exponential Families): Letac & Casalis

- Let $Z^n(G,E^*) = Ker(d^n), B(G,E^*) = \mathrm{Im}(d^{n-1})$, with Z^n inhomogneous 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:

$$d^0 : E^* \to \Im(G,E^*) \qquad Z^0 = \{x \in E^*; g.x = x, \forall g \in G\}$$

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

$$d^1 : \Im(G,E^*) \to \Im(G^2,E^*)$$

$$F \mapsto d^1F \quad , \quad d^1F(g_1,g_2) = g_1.F(g_2) - F(g_1g_2) + F(g_1)$$

$$Z^1 = \{F \in \Im(G,E^*); F(g_1g_2) = g_1.F(g_2) + F(g_1), \forall (g_1,g_2) \in G^2\}$$

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

NEF (Natural Exponential Families): Letac & Casalis

- When the Cohomology Group $H^1(G,E^*)=0$ then $Z^1(G,E^*)=B^1(G,E^*)$

$$\Rightarrow \exists c \in E^*, \text{such that } \forall g \in G, a(g)=\left(I_d - {}^t g^{-1}\right)c$$

Then if $F=F(\mu)$ is an exponential familly 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

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

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

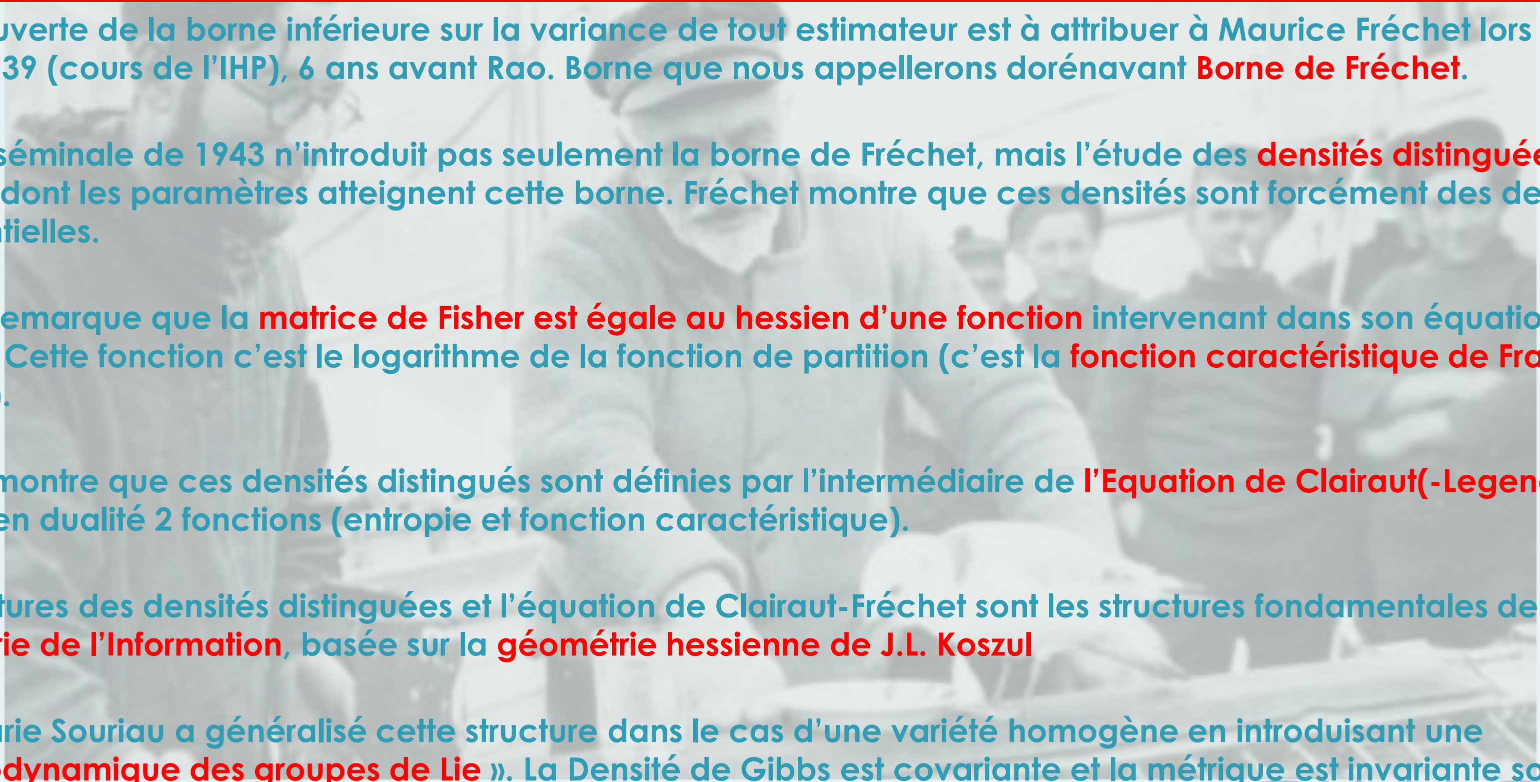
$$\exists \text{fixed point} \Rightarrow \forall g \in G, A_g(c) = {}^t g^{-1}c + a(g) = c \Rightarrow a(g) = \left(I_d - {}^t g^{-1}\right)c$$

THALES

Conclusion

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 distingués 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.**

References

Reference Book: Libermann & Marle

Symplectic Geometry and Analytical Mechanics

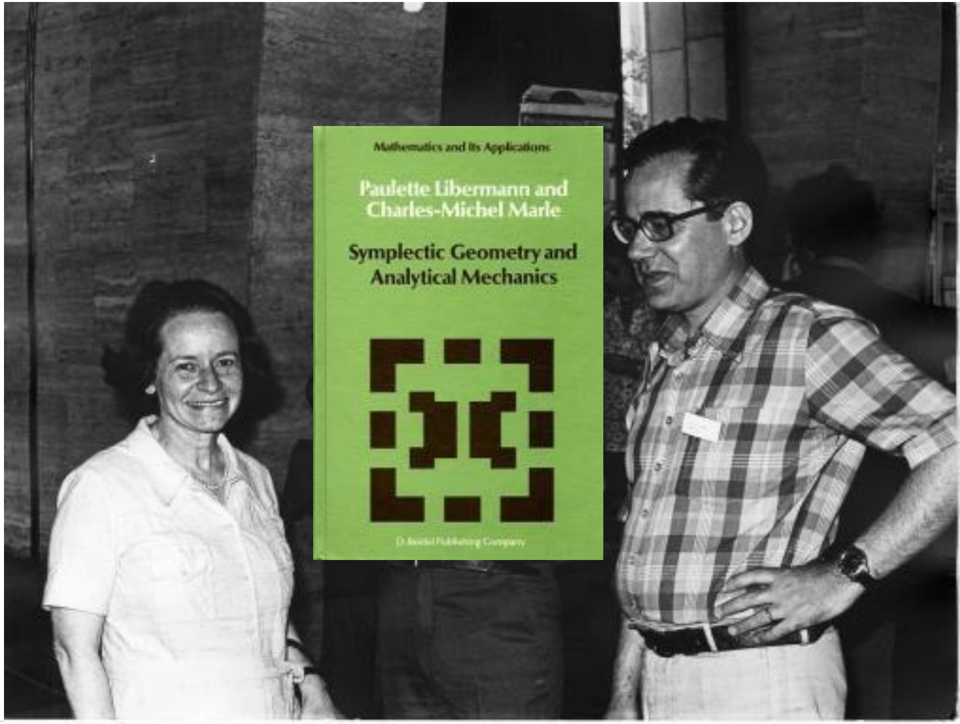
- Paulette Libermann & Charles-Michel Marle

https://www.agnesscott.edu/lriddle/WOMEN/abstracts/libermann_abstract.htm

Paulette Libermann, Legendre foliations on contact manifolds, Differential Geometry and Its Applications, n°1, pp.57-76, 1991

See also:

- Marle, C.-M. From Tools in Symplectic and Poisson Geometry to J.-M. Souriau's Theories of Statistical Mechanics and Thermodynamics. Entropy 2016, 18, 370.
- http://www.mdpi.com/1099-4300/18/10/370

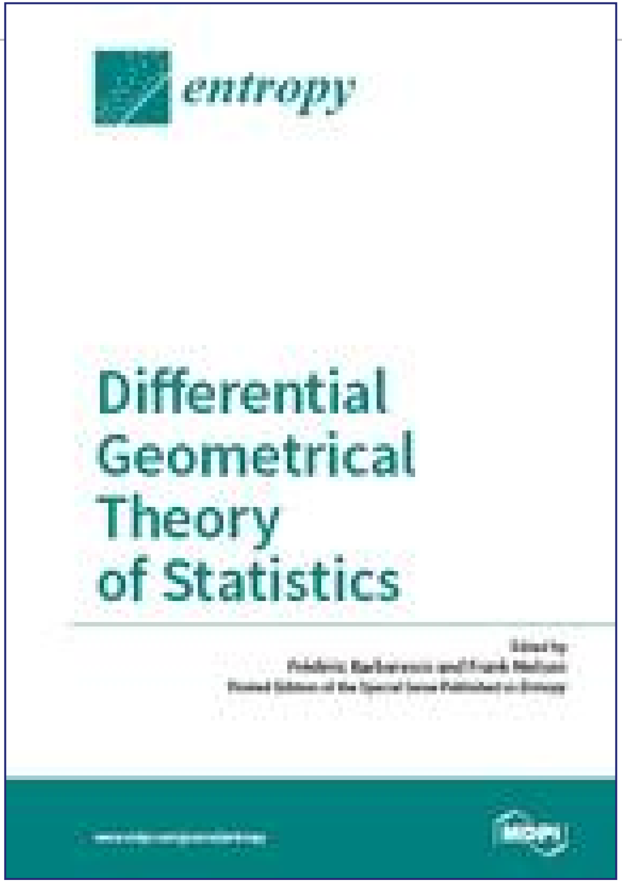


MDPI Book
« Information, Entropy and their Geometric Structures »

http://books.mdpi.com/pdfview/book/127
Could be downloaded at weblink:
http://www.mdpi.com/books/pdfdownload/book/127/1

MDPI Book « Differential Geometrical Theory of Statistics »

http://www.mdpi.com/books/pdfview/book/313



Tutorial sur la géométrie de l'Information

- **Exposé de Frank Nielsen (LIX, Ecole Polytechnique) au workshop SHANNON 100 à l'Institut Henri Poincaré, fin Octobre 2016**

The dual geometry of Shannon information

100 claude shannon 1916-2016

- **Vidéo:** https://www.youtube.com/watch?v=aGxZoKSk6CQ&index=11&list=PL9kd4mpdvWcDMCJ-SP72HV6Bme6CSqk_k

- **Planches:** https://www.lix.polytechnique.fr/~nielsen/CIG-slides.pdf

To Go Deeper

UNESCO UNITWIN Website on « Geometric Science of Information »:

- http://forum.cs-dc.org/category/72/geometric-science-of-information
- http://forum.cs-dc.org/topic/369/geometric-science-of-information-presentation-organisation-subscription

GSI « Geometric Science of Information » Conferences:

- GSI'13, Ecole des Mines de Paris: https://www.see.asso.fr/gsi2013
- GSI'15, Ecole Polytechnique: http://forum.cs-dc.org/category/90/gsi2015 + videos: http://forum.cs-dc.org/category/90/gsi2015
- GSI'17 : www.si2017.org + Videos : https://www.youtube.com/channel/UCnE9-LbfFRqtaes49cN2DVg/videos

TGSI'17 "Topological & Geometrical Structures of Information", CIRM Luminy, August 2017

- TGSI'17: http://forum.cs-dc.org/category/94/tgsi2017

34th International Workshop on Bayesian Inference and Maximum Entropy

- MaxEnt'14, : https://www.see.asso.fr/maxent14

Leon Brillouin Seminar on "Geometric Science of Information"

- http://repmus.ircam.fr/brillouin/home + http://repmus.ircam.fr/brillouin/past-events

Information geometry and probability tools in abstract space for signal and image analysis

- GDR ISIS Technical Day: http://forum.cs-dc.org/topic/410/gdr-isis-gsi-day-information-geometry-and-probabilty-tools-in-abstract-space-for-signal-and-image-analysis

Google+ and Google Scholar

- https://plus.google.com/u/0/100141956413652325744/posts
- http://scholar.google.fr/citations?hl=fr&user=Pe_EE9UAAAAJ&view_op=list_works&sortby=pubdate

References

- [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. Entropy 2016, 18, 386.
- [9] Barbaresco F. , Higher Order Souriau Lie Group Thermodynamics, GSI'17 Conference, Springer LNCS, 2017
- [10] Casalis, M. Familles exponentielles naturelles invariantes par un groupe. Thèse de l'Univ. Paul Sabatier, Toulouse, 1990
- [11] Casalis, M. Familles Exponentielles Naturelles sur Rd Invariantes par un Groupe, International Statistical Review (Revue Internationale de Statistique), Vol. 59, No. 2, pp. 241-262, Aug. 1991

References

- [12] Duhem, P. Sur les équations générales de la thermodynamique, Annales scientifiques de l'ENS, 3ème série, tome 9, 1891, pp.231-266.
- [13] Eriksen, P.S. Geodesics connected with the Fisher metric on the multivariate normal manifold. Lancaster : Proceedings of the GST Workshop, 1987
- [14] 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.
- [15] 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).
- [16] 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).
- [17] Koszul, J.L. Sur la forme hermitienne canonique des espaces homogènes complexes. Can. J. Math. 1955, 7, 562–576.
- [18] Koszul, J.L. Exposés sur les Espaces Homogènes Symétriques; Publicação da Sociedade de Matematica de São Paulo: São Paulo, Brazil, 1959.
- [19] Koszul, J.L. Domaines bornées homogènes et orbites de groupes de transformations affines. Bull. Soc. Math. Fr. 1961, 89, 515–533.
- [20] Koszul, J.L. Ouverts convexes homogènes des espaces affines. Math. Z. 1962, 79, 254–259.
- [21] Koszul, J.L. Variétés localement plates et convexité. Osaka. J. Math. 1965, 2, 285–290.
- [22] Koszul, J.L., Introduction to Symplectic Geometry, Science Press, Beijing, 1986 (Chinese)

THALES

References

- [23] Koszul, J.L. Déformations des variétés localement plates. Ann. Inst. Fourier 1968, 18, 103–114.
- [24] Koszul, J.L. Trajectoires Convexes de Groupes Affines Unimodulaires. In Essays on Topology and Related Topics; Springer: Berlin, Germany, 1970; pp. 105–110.
- [25] Libermann, P.; Marle, C.M. Symplectic Geometry and Analytical Mechanics; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1987.
- [26] 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.
- [27] Marle, C.M. From Tools in Symplectic and Poisson Geometry to J.-M. Souriau's Theories of Statistical Mechanics and Thermodynamics. Entropy 2016, 18, 370.
- [28] Massieu, F. Sur les Fonctions caractéristiques des divers fluides. Comptes Rendus de l'Académie des Sciences 1869, 69, 858–862.
- [29] Massieu, F. Addition au précédent Mémoire sur les Fonctions caractéristiques. Comptes Rendus de l'Académie des Sciences 1869, 69, 1057–1061.
- [30] 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.l. - s.n., 1873
- [31] 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.
- [32] Nencka, H.; Streater, R.F. Information Geometry for some Lie algebras. Infin. Dimens. Anal. Quantum Probab. Relat. Top. 1999, 2, 441–460.

References

- [33] Poincaré, H. Sur les tentatives d'explication mécanique des principes de la thermodynamique. Comptes rendus de l'Académie des sciences 1889, 108, 550–553.
- [34] Poincaré, H. Thermodynamique, Cours de Physique Mathématique; Carré, G., Ed.; 1892; en ligne: http://gallica.bnf.fr/ark:/12148/bpt6k2048983
- [35] Poincaré, H. Calcul des Probabilités; Gauthier-Villars: Paris, France, 1896.
- [36] Poincaré, H. Réflexions sur la théorie cinétique des gaz. J. Phys. Theor. Appl. 1906, 5, 369–403.
- [37] 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.
- [38] 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
- [39] de Saxcé, G.; Vallée, C. Galilean Mechanics and Thermodynamics of Continua; Wiley-ISTE: London, UK, 2016.
- [40] Shima, H. The Geometry of Hessian Structures; World Scientific: Singapore, Singapore, 2007.
- [41] 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.
- [42] Souriau, J.M. Géométrie de l'espace de phases. Comm. Math. Phys. 1966, 374, 1–30
- [43] Souriau, J.M. Définition covariante des équilibres thermodynamiques. Suppl. Nuov. Cimento 1966, 1, pp.203–216. (In French)

THALES

References

- [44] Souriau, J.M. Structure des systèmes dynamiques; Editions Jacques Gabay: Paris, France, 1970.
- [45] 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.
- [46] Souriau, J.M. Géométrie Symplectique et Physique Mathématique; Éditions du C.N.R.S.: Paris, France, 1975.
- [47] Souriau, J.M. Thermodynamique Relativiste des Fluides; Centre de Physique Théorique: Marseille, France, 1977.
- [48] 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.
- [49] 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.
- [50] Souriau, J.M. Mécanique classique et géométrie symplectique. CNRS Marseille. Cent. Phys. Théor. 1984, Report ref. CPT-84/PE-1695.
- [51] Souriau, J.M. On Geometric Mechanics. Discret. Cont. Dyn. Syst. J. 2007, 19, 595–607.
- [52] Vallée, C. Relativistic thermodynamics of continua. Int. J. Eng. Sci. 1981, 19, 589–601.
- [53] Vallée, C.; Lerintiu, C., Convex analysis and entropy calculation in statistical mechanics, Proceedings of A. Razmadze mathematical Institute 137, 2005, pp. 111-129.
- [54] Vey, J. Sur les Automorphismes Affines des Ouverts Convexes Saillants. Annali della Scuola Normale Superiore di Pisa, Classe di Science 1970, 24, 641–665.
- [55] 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

References

- [56] Pierre Cartier, Inégalités de corrélation en mécanique statistique, Séminaire N. Bourbaki, 1972-1973, exp. N°431, p.242-264
- [57] Pierre Cartier, La durée cosmologique est-elle thermodynamique ?, séminaire ENS «Temps et durée entre science et philosophie », 2008
- [58] Pierre Cartier, Le Calcul des Probabilités de Poincaré, IHES, 2006
- [59] Moreau, J.J. Fonctions convexes duales et points proximaux dans un espace hilbertien. C. R. Acad. Sci. 1962, 255, 2897–2899
- [60] Guichardet, A. Cohomologie des Groups Topologiques et des Algèbres de Lie; Ce-dic/Fernand Nathan: Paris, France, 1980
- [61] Paul Belgodère, Extremales d'une intégrale de surface, note présentée par Elie Cartan, CRAS, séance su 11 Septembre 1944
- [62]Paulette Libermann, Legendre foliations on contact manifolds, Differential Geometry and Its Applications, n°1, pp.57-76, 1991
- [63] Paul Gauduchon, Calabi's extremal Kähler metrics: An elementary introduction, Preprint, Ecole Polytechnque
- [64] Souriau J.-M., Groupes différentiels, Lect. Notes in Math., 836, pp. 91–128. Springer Ver-lag, New-York, 1980.
- [65] Souriau J.-M., Un algorithme générateur de structures quantiques, Soc. Math. Fr., As-trisque, Actes du colloque Henri Cartan Lyon, 1985, 341–399
- [66] Souriau, J.-M., Groupes différentiels de physique mathématique, in "Feuilletages et Quan-tification Géometrique," Dazord, P. and N. Desolneux-Moulis (Eds.), Journ. lyonnaises Soc. math. France 1983, Sémin. sud-rhodanien de géom. II,Hermann, Paris, 1984; 73–119.

THALES

References

- [67] Donato P., Géometrie des orbites coadjointes des groupes de difféomorphismes, Volume 1416 of the series Lecture Notes in Mathematics , Géométrie Symplectique et Mécanique, pp 84-104, 2006
- [68] Paul Donato. Revêtement et groupe fondamental des espaces différentiels homogènes, Thèse de doctorat d'état, Université de Provence, Marseille, 1984.
- [69] Dazord P., Lie groups and algebras in infinite dimension: a new approach, Symplectic geometry and quantization, Providence, R.I. : American Mathematical Society, 1994
- [70] Hector, G., and E. Macias-Virgos, Diffeological groups, Research Exp. Math. 25 (2002), 247–260.
- [71] M. V. Losik, Fréchet manifolds as diffeological spaces, Soviet. Math. 5 (1992)
- [72] Alexandre A. Kirillov. Elements de la théorie des représentations. Editions MIR, Moscou, 1974.
- [73] Simon K. Donaldson. Moment maps and diffeomorphisms. Asian Journal of Math, vol. 3, pp. 1–16, 1999.
- [74] M. Condevaux, P. Dazord and P. Molino. Géométrie du moment. Travaux du Séminaire Sud-Rhodanien de Géométrie, I, Publ. D´ep. Math. Nouvelle Sér. B 88-1, Univ. Claude-Bernard, pp. 131 – 160, Lyon, 1988.
- [75] Augustin Banyaga. Sur la structure du groupe des difféomorphismes qui préservent une forme symplectique, Comment. Math. Helv., volume 53, pages 174–227, 1978.
- [76] Patrick Iglesias-Zemmour, The Moment Maps in Diffeology
- [77] Helein, F., Multisymplectic geometry: some perspectives, Unfold mechanics for sound & music, IRCAM, http://medias.ircam.fr/x17b9b0
- [78] Gallissot, François, Les formes extérieures en mécanique, Annales de l'institut Fourier, Tome 4 (1952) , p. 145-297
- [79] Gallissot, François, Les formes extérieures et la mécanique des milieux continusAnnales de l'institut Fourier, Tome 8 (1958) , p. 291-335

References

- [78] V. Volterra, Sulle equazioni differenziali che provengono da questiono di calcolo delle va-riazioni, Rend. Cont. Acad. Lincei, ser. IV, vol. VI, 1890, 42–54.
- [79] V. Volterra, Sopra una estensione della teoria Jacobi–Hamilton del calcolo delle varizioni, Rend. Cont. Acad. Lincei, ser. IV, vol. VI, 1890, 127–138.
- [80] J. Kijowski, A finite dimensional canonical formalism in the classical field theory, Comm. Math. Phys. 30 (1973), 99-128.
- [81] J. Kijowski, Multiphase spaces and gauge in the calculus of variations, Bull. de l'Acad. Polon. des Sci., Série sci. Math., Astr. et Phys. XXII (1974), 1219-1225.
- [82] J. Kijowski, W. Szczyrba, A canonical structure for classical field theories, Commun. Math Phys. 46 (1976), 183–206.
- [83] Jaworski W., Ingarden R.S., On the partition function in information thermodynamics with higher order temperatures, Bull. Acad.. Polon. Sci. Sér. phys. astr., pp. 28-119, 1980
- [84] Jaworski, W. On information thermodynamics with temperatures of the second order (in Polish), Master Thesis, Inst. Phys., N. Copernicus Univ., Torun, 1981
- [85] Jaworski, W., Information thermodynamics with the second order temperatures for the simplest classical systems, Acta Phys. Polon. A60, pp. 645-659., 1981
- [86] Jaworski, W., On the thermodynamic limit in information thermodynamics with higher-order temperatures, Acta Phys. Polon., A63, pp. 3-19, 1983
- [87] Jaworski, W. , Investigation of the thermodynamic limit for the states maximizing entropy under auxiliary conditions for higher-order statistical moments (in Polish). Doctor Thesis. Inst. Phys., N. Copernicus Univ. Torun, 1983
- [88] Jaworski, W., Higher-order moments and the maximum entropy inference: the thermody-namical limit approach, J. Phys. A: Math. Gen. 20, pp. 915-926, 1987

References

- [89] Ingarden, R.S. and Kossakowski A., Statistical thermodynamics with higher order temper-atures for ideal gases of bosons and fermions, Acta Phys. Polon. 28, pp.499•-511., 1965
- [90] R.S. Ingarden, A. Kossakowski, M. Ohya, Information Dynamics and Open Systems: Classical and Quantum Approach, Fundamental Theories of Physics, vol. 86, Springer, 1997
- [91] Ingarden, R.S. , Towards mesoscopic thermodynamics: small systems in higher-order states, Open Sys. Information Dynamics. 1, 75-102, 1993, errata 309.
- [92] Ingarden, R.S., and Nakagomi, T., The second order extension of the Gibbs state, Open Sys. Lnforlllalion, Jyn. 1, 259-:268., 1992
- [93] Ingarden, H.S., Meller J. , Temperatures in linguistics as a model of thermodynamics, Open Sys. Information Dynamics 2, pp.211-230, 1991
- [94] Nakagomi, T.; Mesoscopic version of thermodynamic equilibrium condition. Another ap-proach to higher order temperatures. Open systems Dyn., 1, pp. 233-2•11, 1992
- [95] J Casas-Vazquez1 and D Jou, Temperature in non-equilibrium states: a review of open problems and current proposals, Rep. Prog. Phys. 66 (2003) 1937–2023
- [96] Gunther C., The polysymplectic Hamiltonian formalism in field theory and calculus of variations I: The local case, J. Differential Geom. n°25, pp. 23-53, 1987
- [97] Munteanu F., Rey A.M., Salgado M., The Günther's formalism in classical field theory: momentum map and reduction, J. Math. Phys. bf, n°45, vol. 5, pp.1730–1751, 2004
- [98] Awane A., k-symplectic structures, J. Math. Phys., vol. 33, pp. 4046-4052, 1992
- [99] A. Awane, M. Goze, Pfaffian systems, k-symplectic systems. Springer , 2000

THALES