Jean-Louis Koszul and the Elementary Structures of Information Geometry

Frédéric Barbaresco

Abstract This paper is a scientific exegesis and admiration of Jean-Louis Koszul's works on homogeneous bounded domains that have appeared over time as elementary structures of Information Geometry. Koszul has introduced fundamental tools to characterize the geometry of sharp convex cones, as Koszul-Vinberg characteristic Function, Koszul Forms, and affine representation of Lie Algebra and Lie Group. The 2nd Koszul form is an extension of classical Fisher metric. Koszul theory of hessian structures and Koszul forms could be considered as main foundation and pillars of Information Geometry.

Keywords Koszul-Vinberg characteristic function · Koszul forms
Affine representation of lie algebra and lie group
Homogeneous bounded domains

1 Preamble

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

In this article, I will pay tribute to a part of Professor Jean-Louis Koszul's work and fundamental and deep contributions of this great algebraist and geometer in the field of Information Geometry, which have many applications in the domain of applied mathematics, and in the emerging applications of Artificial Intelligence where the most efficient and robust algorithms are based on the natural gradient of the information geometry deduced from the Fisher matrix, as Yann Ollivier recently showed [1, 2].

F. Barbaresco (✉)
Thales Land & Air Systems, Voie Pierre-Gilles de Gennes, F91470 Limours, France
e-mail: frederic.barbaresco@thalesgroup.com

F. Nielsen (ed.), *Geometric Structures of Information*, Signals and Communication Technology, https://doi.org/10.1007/978-3-030-02520-5_12

After the seminal papers of Fréchet [3], Rao [4] and Chentsov [5], many mathematicians and physicists have studied Information Geometry. One can quote in mathematics, the works of Amari [6, 7] in the 80 s, which does not refer to the Koszul publications of the 50s and 60s where Koszul introduced the elementary structures of the Hessian geometries, and generalized the Fisher metric for homogeneous convex domains. In the physical field, many physicists have also addressed Information Geometry, without references to Koszul. Weinhold [8] in 1976 and Ruppeiner [9] in 1979 empirically introduced the inverse dual metric defined by the Hessian of Entropy, or Ingarden [10, 11] in 1981 in Statistical Physics. Mrugala [12, 13] in 1978, and Janyszek [14] in 1989, tried to geometrize Thermodynamics by jointly addressing Information Geometry and Contact Geometry. All these authors were not familiar with Representations Theory introduced by Kirillov, and more particularly the affine representation of Lie groups and Lie algebras, used and developed by Koszul in mathematics and by Souriau in statistical mechanics [79–84]. It thus appears that the first foundations of the information geometry goes back to Fréchet's paper of 1943 [3] (and his Lecture given during the winter of 1939 at the Institut Henri Poincaré), who first introduced the Clairaut(-Legendre) equation (fundamental equation in Information Geometry between dual potentials) and Fisher metric as the Hessian of a convex function. This Fréchet's seminal work was followed by Koszul's 50's papers [15, 16] which introduced new forms that generalize Fisher metric for sharp convex cones. It was not until 1969 that Souriau completed this extension in the framework of the Lie Group Thermodynamics with a cohomological definition of Fisher metric [17]. This last extension was developed by Koszul at the beginning of 80's in his Lecture "Introduction to Symplectic Geometry" [18]. I will conclude this survey by making reference to Balian [19], who has developed during 80's Information Geometry in Quantum Physics with a Quantum Fisher metric given by Von Neumann Entropy hessian [20].

Inspired by the French mathematical tradition, and the teachings of his master Elie Cartan (Koszul was PhD student of Henri Cartan but was greatly influenced by Elie Cartan), Jean-Louis Koszul was a real "avant-garde", if we take the definition given by Clausewitz « *An avant-garde is a group of units intended to move in front of the army to: explore the terrain to avoid surprises, quickly occupy the strong positions of the battlefield (high points), screen and contain the enemy the time the army can deploy*". Indeed, Jean-Louis Koszul was a pioneer, who explored and cleared many areas of mathematics, detailed in the book "Selected papers of JL Koszul" [21]. What I will expose, in this paper, is therefore only one part of his work which concerns homogeneous bounded domains geometry, from seminal Elie Cartan's earlier work on symmetric bounded domains. In a letter from André Weil to Henri Cartan, cited in the proceedings of the conference "*Elie Cartan and today's mathematics*" in 1984, it says "*As to the symmetrical spaces, and more particularly to the symmetric bounded domains at the birth of which you contributed, I have kept alive the memory of the satisfaction I felt in finding some incarnations in Siegel from his first works on quadratic forms, and later to convince Siegel of the value of your father's ideas on the subject*". At this 1984 conference, two disciples of Elie Cartan gave a conference, Jean-Louis Koszul [22] and Jean-Marie Souriau (Fig. 1).

Fig. 1 (on the left) Jean-Louis Koszul student at ENS ULM in 1940, (on the right) Jean-Louis Koszul at GSI'13 "Geometric Science of Information" conference at the École des Mines de Paris August 2013

In the book "*Selected papers of JL Koszul*" [21], Koszul summarizes the work, I will detail in the following: "*It is with the problem of the determination of the homogeneous bounded domains posed by E. Cartan around 1935 that are related [my papers]. The idea of approaching the question through invariant Hermitian forms already appears explicitly in Cartan. This leads to an algebraic approach which constitutes the essence of Cartan's work and which, with the Lie J-algebras, was pushed much further by the Russian School* [23–36]. *It is the work of Piatetski Shapiro on the Siegel domains, then those of E.B. Vinberg on the homogeneous cones that led me to the study of the affine transformation groups of the locally flat manifolds and in particular to the convexity criteria related to invariant forms"*. In particular, J.L. Koszul source of inspiration is given in this last sentence of Elie Cartan's 1935 article [37]:

"*It is clear that if one could demonstrate that all homogeneous domains whose form* $\Phi = \sum_{i,j} \frac{\partial^2 \log K(z,z^*)}{\partial z_i \partial z_j^*} dz_i dz_j^*$ *is positive definite are symmetric, the whole theory of homogeneous bounded domains would be elucidated. This is a problem of Hermitian geometry certainly very interesting*". It was not until 1953 that the classification of non-Riemannian symmetric spaces has been achieved by Marcel Berger [38]. The work of Koszul has also been extended and deepened by one of his student Jacques Vey in [39, 40]. Jacques Vey has transposed the notion of hyperbolicity, developed

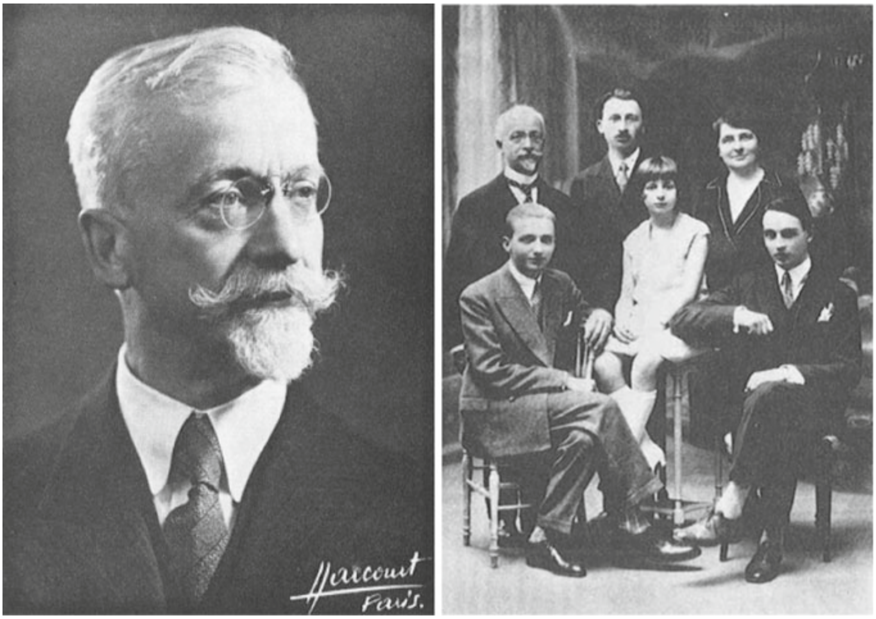


Fig. 2 (on the left) Professor Elie Cartan, (on the right) the Cartan family

by W. Kaup for Riemann surfaces, into the category of differentiable manifolds with flat linear connection (locally flat manifolds), which makes it possible to completely characterize the locally flat manifolds admitting as universal covering a convex open sharp cone of R^n, which had been studied by Koszul in [41]. The links between Koszul's work and those of Ernest B. Vinberg [23–30] were recently developed at the conference "*Transformation groups 2017*" in Moscow dedicated to the 80$^{\text{th}}$ anniversary of Professor EB Vinberg, in Dmitri Alekseevsky's talk on "*Vinberg's theory of homogeneous convex cones: developments and applications*" [42]. Koszul and Vinberg are actually associated with the concept of Koszul-Vinberg's characteristic function on convex cones, which I will develop later in the paper. Koszul introduced the so-called "*Koszul forms*" and a canonical metric given by the Hessian of the opposite of the logarithm of this Koszul-Vinberg characteristic function, from which I will show the links with Fisher's metric in Information Geometry, and its extension (Fig. 2).

Professor Koszul's main papers, which form the elementary structures of information geometry, are as follows:

- ***«Sur la forme hermitienne canonique des espaces homogènes complexes»*** [15] of 1955: Koszul considers the Hermitian structure of a homogeneous *G/B* manifold (*G* related Lie group and *B* a closed subgroup of *G*, associated, up to a constant factor, to the single invariant *G*, and to the invariant complex structure by the operations of *G*). Koszul says "*The interest of this form for the determination of* *homogeneous bounded domains has been emphasized by Elie Cartan: a necessary condition for G/B to be a bounded domain is indeed that this form is positive definite*". Koszul calculated this canonical form from infinitesimal data Lie algebra of *G,* the sub-algebra corresponding to *B* and an endomorphism algebra defining the invariant complex structure of *G/B.* The results obtained by Koszul proved that the homogeneous bounded domains whose group of automorphisms is semi-simple are bounded symmetric domains in the sense of Elie Cartan. Koszul also refers to André Lichnerowicz's work on Kählerian homogeneous spaces [43]. In this seminal paper, Koszul also introduced a left invariant form of degree 1 on *G*: $\Psi(X) = Tr_{g/b}[ad(JX) - J.ad(X)] \quad \forall X \in g$ with *J* an endomorphism of the Lie algebra space and the trace $Tr_{g/b}[.]$ corresponding to that of the endomorphism *g/b.* The Kähler form of the canonical Hermitian form is given by the differential of $-1/4\Psi(X)$ of this form of degree 1.

- ***«Exposés sur les espaces homogènes symétriques»*** [16] of 1959 is a Lecture written as part of a seminar held in September and October 1958 at the University of Sao Paulo, which details the determination of homogeneous bounded domains. He returned to [15] and showed that any symmetric bounded domain is a direct product of irreducible symmetric bounded domains, determined by Elie Cartan (4 classes corresponding to classical groups and 2 exceptional domains). For the study of irreducible symmetric bounded domains, Koszul refered to Elie Cartan, Carl-Ludwig Siegel and Loo-Keng Hua. Koszul illustrated the subject with two particular cases, the half-plane of Poincaré and the half-space of Siegel, and showed that with its trace formula of endomorphism *g/b,* he found that the canonical Kähler hermitian form and the associated metrics are the same as those introduced by Henri Poincaré and Carl-Ludwig Siegel [44] (who introduced them as invariant metric under action of the automorphisms of these spaces).
- ***«Domaines bornées homogènes et orbites de groupes de transformations affines»*** [45] of 1961 is written by Koszul at the Institute for Advanced Study at Princeton during a stay funded by the National Science Foundation. On a complex homogeneous space, an invariant volume defines with the complex structure the canonical invariant Hermitian form introduced in [15]. If the homogeneous space is holomorphically isomorphic to a bounded domain of a space C^n, this Hermitian form is positive definite because it coincides with the Bergmann metric of the domain. Koszul demonstrated in this article the reciprocal of this proposition for a class of complex homogeneous spaces. This class consists of some open orbits of complex affine transformation groups and contains all homogeneous bounded domains. Koszul addressed again the problem of knowing if a complex homogeneous space, whose canonical Hermitian form is positive definite is isomorphic to a bounded domain, but via the study of the invariant bilinear form defined on a real homogeneous space by an invariant volume and an invariant flat connection. Koszul demonstrated that if this bilinear form is positive definite then the homogeneous space with its flat connection is isomorphic to a convex open domain containing no straight line in a real vector space and extended it to the initial problem for the complex homogeneous spaces obtained in defining a complex structure in the variety of vectors of a real homogeneous space provided with an invariant

flat connection. It is in this article that Koszul used the affine representation of Lie groups and algebras. By studying the open orbits of the affine representations, he introduced an affine representation of G, written (f, q), and the following equation setting f the linear representation of the Lie algebra g of G, defined by f and q the restriction to g and the differential of q (f and q are differential respectively of f and q):

$$f(X)q(Y) - f(Y)q(X) = q([X, Y]) \quad \forall X, Y \in \mathsf{g}$$
$$\text{with} \quad f : \mathsf{g} \rightarrow gl(E) \quad \text{and} \quad q : \mathsf{g} \mapsto E$$

- ***«Ouverts convexes homogènes des espaces affines»*** [46] of 1962. Koszul is interested in this paper by the structure of the convex open non-degenerate Ω (with no straight line) and homogeneous (the group of affine transformations of E leaving stable Ω operates transitively in Ω) in a real affine space of finite dimension. Koszul demonstrated that they can be all deduced from non-degenerate and homogeneous convex open cones built in [45]. He used for this the properties of the group of affine transformations leaving stable a non-degenerate convex open domain and an homogeneous domain.
- ***«Variétés localement plates et convexité»*** [41] of 1965. Koszul established the following theorem: let *M be* a locally related differentiable manifold. If the universal covering of M is isomorphic as a flat manifold with a convex open domain containing no straight line in a real affine space, then there exists on M a closed differential form α such that $D\alpha$ *(D* linear covariant derivative of zero torsion) is positive definite in all respects and which is invariant under every automorphism of M. If G is a group of automorphisms of M such that $G\backslash M$ is quasi-compact and if there exists on M a closed 1-differential form α invariant by G and such that $D\alpha$ is positive definite at any point, then the universal covering of M is isomorphic as a flat manifold with a convex open domain that does not contain a straight line in a real affine space.
- ***«Lectures on Groups of Transformations»*** [47] of 1965. This is lecture notes given by Koszul at Bombay "Tata Institute of Fundamental Research " on transformation groups. In particular in Chap. 6, Koszul studied discrete linear groups acting on convex open cones in vector spaces based on the work of C.L. Siegel (work on quadratic forms [48]). Koszul used what I will call in the following Koszul-Vinberg characteristic function on convex sharp cone.
- ***«Déformations des variétés localement plates»*** [49] of 1968. Koszul provided other proofs of theorems introduced in [41]. Koszul considered related differentiable manifolds of dimension n and TM the fibered space of M. The linear connections on M constitute a subspace of the space of the differentiable applications of the *TM*x*TM* fiber *product* in the space *T(TM)* of the *TM* vectors. Any locally flat connection D (the curvature and the torsion are zero) defines a locally flat connection on the covering of *M,* and is hyperbolic when universal covering of *M,* with this connection, is isomorphic to a sharp convex open domain (without straight lines) in R^n. Koszul showed that, if M is a compact manifold, for a locally

flat connection on M to be hyperbolic, it is necessary and sufficient that there exists a closed differential form of degree 1 on M whose covariant differential is positive definite.

- ***«Trajectoires Convexes de Groupes Affines Unimodulaires»*** [50] in 1970 . Koszul demonstrated that a convex sharp open domain in R^n that admits a unimodular transitive group of affine automorphisms is an auto-dual cone. This is a more geometric demonstration of the results shown by Ernest Vinberg [29] on the automorphisms of convex cones.

The elementary geometric structures discovered by Jean-Louis Koszul are the foundations of Information Geometry. These links were first established by Professor Hirohiko Shima [51–56]. These links were particularly crystallized in Shima book 2007 "*The Geometry of Hessian Structures*" [57], which is dedicated to Professor Koszul. The origin of this work followed the visit of Koszul in Japan in 1964, for a mission coordinated with the French government. Koszul taught lectures on the theory of flat manifolds at Osaka University. Hirohiko Shima was then a student and attended these lectures with the teachers Matsushima and Murakami. This lecture was at the origin of the notion of Hessian structures and the beginning of the works of Hirohiko Shima. Henri Cartan noted concerning Koszul's ties with Japan, "*Koszul has attracted eminent mathematicians from abroad to Strasbourg and Grenoble. I would like to mention in particular the links he has established with representatives of the Japanese School of Differential Geometry*". Shima's book [57] is a systematic introduction to the theory of Hessian structures (provided by a pair of a flat connection D and an Hessian metric g). Koszul studied flat manifolds with a closed 1-form α, such that $D\alpha$ be positive definite, where $D\alpha$ is a hessian metric. However, not all Hessian metrics are globally of the form $g = D\alpha$. Shima introduces the notion of Codazzi structure for a pair (D,g), with D a torsion-free connection, which verifies the Codazzi equation $(D_X g)(Y, Z) = (D_Y g)(X, Z)$. A Hessian structure is a Codazzi structure for which connection D is flat. This is an extension of Riemannian geometry. It is then possible to define a connection D' and a dual Codazzi structure (D',g) with $D' = \nabla - D$ where ∇ is the Levi-Civita connection. For a hessian structure (D, g) with $g = Dd\varphi$, the dual Codazzi structure $\left(D', g\right)$ is also a Hessian structure and $g = D'd\varphi'$, where φ' is the Legendre transform of $\varphi : \varphi' = \sum_i x^i \frac{\partial \varphi}{\partial x^i} - \varphi$. Shima observed that Information Geometry framework could be introduced by dual connections, and not only founded on Fréchet, Rao and Chentsov works [5]. A hessian structure (D, g) is of Koszul type, if there is a closed 1-form ω as $g = D\omega$. Using D and the volume element of g, Koszul introduced a 2nd form, which plays a similar role to the Ricci tensor for a Kählerian metric. Let υ be the volume element of g, we define a closed 1-form α such that $D_X \upsilon = \alpha(X)\upsilon$ and a symmetric bilinear form $\gamma = D\alpha$. In the following, α and γ forms are called 1st and 2nd form of Koszul for Hessian structure (D, g). We can consider the forms associated with the Hessian dual structure (D', g) by $\alpha' = -\alpha$ and $\gamma' = \gamma - 2\nabla\alpha$. In the case of a homogeneous regular convex cone Ω, with D the canonical flat connection of the ambient vector space, the Koszul forms α and γ for the canonical Hessian structure $(D, g = Dd\psi)$

Fig. 3 From left to right, Jean-Louis Koszul, Hirohiko Shima and Michel Nguiffo Boyom at GSI'13 (Geometric Science of Information) conference at the École des Mines of Paris in August 2013

are given by $\alpha = d \log \psi$ and $\gamma = g$. The volume element υ determined by g is invariant under the action of the group of automorphisms G of Ω.

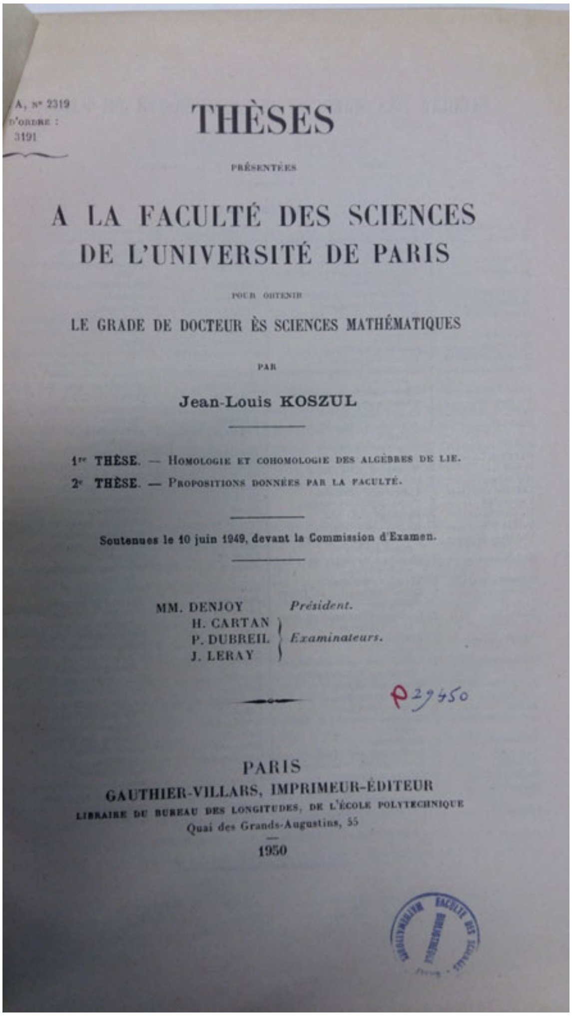
Jean-Louis Koszul attended the 1st GSI "*Geometric Science of Information*" conference in August 2013 at the Ecole des Mines in Paris, where he attended the presentation of Hirohiko Shima, given for his honor on the topic "*Geometry of Hessian Structures* " [58]. In the photo below, we can see from left to right, Jean-Louis Koszul, Hirohiko Shima and Michel Nguiffo Boyom. Professor Michel Boyom has extensively studied and developed, at the University of Montpellier, Koszul models [59–66] in relation to symplectic flat affine manifolds and to the cohomology of Koszul-Vinberg algebras (KV Cohomology). Professor Boyom with his PhD student Byande [67, 68] have explored other links with Information Geometry. André Lichnerowicz worked in parallel on a closed topic about homogeneous Kähler manifolds [69] (Fig. 3).

2 Biographical Reminder of Jean-Louis Koszul Scientific Life

Jean Louis André Stanislas Koszul born in Strasbourg in 1921, is the child of a family of four (with three older sisters, Marie Andrée, Antoinette and Jeanne). He is the son of André Koszul (born in Roubaix on November 19th 1878, professor at the

Strasbourg university), and Marie Fontaine (born in Lyon on June 19th 1887), who was a friend of Henri Cartan's mother. Henri Cartan writes on this friendship "*My mother in her youth, had been a close friend of the one who was to become Jean-Louis Koszul's mother*" [70]. His paternal grandparents were Julien Stanislas Koszul and Hélène Ludivine Rosalie Marie Salomé. He attended high school in Fustel-de-Coulanges in Strasbourg and the Faculty of Science in Strasbourg and in Paris. He entered ENS Ulm in the class of 1940 and defended his thesis with Henri Cartan. Henri Cartan noted "*This promotion included other mathematicians like Belgodère or Godement, and also physicists and some chemists, like Marc Julia and Raimond Castaing*" [70] (for the anecdote, the maiden name of my wife Anne, is *Belgodère*, with a filial link with Paul Belgodère of the Koszul ENS promotion). Jean-Louis Koszul married on July 17th 1948 with Denise Reyss-Brion, student of ENS Sèvres, entered in 1941. They have three children, Michel (married to Christine Duchemin), Anne (wife of Stanislas Crouzier) and Bertrand. He then taught in Strasbourg and was appointed Associate Professor at the University of Strasbourg in 1949, and had for colleagues R. Thom, M. Berger and B. Malgrange. He was promoted to professor status in 1956. He became a member of Bourbaki with the 2nd generation, J. Dixmier, R. Godement, S. Eilenberg, P. Samuel, J. P. Serre and L. Schwartz. Henri Cartan remarked in [70] "*In the vehement discussions within Bourbaki, Koszul was not one of those who spoke loudly; but we learned to listen to him because we knew that if he opened his mouth he had something to say*". About this Koszul's period at Strasbourg University, Pierre Cartier [71] said "*When I arrived in Strasbourg, Koszul was returning from a year spent in Institute for Advanced Studies in Princeton, and he was after the departure of Ehresman and Lichnerowicz to Paris the paternal figure of the Department of Mathematics (despite his young age). I am not sure of his intimate convictions, but he represented for me a typical figure of this Alsatian Protestantism, which I frequented at the time. He shared the seriousness, the honesty, the common sense and the balance. In particular, he knew how to resist the academic attraction of Paris. He left us after 2 years to go to Grenoble, in a maneuver uncommon at the time of exchange of positions with Georges Reeb*". He became Senior Lecturer at the University of Grenoble in 1963, and then an honorary professor at the Joseph Fourier University [72] and integrated in Fourier Institute led by C. Chabauty. During this period, B. Malgrange [73] remembered Koszul seminar on "algebra and geometry" with his three students J. Vey, D. Luna and J. Helmstetter. In Grenoble, he practiced mountaineering and was a member of the French Alpine Club. Koszul was awarded by Jaffré Prize in 1975 and was elected correspondent at the Academy of Sciences on January 28th 1980. Koszul was one of the CIRM conference center founder at Luminy. The following year, he was elected to the Academy of São Paulo. Jean-Louis Koszul died on January 12th 2018, at the age of 97.

As early as 1947, Jean-Louis Koszul published three articles in CRAS of the Academy of Sciences, on the Betti number of a simple compact Lie group, on cohomology rings, generalizing ideas of Jean-Leray, and finally on the homology of homogeneous spaces. Koszul's thesis, defended in June 10th 1949 under the direction of Henri Cartan, dealt with the homology and cohomology of Lie algebras [74]. The jury was composed of M. Denjoy (President), J. Leray, P. Dubreil and H. Cartan.


A, n° 2319
D'ORDRE :
3191

THÈSES

PRÉSENTÉES

A LA FACULTÉ DES SCIENCES
DE L'UNIVERSITÉ DE PARIS

POUR OBTENIR

LE GRADE DE DOCTEUR ÈS SCIENCES MATHÉMATIQUES

PAR

Jean-Louis KOSZUL

1re THÈSE. — Homologie et cohomologie des algèbres de Lie.
2e THÈSE. — Propositions données par la Faculté.

Soutenues le 10 juin 1949, devant la Commission d'Examen.

MM. DENJOY *Président.*
H. CARTAN
P. DUBREIL *Examinateurs.*
J. LERAY

PARIS
GAUTHIER-VILLARS, IMPRIMEUR-ÉDITEUR
LIBRAIRE DU BUREAU DES LONGITUDES, DE L'ÉCOLE POLYTECHNIQUE
Quai des Grands-Augustins, 55

1950

Fig. 4 Cover page of Koszul's PhD report defended June 10th 1949 with a Jury composed of Professors Arnaud Denjoy, Henri Cartan, Paul Dubreil and Jean Leray, published in [74]

Under the title "Works of Koszul I, II and III", Henri Cartan reported Koszul's PhD results to Bourbaki seminar [75–77]. See also, André Haefliger paper [78] (Fig. 4).

In 1987, an International Symposium on Geometry was held in Grenoble in honor of Jean-Louis Koszul, whose proceedings were published in "*les Annales de l'Institut Fourier*", Volume 37, No. 4. This conference began with a presentation by Henri Cartan, who remembered the mention given to Koszul for his aggregation [70]: "*Distinguished Spirit; he is successful in his problems. Should beware, orally, of overly systematic trends. A little less subtle complications, baroque ideas, a little more common sense and balance would be desirable*". About his supervision of Koszul's PhD, Henri Cartan writed "*Why did he turn to guide him (so-called)? Is it because he found inspiration in Elie Cartan's work on the topology of Lie groups?*

Perhaps he was surprised to note that mathematical knowledge is not necessarily transmitted by descent. In any case, he helped me to better know what my father had brought to the theory" [70]. On the work of Koszul algebrisation, Henri Cartan notes " *Koszul was the first to give a precise algebraic formalization of the situation studied by Leray in his 1946 publication, which became the theory of the spectral sequence. It took a good deal of insight to unravel what lay behind Leray's study. In this respect, Koszul's Note in the July 1947 CRAS is of historical significance.*" [70]. From June 26th to July 2nd 1947, CNRS, received an International conference in Paris, on "*Algebraic Topology*". This was the first postwar international diffusion of Leray's ideas. Koszul writes about this lecture "*I can still see Leray putting his chalk at the end of his talk by saying (modestly?) that he definitely did not understand anything about Algebraic Topology*". In writing his lectures at the Collège de France, Leray adopted the algebraic presentation of the spectral suite elaborated by Koszul. As early as 1950, J.P. Serre used the term "Leray-Koszul suite". Speaking of Leray, Koszul wrote "*around 1955 I remember asking him what had put him on the path of what he called the ring of homology of a representation in his Notes to the CRAS of 1946. His answer was Künneth's theorem; I could not find out more*". The sheaf theory, introduced by Jean-Leray, followed in 1947, at the same time as the spectral sequences.

In 1950, Koszul published an important book of 62 pages entitled "*Homology and Cohomology of Lie Algebras*" [74] based on his PhD work, in which he studied the links between homology and cohomology (with real coefficients) of a compact connected Lie group and purely algebraic problems of Lie algebra. Koszul then gave a lecture in São Paulo on the topic "*sheaves and cohomology*". The superb lecture notes were published in 1957 and dealt with the cohomology of Čech with coefficients in a sheaf. In the autumn of 1958, he again organized a series of seminars in São Paulo, this time on symmetric spaces [16]. R. Bott commented on these seminars "*very pleasant. The pace is fast, and the considerable material is covered elegantly. In addition to the more or less standard theorems on symmetric spaces, the author discusses the geometry of geodesics, Bergmann's metrics, and finally studies the bounded domains with many details*". In the mid-1960s, Koszul taught at the Tata Institute in Bombay on transformation groups [47] and on fiber bundles and differential geometry. The second lecture dealt with the theory of connections and the lecture notes were published in 1965. In 1986 he published "*Introduction to symplectic geometry*" [18] following a Chinese course in China (with the agreement of Jean-Louis Koszul given in 2017, this lecture given at the University of Nanjing will be translated into English by Springer and will be published in 2018). This book takes up and develops works of Jean-Marie Souriau [17, 79] on homogeneous symplectic manifolds and the affine representation of Lie algebras and Lie groups in geometric mechanics (another fundamental source of Information Geometry structures extended on homogeneous varieties [80–84]). Chuan Yu Ma writes in a review, on this latest book in Chinese, that "*This work coincided with developments in the field of analytical mechanics. Many new ideas have also been derived using a wide variety of notions of modern algebra, differential geometry, Lie groups, functional analysis, differentiable manifolds, and representation theory. [Koszul's book] emphasizes the* *differential-geometric and topological properties of symplectic manifolds. It gives a modern treatment of the subject that is useful for beginners as well as for experts*".

In 1994, in [21], a comment by Koszul explains the problems he was preoccupied with when he invented what is now called the "*Koszul complex*". This was introduced to define a theory of cohomology for Lie algebras and proved to be a general structure useful in homological algebra.

3 Koszul-Vinberg Characteristic Function, Koszul Forms and Maximum Entropy Density

Through the study of the geometry of bounded homogeneous domains initiated by Elie Cartan [37, 85], Jean-Louis Koszul discovered that the elementary structures are associated with Hessian manifolds on sharp convex cones [15, 16, 41, 45–47, 49, 50]. In 1935, Elie Cartan proved in [37] that the symmetric homogeneous irreducible bounded domains could be reduced to 6 classes, 4 canonical models and 2 exceptional cases. Ilya Piatetski-Shapiro [31–35], after Luogeng Hua [86], extended Siegel's description [44, 48] to other symmetric spaces, and showed by a counterexample that Elie Cartan's conjecture, that all transitive domains are symmetrical, was false. At the same time, Ernest B. Vinberg [23–30] worked on the theory of homogeneous convex cones and the construction of Siegel domains [44, 48]. More recently, the classical complex symmetric spaces were studied by F. Berezin [87, 88] in the context of quantification. In parallel, O.S. Rothaus [89] and Piatetski-Shapiro [31–35] with Karpelevitch, explored the underlying geometry of these complexes homogeneous fields, and more particularly the fibration areas on the components of the shilov boundary. In Italy, I note the work of E. Vessentini [90] and U. Sampieri [91, 92]. The Siegel domains, which fit into these classes of structures, nowadays play an important role in the processing of radar spatio-temporal signals and, more broadly, in learning from structured covariance matrices.

Jean-Louis Koszul and Ernest B. Vinberg have introduced a hessian metric invariant by the group of linear automorphisms on a sharp convex cone Ω through a function, called characteristic function ψ. In the following Ω is a sharp convex cone in a vector space E of finite size on R (a convex cone is sharp if there is no straight lines). In dual space E^* of E, Ω^* is the set of linear strictly positive forms on $\overline{\Omega}-\{0\}$. Ω^*, dual cone of Ω, is also a sharp convex cone. If $\xi \in \Omega^*$, then intersection $\Omega \cap \{x \in E / \langle x, \xi \rangle = 1\}$ is bounded. $G = Aut(\Omega)$ is the group of linear transformation from *E that preserves* Ω (group of automorphisms). $G = Aut(\Omega)$ acts on Ω^* such that, $\forall g \in G = Aut(\Omega), \forall \xi \in E^*$ then $\bar{g}.\xi = \xi \circ g^{-1}$. Koszul introduce an integral, of Laplace kind, on sharp dual convex cone, as:

Koszul-Vinberg Characteristic definition:
Let $d\xi$ Lebesque measure on E^*, following integral:

$$\psi_\Omega(x) = \int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi \quad \forall x \in \Omega \tag{1}$$

with Ω^* the dual cone, is analytical function on Ω, with $\psi_\Omega(x) \in]0, +\infty[$, called Koszul-Vinberg characteristic function of cone Ω.

Nota: the logarithm of the characteristic function is called « barrier function » for convex optimization algorithms. Yurii Nesterov and Arkadii Nemirovskii [93] have proved in modern theory of « *interior point* », using function $\Theta_\Omega(x) = \log(vol_n\{s \in \Omega^* / \langle s, x\rangle \leq 1\})$, that all convex cones in R^n have a self-dual barrier, linked with Koszul characteristic function.

Koszul-Vinberg Characteristic function has the following properties:

- Bergman kernel of $\Omega + iR^{n+1}$ is written $K_\Omega(\mathrm{Re}(z))$ up to a constant. K_Ω is defined by integral:

$$K_\Omega(x) = \int_{\Omega^*} e^{-\langle \xi, x\rangle} \psi_{\Omega^*}(\xi)^{-1} d\xi \tag{2}$$

- ψ_Ω is an analytical function defined in the interior of Ω and $\psi_\Omega(x) \to +\infty$ when $x \to \partial\Omega$. If $g \in Aut(\Omega)$ then $\psi_\Omega(gx) = |\det g|^{-1} \psi_\Omega(x)$ and as $tI \in G = Aut(\Omega)$ for all $t > 0$, we have:

$$\psi_\Omega(tx) = \psi_\Omega(x)/t^n \tag{3}$$

- ψ_Ω is strictly log convex, such that $\phi_\Omega(x) = \log(\psi_\Omega(x))$ is strictly convex.

From this characteristic function, Koszul introduced two forms:

1st Koszul form $\boldsymbol{\alpha}$: Differential 1-form

$$\alpha = d\phi_\Omega = d \log \psi_\Omega = d\psi_\Omega / \psi_\Omega \tag{4}$$

is invariant with respect to all automorphisms $G = Aut(\Omega)$ of Ω. If $x \in \Omega$ and $u \in E$ then:

$$\langle \alpha_x, u\rangle = -\int_{\Omega^*} \langle \xi, u\rangle . e^{-\langle \xi, x\rangle} d\xi \text{ and } \alpha_x \in -\Omega^* \tag{5}$$

and

2nd Koszul form $\boldsymbol{\gamma}$: Differential symmetric 2-form

$$\gamma = D\alpha = Dd \log \psi_\Omega \tag{6}$$

is a bilinear symmetric positive definite form invariant with respect to the action of $G = Aut(\Omega)$ and $D\alpha > 0$

Positivity is given by Schwarz inequality and:

$$Dd\ \log \psi_\Omega(u,v) = \int\limits_{\Omega^*} \langle \xi, u\rangle \langle \xi, v\rangle e^{-\langle \xi, u\rangle} d\xi \tag{7}$$

Koszul has proved that from this 2nd form, we can introduce an invariant Riemannian metric with respect to the action of cone automorphisms:
Koszul Metric: $D\alpha$ defines a Riemannian invariant structure by $Aut(\Omega)$, and the Riemannian metric is given by:

$$g = Dd\ \log \psi_\Omega \tag{8}$$

$$(Dd\ \log \psi(x))(u) = \frac{1}{\psi(u)^2}\left[\int\limits_{\Omega^*} F(\xi)^2 d\xi . \int\limits_{\Omega^*} G(\xi)^2 d\xi - \left(\int\limits_{\Omega^*} F(\xi).G(\xi) d\xi\right)^2\right] > 0$$

$$\text{with } F(\xi) = e^{-\frac{1}{2}\langle x, y\rangle} \quad \text{and} \quad G(\xi) = e^{-\frac{1}{2}\langle x, \xi\rangle}\langle u, \xi\rangle \tag{9}$$

The positivity could be proved by using Schwarz inequality, and the following properties for the derivative given by $d\ \log \psi = \frac{d\psi}{\psi}$ and $Dd\ \log \psi = \frac{Dd\psi}{\psi} - \left(\frac{d\psi}{\psi}\right)^2$ where $(d\psi(x))(u) = -\int\limits_{\Omega^*} e^{-\langle x, \xi\rangle}\langle u, \xi\rangle d\xi$ and $(Dd\psi(x))(u) = -\int\limits_{\Omega^*} e^{-\langle x, \xi\rangle}\langle u, \xi\rangle^2 d\xi$.

Koszul uses this diffeomorphism to define dual coordinates:

$$x^* = -\alpha_x = -d\ \log \psi_\Omega(x) \tag{10}$$

with $\langle df(x), u\rangle = D_u f(x) = \frac{d}{dt}\Big|_{t=0} f(x + tu)$. When the cone Ω is symmetric, the map $x \mapsto x^* = -\alpha_x$ is a bijection and an isometry with only one fixed point (the manifold is a symmetric Riemannian space given by its isometry):

$$(x^*)^* = x, \left\langle x, x^*\right\rangle - n \text{ et } \psi_\Omega(x)\psi_{\Omega^*}(x^*) = cste \tag{11}$$

x^* is characterized by $x^* = \arg\min\{\psi(y)/y \in \Omega^*, \langle x, y\rangle = n\}$ and x^* is the gravity center of the transverse cut $\{y \in \Omega^*, \langle x, y\rangle = n\}$ of Ω^*:

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

$$\text{and} \quad \left\langle -x^*, h\right\rangle = d_h \log \psi_\Omega(x) = -\int\limits_{\Omega^*} \langle \xi, h\rangle e^{-\langle \xi, x\rangle} d\xi \Big/ \int\limits_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi \tag{12}$$

In [94–97], Misha Gromov was interested by these structures. If we set $\Phi(x) = -\log \psi_\Omega(x)$, Gromov has observed that $x^* - d\Phi(x)$ is an injection where the image

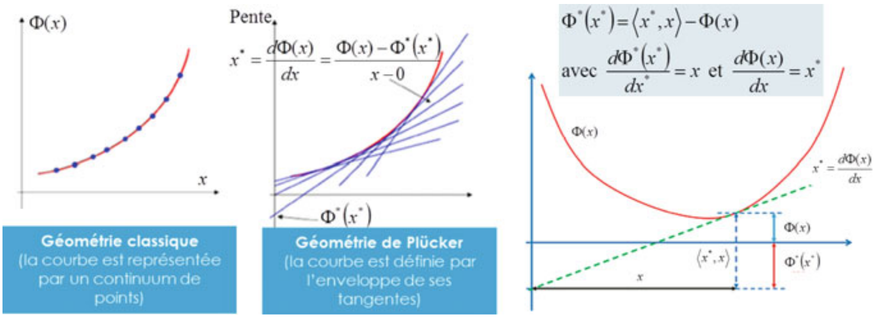


Fig. 5 Legendre transform and Plücker geometry

closure is equal to the convex envelop of the support and the volume of this envelop is the n-dimensionnel volume defined by the integral of hessian determinant of this function, $\Phi(x)$, where the map $\Phi \mapsto M(\Phi) = \int_{\Omega} \det(Hess(\Phi(x))).dx$ obeys a non-trivial inequality given by Brunn-Minkowsky:

$$[M(\Phi_1 + \Phi_2)]^{1/2} \geq [M(\Phi_1)]^{1/n} + [M(\Phi_2)]^{1/n} \tag{13}$$

These relations appear also in statistical physics. As the physicist Jean-Marie Souriau [17, 80–84, 98] did, it is indeed possible to define the concept of Shannon's Entropy via the Lengendre transform associated with the opposite of the logarithm of this Koszul-Vinberg characteristic function. Taking up the seminal ideas of François Massieu [99–102] in Thermodynamics (classmate of the Corps des Mines, it is François Massieu who influenced Henri Poincaré [103] who introduced the characteristic function in Probability, with a Laplace transform, and not a Fourier transform as did then Paul Levy), which were recently developed by Roger Balian in Quantum Physics [19, 20, 104–111], replacing Shannon Entropy by von Neumann Entropy. I will also note the work of Jean-Leray on the extensions of the Laplace transform in [112]. Starting from the characteristic function of Koszul-Vinberg, it is thus possible to introduce an entropy of Koszul defined as the Legendre transform of this function, which is the opposite of the logarithm of the characteristic function of Koszul-Vinberg (a logarithm lies the characteristic function of Massieu and the characteristic function of Koszul or Poincaré). Starting from the Koszul function, its Legendre transform gives a dual potential function in the dual coordinate system.x^* (Fig. 5):

$$\Phi^*(x^*) = \langle x, x^* \rangle - \Phi(x) \text{ with } x^* = D_x \Phi \text{ and } x = D_{x^*} \Phi^* \text{ where } \Phi(x) = -\log \psi_\Omega(x) \tag{14}$$

Concerning the Legendre transform [113], Darboux gives in his book an interpretation of Chasles: "*What comes back according to a remark of M. Chasles, to replace the surface with its polar reciprocal with respect to a paraboloid*". We have the same reference to polar reciprocal in "*Lessons on the calculus of variations*" by Jacques

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

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

Fig. 6 Legendre-Clairaut equation in 1943 Fréchet's paper

Hadamard, written by Maurice Fréchet (student of Hadamard), with references to M.E. Vessiot, which uses the "*figuratrice*", as polar reciprocal of the "*figurative*".

It is possible to express this Legendre transform only from the dual coordinate system x^*, using that $x = D_{x^*}\Phi^*$. We then obtain the Clairaut equation:

$$\Phi^*(x^*) - \left\langle (D_x\Phi)^{-1}(x^*), x^* \right\rangle - \Phi\left[(D_x\Phi)^{-1}(x^*)\right] \forall x^* \in \{D_x\Phi(x)/x \in \Omega\} \quad (15)$$

This equation was discovered by Maurice Fréchet in his 1943 paper [3] (see also in the appendix), in which he introduced for the first time the bound on the variance of any statistical estimator via the Fisher matrix, wrongly attributed to Cramer and Rao [4]. Fréchet was looking for "*distinguished densities*" [98], densities whose covariance matrix of the estimator of these parameters reaches this bound. Fréchet there showed that these densities were expressed while using this characteristic function $\Phi(x)$, and that these densities belong to the exponential densities family (Fig. 6).

Apparently, this discovery by Fréchet dates from winter of 1939, because Fréchet writes at the bottom of the page [3] "*The content of this dissertation formed part of our mathematical statistics Lecture at the Institut Henri Poincaré during the winter of 1939–1940. It is one of the chapters of the second edition (in preparation) of our ' Lessons in Mathematical Statistics', the first of which is 'Introduction: Preliminary Lecture on the Probability Calculation' (119 pages in quarto, typed in) has just been published at the University Documentation Center, Tournaments and Constans. Paris*". More details are given in appendix.

More recently Muriel Casalis [114, 115], the PhD student of Gérard Letac [116], has studied in her PhD, invariance of probability densities with respect to the affine group, and the links with densities of exponential families.

To make the link between the characteristic function of Koszul-Vinberg and Entropy of Shannon, we will detail the formulas of Koszul in the following developments. Using the fact that $-\langle \xi, x \rangle = \log e^{-\langle \xi, x \rangle}$, we can write:

$$-\left\langle x^*, x \right\rangle = \int_{\Omega^*} \log\, e^{-\langle \xi, x \rangle}.e^{-\langle \xi, x \rangle} d\xi \Big/ \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi \quad (16)$$

and then developing the Legendre transform to make appear the density of maximum entropy in $\Phi^*(x^*)$, and also the Shannon entropy:

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

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

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

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

$$\text{with} \int\limits_{\Omega^*} \frac{e^{-\langle \xi, x \rangle}}{\int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} d\xi = 1$$

$$\Phi^*(x^*) = \left[-\int\limits_{\Omega^*} \frac{e^{-\langle \xi, x \rangle}}{\int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} . \log \left(\frac{e^{-\langle \xi, x \rangle}}{\int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} \right) d\xi \right] \quad (17)$$

In this last equation, $p_x(\xi) = e^{-\langle \xi, x \rangle} / \int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$ plays the role of maximum entropy density as introduced by Jaynes [117–119] (also called, Gibbs density in Thermodynamics). I call the associated entropy, Koszul Entropy:

$$\Phi^* = -\int\limits_{\Omega^*} p_x(\xi) \log\, p_x(\xi) d\xi \quad (18)$$

with

$$p_x(\xi) = e^{-\langle \xi, x \rangle} / \int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi = e^{-\langle x, \xi \rangle - \log \int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} = e^{-\langle x, \xi \rangle + \Phi(x)} \text{and } x^* = \int\limits_{\Omega^*} \xi . p_x(\xi) d\xi \quad (19)$$

This Koszul density $p_x(\xi) = \frac{e^{-\langle \xi, x \rangle}}{\int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi}$ help us to develop the log likelihood:

$$\log\, p_x(\xi) = -\langle x, \xi \rangle - \log \int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi = -\langle x, \xi \rangle + \Phi(x) \quad (20)$$

and deduce from the expectation:

$$E_\xi\left[-\log p_x(\xi) \right] = \langle x, x^* \rangle - \Phi(x) \quad (21)$$

We also obtain the equation about normalization:

$$\Phi(x) = -\log \int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi = -\log \int_{\Omega^*} e^{-[\Phi^*(\xi)+\Phi(x)]} d\xi = \Phi(x) - \log \int_{\Omega^*} e^{-\Phi^*(\xi)} d\xi$$
$$\Rightarrow \int_{\Omega^*} e^{-\Phi^*(\xi)} d\xi = 1 \tag{22}$$

But we have to make appear the variable x^* in $\Phi^*(x^*)$. We have then to write:

$$\log\ p_x(\xi) = \log\ e^{-\langle x, \xi\rangle + \Phi(x)} = \log e^{-\Phi^*(\xi)} = -\Phi^*(\xi)$$
$$\Rightarrow \Phi^* = -\int_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi = \int_{\Omega^*} \Omega^*(\xi) p_x(\xi) d\xi = \Phi^*(x^*) \tag{23}$$

Last equality is true, if we have:

$$\int_{\Omega^*} \Phi^*(\xi) p_x(\xi) d\xi - \Phi^*\left(\int_{\Omega^*} \xi . p_x(\xi) d\xi\right) \text{ with } x^* = \int_{\Omega^*} \xi . p_x(\xi) d\xi \tag{24}$$

This last relation is associated to classical Jensen inequality. Equality is obtained for Maximum Entropy density for $x^* = D_x\Phi$ [120]:

$$\text{Legendre - Moreau Transform:} \quad \Phi^*(x^*) = \underset{x}{Sup}\left[\left\langle x, x^*\right\rangle - \Phi(x)\right]$$
$$\Rightarrow \begin{cases} \Phi^*(x^*) \geq \langle x, x^*\rangle - \Phi(x) \\ \Phi^*(x^*) \geq \int_{\Omega^*} \Phi^*(\xi) p_x(\xi) d\xi \end{cases} \Rightarrow \begin{cases} \Phi^*(x^*) \geq E\left[\Phi^*(\xi)\right] \\ \text{equality if } x^* = \frac{d\Phi}{dx} \end{cases} \tag{25}$$

We obtain for the maximum entropy density, the equality:

$$E\left[\Phi^*(\xi)\right] = \Phi^*(E[\xi]), \xi \in \Omega^* \tag{26}$$

To make the link between this Koszul model and maximum entropy density [121–123] introduced by Jaynes [117–119], I use previous notation and I look for the density $p_x(\xi)$ that is the solution to this maximum entropy variational problem. Find the density that maximizes the Shannon entropy with constraint on normalization and on the knowledge of first moment:

$$\underset{p_x(.)}{Max}\left[-\int_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi\right] \text{ such that } \begin{cases} \int_{\Omega^*} p_x(\xi) d\xi = 1 \\ \int_{\Omega^*} \xi . p_x(\xi) d\xi = x^* \end{cases} \tag{27}$$

If we consider the density $q_x(\xi) = e^{-\langle \xi, x \rangle} / \int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi = e^{-\langle x, \xi \rangle - \log \int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi}$ such that:

$$\begin{cases} \int\limits_{\Omega^*} q_x(\xi).d\xi = \int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi / \int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi = 1 \\ \log q_x(\xi) = \log e^{-\langle x, \xi \rangle - \log \int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi} = -\langle x, \xi \rangle - \log \int\limits_{\Omega^*} e^{-\langle x, \xi \rangle} d\xi \end{cases} \tag{28}$$

By using the inequality $\log x \geq \left(1 - x^{-1}\right)$ with equality if $x = 1$, we can then write that:

$$-\int\limits_{\Omega^*} p_x(\xi) \log \frac{p_x(\xi)}{q_x(\xi)} d\xi \leq -\int\limits_{\Omega^*} p_x(\xi) \left(1 - \frac{q_x(\xi)}{p_x(\xi)}\right) d\xi \tag{29}$$

We develop the right term of the equation:

$$\int\limits_{\Omega^*} p_x(\xi) \left(1 - \frac{q_x(\xi)}{p_x(\xi)}\right) d\xi = \int\limits_{\Omega^*} p_x(\xi) d\xi - \int\limits_{\Omega^*} q_x(\xi) d\xi = 0 \tag{30}$$

knowing that $\int\limits_{\Omega^*} p_x(\xi) d\xi = \int\limits_{\Omega^*} q_x(\xi) d\xi = 1$, we can deduce that:

$$-\int\limits_{\Omega^*} p_x(\xi) \log \frac{p_x(\xi)}{q_x(\xi)} d\xi \leq 0 \Rightarrow -\int\limits_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi \leq -\int\limits_{\Omega^*} p_x(\xi) \log q_x(\xi) d\xi \tag{31}$$

We have then to develop the right term by using previous expression of $q_x(\xi)$:

$$-\int\limits_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi \leq -\int\limits_{\Omega^*} p_x(\xi) \left[-\langle x, \xi \rangle - \log \int\limits_{\Omega^*} e^{-\langle x, \xi \rangle} d\xi \right] d\xi \tag{32}$$

$$-\int\limits_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi \leq \left\langle x, \int\limits_{\Omega^*} \xi . p_x(\xi) d\xi \right\rangle + \log \int\limits_{\Omega^*} e^{-\langle x, \xi \rangle} d\xi \tag{33}$$

If we use that $x^* = \int\limits_{\Omega^*} \xi . p_x(\xi) d\xi$ and $\Phi(x) = -\log \int\limits_{\Omega^*} e^{-\langle x, \xi \rangle} d\xi$, then we obtain that the density $q_x(\xi) = e^{-\langle \xi, x \rangle} / \int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi = e^{-\langle x, \xi \rangle - \log \int\limits_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi}$ is the maximum entropy density constrained by $\int\limits_{\Omega^*} p_x(\xi) d\xi$ and $\int\limits_{\Omega^*} \xi . p_x(\xi) d\xi = x^*$:

$$-\int\limits_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi \leq \left\langle x, x^* \right\rangle - \Phi(x) \tag{34}$$

$$-\int_{\Omega^*} p_x(\xi)\log p_x(\xi)d\xi \le \Phi^*(x^*) \tag{35}$$

In the following, we will write $x^* = \hat{\xi}$, to give to this variable the link with momentum $\hat{\xi} = \int_{\Omega^*} \xi . p_{\hat{\xi}}(\xi)d\xi$. To express the density with respect to the 1st moment as variable, we have to inverse $\hat{\xi} = \Theta(x) = \frac{d\Phi(x)}{dx}$, by writting $x = \Theta^{-1}(\hat{\xi})$ the inverse function (given by Legendre transform):

$$p_{\hat{\xi}}(\xi) = \frac{e^{-\left\langle \xi, \Theta^{-1}(\hat{\xi})\right\rangle}}{\int_{\Omega^*} e^{-\left\langle \xi, \Theta^{-1}(\hat{\xi})\right\rangle} d\xi} \text{ with } \hat{\xi} = \int_{\Omega^*} \xi . p_{\hat{\xi}}(\xi)d\xi \text{ and } \Phi(x) = -\log \int_{\Omega^*} e^{-\langle x, \xi\rangle} d\xi \tag{36}$$

We find finally the Maximum entropy density parametrized by 1st moment $\hat{\xi}$.

4 Links Between Koszul-Vinberg Characteristic Function, Koszul Forms and Information Geometry

Koszul Hessian Geometry Structure is the key tool to define elementary structures of Information Geometry, that appears as one particular case of more general framework studied by Koszul. In the Koszul-Vinberg Characteristic function $\psi_\Omega(x) = \int_{\Omega^*} e^{-\langle x, \xi\rangle} d\xi$, $\forall x \in \Omega$ where Ω is a sharp convex cone and Ω^* its dual cone, the duality bracket $<.,.>$ has to be defined. I will introduce it by using Cartan-Killing form $\langle x, y\rangle = -B(x, \theta(y))$ with $B(.,.)$ killing form and $\theta(.)$ Cartan involution. The inner product is then invariant with respect to automorphims of cone Ω. Koszul-Vinberg characteristic function could be developed as [124]:

$$\psi_\Omega(x + \lambda u) = \psi_\Omega(x) - \lambda\left\langle x^* + u\right\rangle + \frac{\lambda^2}{2}\langle K(x)u, u\rangle + \ldots \tag{37}$$

with $x^* = \frac{d\Phi(x)}{dx}$, $\Phi(x) = -\log \psi_\Omega(x)$ and $K(\mathrm{x}) = \frac{d^2\Phi(x)}{dx^2}$

In the following developments, I will write β, previous variable written x, because in thermodynamics, this variable corresponds to the Planck temperature, classically $\beta = \frac{1}{T}$. The variable β will be the dual variable of $\hat{\xi}$.

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

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

$$S(\hat{\xi})=\left\langle\hat{\xi},\beta\right\rangle-\Phi(\beta)\tag{38}$$

Inversion of the function $\Theta(.)$ is given by $\beta=\Theta^{-1}(\hat{\xi})$ is achieved by Legendre transform using relation between Entropy $S(\hat{\xi})$ and the function $\Phi(\beta)$ (opposite of the logarithm of the Koszul-Vinberg characteristic function):

$$\begin{aligned}&S(\hat{\xi})=\left\langle\beta,\hat{\xi}\right\rangle-\Phi(\beta)\\&\text{with}\quad\Phi(\beta)=-\log\int\limits_{\Omega^*} e^{-\langle\xi,\beta\rangle}d\xi\quad\forall\beta\in\Omega\quad\text{and}\quad\forall\xi,\hat{\xi}\in\Omega^*\end{aligned}\tag{39}$$

We will prove that the 2nd Koszul form $-\frac{\partial^2\Phi(\beta)}{\partial\beta^2}$ is linked with Fisher Metric of Information Geometry:

$$I(\beta)=-E\left[\frac{\partial^2\log p_\beta(\xi)}{\partial\beta^2}\right]\tag{40}$$

To compute the Fisher metric $I(\beta)$, we use the following relations between variable

$$\begin{aligned}&\begin{cases}\log p_{\hat{\xi}}(\xi)=-\langle\xi,\beta\rangle+\Phi(\beta)\\ S(\widehat{\xi})=-\int\limits_{\Omega^*} p_{\hat{\xi}}(\xi).\log p_{\hat{\xi}}(\xi).d\xi=-E\left[\log p_{\hat{\xi}}(\xi)\right]\end{cases}\\&\Rightarrow S(\widehat{\xi})=\langle E[\xi],\beta\rangle-\Phi(\beta)=\left\langle\hat{\xi},\beta\right\rangle-\Phi(\beta)\end{aligned}\tag{41}$$

We can observe that the logarithm of the density is affine with respect to the variable β, and that the Fisher matrix is given by the hessian. We can then deduce that the Fisher Metric is given by the hessian.

$$I(\beta)=-E\left[\frac{\partial^2\log p_\beta(\xi)}{\partial\beta^2}\right]=-E\left[\frac{\partial^2(-\langle\xi,\beta\rangle+\Phi(\beta))}{\partial\beta^2}\right]=-\frac{\partial^2\Phi(\beta)}{\partial\beta^2}=\frac{\partial^2\log\psi_\Omega(\beta)}{\partial\beta^2}\tag{42}$$

We can also identify the Fisher metric as a variance:

$$\log\Psi_\Omega(\beta)=\log\int\limits_{\Omega^*} e^{-\langle\xi,\beta\rangle}d\xi\Rightarrow\frac{\partial\log\Psi_\Omega(\beta)}{\partial\beta}=-\frac{1}{\int\limits_{\Omega^*} e^{-\langle\xi,\beta\rangle}d\xi}\int\limits_{\Omega^*}\xi.e^{-\langle\xi,\beta\rangle}d\xi\tag{43}$$

$$\frac{\partial^2 \log \Psi_\Omega(\beta)}{\partial \beta^2} = -\frac{1}{\left(\int\limits_{\Omega^*} e^{-\langle \xi,\beta \rangle} d\xi\right)^2} \left[-\int\limits_{\Omega^*} \xi^2.e^{-\langle \xi,\beta \rangle} d\xi . \int\limits_{\Omega^*} e^{-\langle \xi,\beta \rangle} d\xi + \left(\int\limits_{\Omega^*} \xi^2.e^{-\langle \xi,\beta \rangle} d\xi \right)^2 \right] \tag{44}$$

$$\frac{\partial^2 \log \Psi_\Omega(\beta)}{\partial \beta^2} = \int\limits_{\Omega^*} \xi^2 . \frac{e^{-\langle \xi,\beta \rangle}}{\int\limits_{\Omega^*} e^{-\langle \xi,\beta \rangle} d\xi} d\xi - \left(\int\limits_{\Omega^*} \xi . \frac{e^{-\langle \xi,\beta \rangle}}{\int\limits_{\Omega^*} e^{-\langle \xi,\beta \rangle} d\xi} d\xi \right)^2$$
$$= \int\limits_{\Omega^*} \xi^2 . p_\beta(\xi) d\xi - \left(\int\limits_{\Omega^*} \xi . p_\beta(\xi) d\xi \right)^2 \tag{45}$$

$$I(\beta) = -E_\xi \left[\frac{\partial^2 \log p_\beta(\xi)}{\partial \beta^2} \right] = \frac{\partial^2 \log \psi_\Omega(\beta)}{\partial \beta^2} = E_\xi\left[\xi^2\right] - E_\xi\left[\xi^2\right] = Var(\xi) \tag{46}$$

In 1977, Crouzeix [125, 126] has identified the following relation between both hessian of entropy and characteristic function $\frac{\partial^2 \Phi}{\partial \beta^2} = \left[\frac{\partial^2 S}{\partial \hat{\xi}^2}\right]^{-1}$ giving a relation between the dual metrics with respect to their dual coordinate systems. The metric could be given by Fisher metric or given by the hessian of Entropy *S:*

$$ds_g^2 = d\beta^T I(\beta) d\beta = \sum_{ij} g_{ij} d\beta_i d\beta_j \quad \text{with} \quad g_{ij} = [I(\beta)]_{ij} \tag{47}$$

Thanks to Crouzeix relation [125] [126], we observe that 2 geodesic distances given by hessian of dual potential functions in dual coordinates systems, are equal:

$$ds_h^2 = d\hat{\xi}^T \left[\frac{\partial^2 S(\hat{\xi})}{\partial \hat{\xi}^2} \right] d\hat{\xi} = \sum_{ij} h_{ij} d\hat{\xi}_i d\hat{\xi}_j \quad \text{with} \quad h_{ij} = \left[\frac{\partial^2 S(\hat{\xi})}{\partial \hat{\xi}^2} \right]_{ij} \tag{48}$$

$$ds_h^2 = ds_g^2 \tag{49}$$

One can ask oneself the question of what is the most natural product of duality. This question has been treated by Elie Cartan in his thesis in 1894, by introducing a form called Cartan-Killing form, a symmetric bilinear form naturally associated with any Lie algebra. This form of Cartan-Killing is defined via the endomorphism ad_x of Lie algebra g via the Lie bracket:

$$ad_x(y) = [x, y] \tag{50}$$

The trace of the composition of these 2 endomorphisms defines this bilinear form by:

$$B(x, y) = Tr\left(ad_x ad_y\right) \tag{51}$$

The Cartan-Killing form is symmetric:

$$B(x, y) = B(y, x) \tag{52}$$

and verify associativity property:

$$B([x, y], z) = B(x, [y, z]) \tag{53}$$

given by:

$$\begin{aligned} B([x, y], z) &= Tr\left(ad_{[x,y]}ad_z\right) = Tr\left(\left[ad_x, ad_y\right]ad_z\right) \\ &= Tr\left(ad_x\left[ad_y, ad_z\right]\right) = B(x, [y, z]) \end{aligned} \tag{54}$$

Elie Cartan proved that if g is a semi-simple Lie algebra (the form of Killing is non-degenerate) then any symmetric bilinear form is a scalar multiple of the Cartan-Killing form. The Cartan-Killing form is invariant under the action of automorphisms $\sigma \in Aut(g)$ of the algebra g:

$$B(\sigma(x), \sigma(y)) = B(x, y) \tag{55}$$

This invariance is deduced from:

$$\begin{cases} \sigma[x, y] = [\sigma(x), \sigma(y)] \\ z = \sigma(y) \end{cases} \Rightarrow \sigma\left[x, \sigma^{-1}(z)\right] = [\sigma(x), z]$$
$$\text{by writting} \quad ad_{\sigma(x)} = \sigma \circ ad_x \circ \sigma^{-1} \tag{56}$$

Then, we can write:

$$B(\sigma(x), \sigma(y)) = Tr\left(ad_{\sigma(x)}ad_{\sigma(y)}\right) = Tr\left(\sigma \circ ad_x ad_y \circ \sigma^{-1}\right) = Tr\left(ad_x ad_y\right) = B(a, y) \tag{57}$$

Cartan has introduced this natural inner product that is invariant by the automorphisms of the Lie algebra, from this Cartan-Killing form:

$$\langle x, y\rangle = -B(x, \theta(y)) \tag{58}$$

with $\theta \in g$ the Cartan involution (an involution on the Lie algebra g is an automorphism θ such that the square is equal to identity).

I summarize all these relations of information geometry from the characteristic function of Koszul-Vinberg, and the duality given via the Cartan-Killing form, as described in the figure below (Fig. 7):

Thanks to the expression of the characteristic function of Koszul-Vinberg and the Cartan-Killing form, one can express the maximum Entropy density in a very general way. For example, by applying these formulas to the cone Ω (self-dual: $\Omega^* = \Omega$) symmetric positive definite matrices $Sym^+(n)$, Cartan-Killing form gives us the product of duality:

$$\langle \eta, \xi\rangle = Tr(\eta^T \xi). \quad \forall \eta, \xi \in Sym^+(n) = \left\{\xi / \xi^T = \xi, \xi > 0\right\} \tag{59}$$

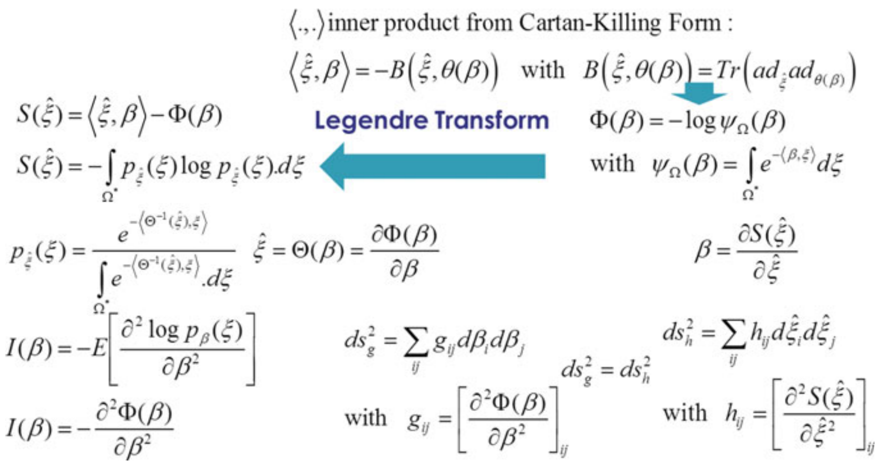


Fig. 7 Relations between cartan-killing form, koszul-vinberg characteristic function, potentials and dual coordinates, and metrics of information geometry

The maximum entropy density is given by:

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

$$\text{and} \quad \hat{\xi} = \frac{\partial \Phi(\beta)}{\partial \beta} = \frac{\partial(-\log \psi_\Omega(\beta))}{\partial \beta} = \frac{n+1}{2}\beta^{-1} \tag{60}$$

From which, I can deduce the final expression:

$$p_{\hat{\xi}}(\xi) = e^{-\left\langle \Theta^{-1}(\hat{\xi}), \xi \right\rangle + \Phi\left(\Theta^{-1}(\hat{\xi})\right)} = \psi_\Omega(I_d).\left[\det\left(\alpha \hat{\xi}^{-1}\right)\right].e^{-Tr\left(\alpha \hat{\xi}^{-1} \xi\right)}$$
$$\text{with} \quad \alpha = \frac{n+1}{2} \tag{61}$$

We can apply this approach for multivariate Gaussian densities. In the case of multivariate Gaussian densities, as noted by Souriau [17, 79], the classical Gibbs expression can be rewritten by modifying the coordinate system and defining a new duality product [80–84, 98]. The multivariate Gaussian density is classically written with the following coordinate system (m, R), with m the mean vector, and R the covariance matrix of the vector z:

$$p_{\hat{\xi}}(\xi) = \frac{1}{(2\pi)^{n/2} \det(R)^{1/2}} e^{-\frac{1}{2}(z-m)^T R^{-1}(z-m)} \text{ with } \begin{cases} m = E(z) \\ R = E\left[(z-m)(z-m)^T\right] \end{cases} \tag{62}$$

By developing the term in the exponential:

$$\frac{1}{2}(z-m)^T R^{-1}(z-m) = \frac{1}{2}\left[z^T R^{-1} z - m^T R^{-1} z - z^T R^{-1} m + m^T R^{-1} m\right]$$
$$= \frac{1}{2} z^T R^{-1} z - m^T R^{-1} z + \frac{1}{2} m^T R^{-1} m \quad (63)$$

I can write this density as a Gibbs density by introducing a new duality bracket between $\left(z, zz^T\right)$ and $\left(-R^{-1}m, \frac{1}{2}R^{-1}\right)$:

$$p_{\hat{\xi}}(\xi) = \frac{1}{(2\pi)^{n/2}\det(R)^{1/2} e^{\frac{1}{2}m^T R^{-1} m}} e^{-\left[-m^T R^{-1} z + \frac{1}{2} z^T R^{-1} z\right]} = \frac{1}{Z} e^{-\langle \xi, \beta \rangle}$$
$$\xi = \begin{bmatrix} z \\ zz^T \end{bmatrix} \text{ and } \beta = \begin{bmatrix} -R^{-1}m \\ \frac{1}{2}R^{-1} \end{bmatrix} = \begin{bmatrix} a \\ H \end{bmatrix}$$
$$\text{with } \langle \xi, \beta \rangle = a^T z + z^T H z = Tr\left[z a^T + H^T z z^T\right] \quad (64)$$

We can then write the density in Koszul form:

$$p_{\hat{\xi}}(\xi) = \frac{1}{\int\limits_{\Omega^*} e^{-\langle \xi, \beta \rangle}.d\xi} e^{-\langle \xi, \beta \rangle} = \frac{1}{Z} e^{-\langle \xi, \beta \rangle}$$
$$\text{with } \log(Z) = n\log(2\pi) + \frac{1}{2}\log\det(R) + \frac{1}{2} m^T R^{-1} m$$
$$\xi = \begin{bmatrix} z \\ zz^T \end{bmatrix}, \hat{\xi} = E[\xi] = \begin{bmatrix} E[z] \\ E\left[zz^T\right] \end{bmatrix} = \begin{bmatrix} m \\ R + mm^T \end{bmatrix}, \beta = \begin{bmatrix} a \\ H \end{bmatrix} = \begin{bmatrix} -R^{-1}m \\ \frac{1}{2}R^{-1} \end{bmatrix}$$
$$\text{with } \langle \xi, \beta \rangle = Tr\left[z a^T + H^T z z^T\right]$$
$$R = E\left[(z-m)(z-m)^T\right] = E\left[zz^T - mz^T - zm^T + mm^T\right] = E\left[zz^T\right] - mm^T \quad (65)$$

We are then able to compute the Koszul-Vinberg characteristic function whose opposite of the logarithm provides the potential function:

$$\psi_\Omega(\beta) = \int\limits_{\Omega^*} e^{-\langle \xi, \beta \rangle}.d\xi$$
$$\text{and } \Phi(\beta) = -\log\psi_\Omega(\beta) = \frac{1}{2}\left[-Tr\left[H^{-1} a a^T\right] + \log\left[(2)^n \det H\right] - n\log(2\pi)\right] \quad (66)$$

that verifies the following relation given by Koszul and linked with 1st Koszul form:

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

$$\frac{\partial \Phi(\beta)}{\partial \beta} = \begin{bmatrix} \frac{\partial \Phi(\beta)}{\partial \alpha} \\ \frac{\partial \Phi(\beta)}{\partial H} \end{bmatrix} = \begin{bmatrix} m \\ R + mm^T \end{bmatrix} = \hat{\xi} \tag{67}$$

The 2nd dual potential is given by the Legendre transform of $\Phi(\beta)$:

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

$$S\left(\hat{\xi}\right) = -\int\limits_{\Omega^*} \frac{e^{-\langle \xi, \beta \rangle}}{\int\limits_{\Omega^*} e^{-\langle \xi, \beta \rangle}.d\xi} \log \frac{e^{-\langle \xi, \beta \rangle}}{\int\limits_{\Omega^*} e^{-\langle \xi, \beta \rangle}.d\xi}.d\xi = -\int\limits_{\Omega^*} p_{\hat{\xi}}(\xi) \log p_{\hat{\xi}}(\xi).d\xi \tag{68}$$

that is explicitly identified with the classical Shannon Entropy:

$$\begin{aligned} S(\hat{\xi}) &= -\int\limits_{\Omega^*} p_{\hat{\xi}}(\xi) \log p_{\hat{\xi}}(\xi).d\xi \\ &= \frac{1}{2}\left[\log(2)^n \det\left[H^{-1}\right] + n \log(2\pi.e)\right] = \frac{1}{2}\left[\log \det[R] + n \log(2\pi.e)\right] \end{aligned} \tag{69}$$

The Fisher metric of Information Geometry is given by the hessian of the opposite of the logarithm of the Koszul-Vinberg characteristic function:

$$\begin{aligned} & ds_g^2 = d\beta^T I(\beta) d\beta = \sum_{ij} g_{ij} d\beta_i d\beta_j \\ & \text{with} \quad g_{ij} = [I(\beta)]_{ij} \text{ and } I(\beta) = -E_\xi\left[\frac{\partial^2 \log p_\beta(\xi)}{\partial \beta^2}\right] = \frac{\partial^2 \log \psi_\Omega(\beta)}{\partial \beta^2} \end{aligned} \tag{70}$$

Then, for the multivariate Gaussian density, we have the following Fisher metric:

$$ds^2 = \sum_{ij} g_{ij} d\theta_i d\theta_j = dm^T R^{-1} dm + \frac{1}{2} Tr\left[\left(R^{-1} dR\right)^2\right] \tag{71}$$

Geodesic equations are given by Euler-Lagrange equations:

$$\begin{aligned} & \sum_{i=1}^n g_{ik} \ddot{\theta}_i + \sum_{i,j=1}^n \Gamma_{ijk} \dot{\theta}_i \dot{\theta}_j = 0, \ k = 1, \ldots, n \\ & \text{with} \quad \Gamma_{ijk} = \frac{1}{2}\left[\frac{\partial g_{jk}}{\partial \theta_i} + \frac{\partial g_{jk}}{\partial \theta_j} + \frac{\partial g_{ij}}{\partial \theta_k}\right] \end{aligned} \tag{72}$$

that can be reduced to the equations:

$$\begin{cases} \ddot{R} + \dot{m}\dot{m}^T - \dot{R}R^{-1}\dot{R} = 0 \\ \ddot{m} - \dot{R}R^{-1}\dot{m} = 0 \end{cases} \tag{73}$$

I use a result of Souriau [17] that the component of « moment map » are constants (geometrization of Emmy Noether theorem), to identify the following constants [83]:

$$\frac{d\Pi_R}{dt} = \begin{bmatrix} \frac{d\left(R^{-1}\dot{R}+R^{-1}\dot{m}m^T\right)}{dt} & \frac{d\left(R^{-1}\dot{m}\right)}{dt} \\ 0 & 0 \end{bmatrix} = 0$$
$$\Rightarrow \begin{cases} R^{-1}\dot{R} + R^{-1}\dot{m}m^T = B = cste \\ R^{-1}\dot{m} = b = cste \end{cases} \tag{74}$$

with Π_R the moment map introduced by Souriau [17]. This moment map could be computed if we consider the following Lie group acting in case of Gaussian densities:

$$\begin{bmatrix} Y \\ 1 \end{bmatrix} = \begin{bmatrix} R^{1/2} & m \\ 0 & 1 \end{bmatrix}\begin{bmatrix} X \\ 1 \end{bmatrix} = \begin{bmatrix} R^{1/2}X + m \\ 1 \end{bmatrix}, \begin{cases} (m, R) \in R^n \times Sym^+(n) \\ M = \begin{bmatrix} R^{1/2} & m \\ 0 & 1 \end{bmatrix} \in G_{aff} \end{cases}$$
$$X \approx \aleph(0, I) \rightarrow Y \approx \aleph(m, R) \tag{75}$$

$R^{1/2}$, square root of R, is given by Cholesky decomposition of R. $R^{1/2}$ is the Lie group of triangular matrix with positive elements on the diagonal. Euler-Poincaré equations, reduced equations from Euler-Lagrange equations, are then given by:

$$\begin{cases} \dot{m} = Rb \\ \dot{R} = R(B - bm^T) \end{cases} \tag{76}$$

Geodesic distance between multivariate Gaussian density is then obtained by "*geodesic shooting*" method that will provide iteratively the final solution from the tangent vector at the initial point:

$$\left(R^{-1}(0)\dot{m}(0), R^{-1}(0)\left(\dot{R}(0) + \dot{m}(0)m(0)^T\right)\right) = (b, B) \in R^n \times Sym^+(n) \tag{77}$$

From which, we then deduce the distance:

$$d = \sqrt{\dot{m}(0)^T R^{-1}(0)\dot{m}(0) + \frac{1}{2}Tr\left[\left(R^{-1}(0)\dot{R}(0)\right)^2\right]} \tag{78}$$

Geodesic shooting is obtained by using equations established by Eriksen [127, 128] for "*exponential map*" using the following change of variables:

$$\begin{cases} \Delta(t) = R^{-1}(t) \\ \delta(t) = R^{-1}(t)m(t) \end{cases} \Rightarrow \begin{cases} \dot{\Delta} = -B\Delta + bm^T \\ \dot{\delta} = -B\delta + \left(1 + \delta^T \Delta^{-1}\delta\right)b \\ \Delta(0) = I_p, \delta(0) = 0 \end{cases} \text{with} \begin{cases} \dot{\Delta}(0) = -B \\ \dot{\delta}(0) = b \end{cases} \tag{79}$$

The method based on geodesic shooting consists in iteratively approaching the solution by geodesic shooting in direction $\left(\dot{\delta}(0), \dot{\Delta}(0)\right)$, using the following exponential map (Fig. 8):

$$\Lambda(t) = \exp(tA) = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!} = \begin{pmatrix} \Delta & \delta & \Phi \\ \delta^T & \varepsilon & \gamma^T \\ \Phi^T & \gamma & \Gamma \end{pmatrix}$$
$$\text{with} \quad A = \begin{pmatrix} -B & b & 0 \\ b^T & 0 & -b^T \\ 0 & -b & B \end{pmatrix} \tag{80}$$

The principle of geodesic shooting is the following. We consider one geodesic χ between θ_0 and θ_1 with an initial tangent vector V from the origin, and assume that V is modified by W, with respect to $V + W$. Variation of final point θ_1 could be obtained by Jacobi vector field $J(0) = 0$ and $\dot{J}(0) = W$:

$$J(t) = \frac{d}{d\alpha} \exp_{\theta_0}(t(V + \alpha W))|_{\alpha=0} \tag{81}$$

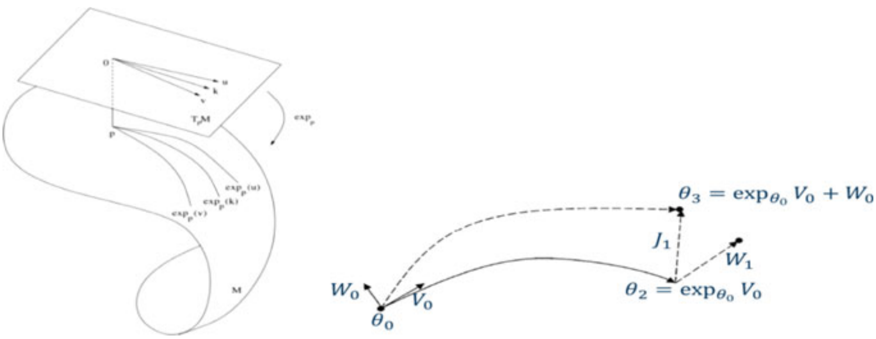


Fig. 8 Principle of geodesic shooting in the direction of the initial vector V_0 at the origin and correction by W_0

5 Koszul's Study of Homogeneous Bounded Domains and Affine Representations of Lie Groups and Lie Algebras

Jean-Louis Koszul [15, 16, 41, 45–47, 49, 50] and his student Jacques Vey [39, 40] introduced new theorems with more general extension than previous results:

Koszul theorem [50]: Let Ω be a sharp convex open in an affine space of E of finite dimension on R. If a unimodular Lie group of affine transformations operates transitively on Ω, Ω is a cone.

Koszul-Vey Theorem [40]: Let M a hessian connected manifold associated with the hessian metric g. Assume that M has a closed 1-form α such that $D\alpha = g$ and that there is a group G of affine automorphisms of M preserving α, then:

- If M/G is almost compact, then the manifold, universal covering of M, is affinely isomorphic to a convex domain of an affine space containing no straight line.
- If M/G is compact, then Ω is a sharp convex cone.

Jean-Louis Koszul developed his theory, studying the homogeneous domains, in particular the homogeneous symmetric bounded domains of Siegel, which we note *DS* [44, 48]. He has proved that there is a subgroup G in the group of complex affine automorphisms of these domains (Iwasawa subgroup), so that G acts on *DS* in a merely transitive way. The Lie algebra g of G has a structure which is an algebraic translation of the Kähler structure *DS*.

Koszul considered on *G/B* an invariant complex structure tensor I. All the invariant volumes on *G/B*, equal up to a constant factor, define with the complex structure the same invariant Hermitian form on *G/B*, called Hermitian canonical form, denoted h. Let E be a differentiable fiber space of base M and let p be the projection of E on M, such that $p^*((pX).f) = X.(p^* f)$. The projection $p : E \to M$ defines an injective homomorphism p^* of the space of differential forms of M in the space of the differential forms of E such that for any form α of degree n on M and any sequence of n projectable vectors fields, we have $p^*(\alpha(pX_1, pX_2, \ldots, pX_n)) = (p^*\alpha)(X_1, X_2, \ldots, X_n)$. Let I be the tensor of an almost complex structure on the basis M, there exists on E a tensor J of type *(1,1)* and only one which possesses the following properties $p(JX) = I(pX)$ and $J^2 X = -X \mod h, X \in g$ for any vector field X on E. Let G be a connected Lie group and B a closed subgroup of G, we note g the Lie algebra left invariant vector fields on G and b sub-algebra of g corresponding to B. The canonical mapping of G on *G/B* is denoted p (defining E as before). We assume that there exists on *G/B* an invariant volume by G, which consist in assuming that, for all $s \in B$, the automorphism $X \to Xs$ of g defines by passing to the quotient an automorphism of determinant *1* in *g/b*. Let r be the dimension of *G/B* and $(X_i)_{1 \le i \le m}$ a base of g such that $X_i \in b$, for $r \le i \le m$. Let $(\xi_i)_{1 \le i \le m}$ the base of the space of differential forms of degree *1* left invariant on G such that $\xi_i(X_j) = \delta_{ij}$. If ω is an invariant volume on *G/B*, then $\Omega = p^*\omega$ is equal, up to a constant factor, to $\xi_1 \wedge \xi_2 \wedge \ldots \wedge \xi_r$. We will assume the base (X_j) chosen so that this factor is equal to *1*, let $\Omega = \xi_1 \wedge \xi_2 \wedge \ldots \wedge \xi_r$. For any vector field that can be projected X on G, we have:

$$p^*(div(pX))\Omega = p^*((div(pX))\omega) = p^*((pX)\omega) = X\Omega = \sum_{j=1}^{r} \xi_j([X_j, X])\Omega \tag{82}$$

$$p^*(div(pX)) = \sum_{j=1}^{r} \xi_j([X_j, X]) \tag{83}$$

These elements being defined, Koszul calculates the Hermitian canonical form of *G/B*, denoted h, more particularly $\eta = p^*h$ on G. Let X and Y both right invariant vector fields on G. They are projectable and the fields pX and pY are conformal vector fields on *G/B* such that $div(pX) = div(pY) = 0$, because the volume and the complex structure of *G/B* are invariant under G. As a result, if κ is the Kähler form of h and if $\alpha = p^*\kappa$, then:

$$4\alpha(X, Y) = 4p^*(\kappa(pX, pY)) = p^*div(I[pX, pY]) \tag{84}$$

and as $p(J[X, Y]) = I[pX, pY]$, we obtain:

$$4\alpha(X, Y) = p^*div(J[X, Y]) = \sum_{i=1}^{2n} \xi_i([X_i, J[X, Y]]) \tag{85}$$

X and Y are two left invariant vectors fields on G. *X '* and *Y'* right invariant vectors fields coinciding with X and Y at the point e, neutral element of G. If $T = \left[X', Y'\right]$ is tight invariant vectors fields which coincide with $-[X, Y]$ on e, then:

$$[X, JT] = J[X, [X, Y]] - [X, J[X, Y]] \text{ at point e} \tag{86}$$

At point e, we have the equality:

$$4\alpha(X, Y) = \sum_{i=1}^{2n} \xi_i([J[X, Y], X_i] - J[[X, Y], X_i]) \tag{87}$$

As the form α is invariant on the left by G, this equality is verified for all points. For any endomorphism Θ of the space g such that $\Theta b \subset b$, we denote by $Tr_b\Theta$ the trace of the restriction of Θ to b and by $Tr_{g/b}\Theta$ the trace of the endomorphism of *g/b* deduced from Θ by passage to the quotient, with $Tr\Theta = Tr_b\Theta + Tr_{g/b}\Theta$. We have:

$$T_{r_{g/b}\Theta} = \sum_{i=1}^{2n} \xi_i(\Theta X_i) \tag{88}$$

Whatever $X \in g$ and $s \in B$, we have $J(Xs) - (JX)s \in b$. If $ad(Y)$ is the endomorphism of g defined by $ad(Y).Z = [Y, Z]$, we have $(Jad(Y) - ad(Y)J)g \subset b$ for all $Y \in b$. We can deduce, for all $X \in g$, the endomorphism $ad(JX) - Jad(X)$ leaves steady the subspace b. Koszul defines a linear form Ψ on the space g by defining:

$$\Psi(X) = Tr_{g/b}(ad(JX) - Jad(X)) \ , \forall X \in g \tag{89}$$

Koszul has finally obtained the following fundamental theorem:

Theorem of Koszul [15]:

The Kähler form of the Hermitian canonical form has for image by p^* the differential of the form $-\frac{1}{4}\Psi(X) = -\frac{1}{4}Tr_{g/b}(ad(JX) - Jad(X)), \forall X \in g$

Koszul note that the form Ψ is independent of the choice of the tensor J. It is determined by the invariant complex structure of *G/B*. The form Ψ is right invariant by B. For all $s \in B$, note the endomorphism $r(s) : X \to Xs$ of g. Since $J(Xs) = (JX)s \mod b$ and that $Tr_{g/b}ad(Y) = 0$, we have:

$$\Psi(Xs) = Tr_{g/b}(ad((JX)s) - Jad(Xs)), \quad \forall X \in g \, , \forall Y \in b \tag{90}$$

$$\Psi(Xs) = Tr_{g/b}(r(s)ad(JX)r(s)^{-1} - Jr(s)ad(X)r(s)^{-1}) \tag{91}$$

$$\Psi(Xs) = \Psi(X) + Tr_{g/b}((J - r(s)^{-1}Jr(s))ad(X)), \quad \forall X \in g, s \in B \tag{92}$$

As $\left(J - r(s)^{-1}Jr(s)\right)$ maps g in b, we get $\Psi(Xs) = \Psi(X)$. The form Ψ is not zero on b. This is not the image by p^* of a differential form of *G/B*. However, the right invariance of Ψ on B is translated, infinitesimally by the relation:

$$\Psi([b, g]) = (0) \tag{93}$$

Koszul proved that the canonical hermitian form h of a homogeneous Kähler manifold *G/B* has the following expression:

$$\eta(X, Y) = \frac{1}{2}\Psi([JX, Y])$$
$$\text{with} \ \begin{cases} \Psi([X, Y]) = \Psi([JX, JY]) \\ \eta([JX, JY]) = \eta(X, Y) \end{cases} \forall X, Y \in g \tag{94}$$

To do, the link with the first chapters, I can summarize the main result of Koszul that there is an integrable structure almost complex J on g, and for $l \in g^*$ defined by a positive J -invariant inner product on g:

$$\langle X, Y \rangle_l = \langle [JX, Y], l \rangle \tag{95}$$

Koszul has proposed as admissible form, $l \in g^*$, the form ξ:

$$\Psi(X) = \langle X, \xi \rangle = T_r[ad(JX) - J.ad(X)]\ \forall X \in g \tag{96}$$

Koszul proved that $\langle X, Y \rangle_\xi$ coincides, up to a positive multiplicative constant; with the real part of the Hermitian inner product obtained by the Bergman metric of symmetric homogeneous bounded domains DS by identifying g with the tangent space of DS. $\Psi(X)$ is the restriction to g of a differential form Ψ of degree 1, with left invariance on G. This form is fully defined by the invariant complex structure of G/B. This form is invariant to the choice of J. This form is invariant on the right by B. We have $\Psi([X, Y]) = 0$ with $X \in g$, $Y \in b$. The exterior differential $d\Psi$ of Ψ is the inverse image by the projection $G \to G/B$ of degree 2 form Ω. This form Ω is, up to a constant, the Kähler form h, defined by the canonical Hermitian form of G/B: $h(\pi.X, \pi.Y) = \frac{1}{2}(d\Psi)(X, J.Y), \forall X, Y \in G$ as it is proved in Bourbaki seminar by Koszul in [129].

The 1st Koszul form is then given by:

$$\alpha = -\frac{1}{4} d\Psi(X) \tag{97}$$

We can illustrate this structure for the simplest example of DS, the Poincaré upper half-plane $V = \{z = x + iy / y > 0\}$ which is isomorphic to the open $zz^* < 1$, which is a bounded domain. The group G of transformations $z \to az + b$ with a and b real values with $a > 0$ is simply transitive in V. We identify G and V by the application passing from $s \in G$ an element to the image $i = \sqrt{-1}$ by s.

Let's define vector fields $X = y\frac{d}{dx}$ and $Y = y\frac{d}{dy}$ which generate the vector space of left invariant vectors fields on G, and J an almost complex structure on V defined by $JX = Y$. As $[X, Y] = -Y$ and $ad(Y).Z = [Y, Z]$ then:

$$\begin{cases} Tr[ad(JX) - Jad(X)] = 2 \\ Tr[ad(JY) - Jad(Y)] = 0 \end{cases} \tag{98}$$

The Koszul forms and the Koszul metric are respectively given by:

$$\Psi(X) = 2\frac{dx}{y} \Rightarrow \alpha = -\frac{1}{4}d\Psi = -\frac{1}{2}\frac{dx \wedge dy}{y^2} \Rightarrow ds^2 = \frac{dx^2 + dy^2}{2y^2} \tag{99}$$

I note that $\alpha = -\frac{1}{4}d\Psi(X)$ is indeed the Kähler form of Poincaré's metric, which is invariant by the automorphisms of the upper half-plane.

The following example concerns $V = \{Z = X + iY / X, Y \in Sym(p), Y > 0\}$ the upper half-space of Siegel (which is the most natural extension of the Poincaré half-plane) with:

$$\begin{cases} SZ = (AZ + B)D^{-1} \\ A^T D = I, B^T D = D^T B \end{cases} \text{with } S = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \text{ and } J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{100}$$

We can then compute Koszul forms and the metric:

$$\Psi(dX + idY) = \frac{3p+1}{2} Tr(Y^{-1}dX)$$
$$\Rightarrow \begin{cases} \alpha = -\frac{1}{4}d\Psi = \frac{3p+1}{8} Tr\left(Y^{-1}dZ \wedge Y^{-1}d\bar{Z}\right) \\ ds^2 = \frac{(3p+1)}{8} Tr\left(Y^{-1}dZY^{-1}d\bar{Z}\right) \end{cases} \quad (101)$$

We recover Carl-Ludwig Siegel metric for the upper half space.

More recent development on Kähler manifolds are described in [130] et [131].

Koszul studied symmetric homogeneous spaces and defines the relation between invariant flat affine connections and the affine representations of Lie algebras and invariant Hessian metrics characterized by affine representations of Lie algebras. Koszul provides a correspondence between symmetric homogeneous spaces with invariant Hessian structures using affine representations of Lie algebras, and proves that a symmetric homogeneous space simply connected with an invariant Hessian structure is a direct product of a Euclidean space and of a homogeneous dual-cone. Let *G* be a connected Lie group and *G/K* a homogeneous space over which *G* acts effectively. Koszul gives a bijective correspondence between all planar *G*-invariantes connections on *G/K* and all of a certain class of affine representations of the Lie algebra of *G*. The main theorem of Koszul is:

Koszul's theorem: Let *G/K* be a homogeneous space of a connected Lie group *G* and be $\mathfrak{g}$ and $\mathfrak{k}$ the Lie algebras of *G* and *K*, assuming that *G/K* has G-invariant connection, then admits an affine representation *(f, q)* on the vector space *E*. Conversely, assume that *G* is simply connected and has an affine representation, then *G/K* admits a flat *G*-invariant connection.

In the foregoing, the basic tool studied by Koszul is the affine representation of Lie algebra and Lie group. To study these structures, Koszul introduced the following developments.

Let Ω a convex domain on R^n without any 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 exist an affine embedding:

$$\ell : x \in \Omega \mapsto \begin{bmatrix} x \\ 1 \end{bmatrix} \in V(\Omega) \quad (102)$$

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

$$s \in A(n, R) \mapsto \begin{bmatrix} \mathsf{f}(s) & \mathsf{q}(s) \\ 0 & 1 \end{bmatrix} \in GL(n + 1, R) \quad (103)$$

and the affine representation of Lie algebra:

$$\begin{bmatrix} f & q \\ 0 & 0 \end{bmatrix} \quad (104)$$

with $A(n, R)$ the group of all affine representations of R^n. We have $\eta(G(\Omega)) \subset G(V(\Omega))$ and the pair (η, ℓ) of homomorphism $\eta : (G(\Omega) \to G(V(\Omega))$ and the application $\ell : \Omega \to V(\Omega)$ is equivariant.

If we observe Koszul affine representations of Lie algebra and Lie group, we have to consider G a convex Lie group and E a real or complex vector space of finite size, Koszul has introduced an affine representation of G in E such that:

$$\begin{aligned} &E \to E \\ &a \mapsto sa \, \forall s \in G \end{aligned} \tag{105}$$

is an affine representation. We set $A(E)$ the set of all affine transformation of a real vector space E, a Lie group called affine representation group of E. The set $GL(E)$ of all regular linear representation of E, a sub-group of $A(E)$.

We define a linear representation of G in $GL(E)$:

$$\begin{aligned} &\mathsf{f} : G \to GL(E) \\ &s \mapsto \mathsf{f}(s)a = sa - so \quad \forall a \in E \end{aligned} \tag{106}$$

and a map from G to E:

$$\begin{aligned} &\mathsf{q} : G \to E \\ &s \mapsto \mathsf{q}(s)so \quad \forall s \in G \end{aligned} \tag{107}$$

then, we have $\forall s, t \in G$:

$$\mathsf{f}(s)\mathsf{q}(t) + \mathsf{q}(s) = \mathsf{q}(st) \tag{108}$$

deduced from $\mathsf{f}(s)\mathsf{q}(t) + \mathsf{q}(s) = s\mathsf{q}(t) - s\mathsf{q} + so = s\mathsf{q}(t) = sto = \mathsf{q}(st)$.

Inversely, if a map q from G to E and a linear representation f from G to $GL(E)$ verifying previous equation, then we can define an affine representation from G in E, written by (f, q):

$$\mathrm{A}ff(s) : a \mapsto sa = \mathsf{f}(s)a + \mathsf{q}(s) \, \forall s \in G, \forall a \in E \tag{109}$$

The condition $\mathsf{f}(s)\mathsf{q}(t) + \mathsf{q}(s) = \mathsf{q}(st)$ is equal to the request that the following mapping is an homomorphism:

$$\mathrm{A}ff : s \in G \mapsto \mathrm{A}ff(s) \in A(E) \tag{110}$$

We write f the affine representation of Lie algebra g of G, defined by f and q the restriction to g to the differential of q (f and q differential of f and q respectively), Koszul proved the following equation:

$$\begin{aligned} &f(X)q(Y) - f(Y)q(X) = q([X, Y]) \forall X, Y \in \mathsf{g} \\ &\text{with} \quad f : \mathsf{g} \to gl(E) \quad \text{and}\, q : \mathsf{g} \mapsto E \end{aligned} \tag{111}$$

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

We use the assumption that:

$$q(Ad_s Y) = \left.\frac{d\mathsf{q}\left(s.e^{tY}.s^{-1}\right)}{dt}\right|_{t=0} = \mathsf{f}(s)f(Y)\mathsf{q}(s^{-1}) + \mathsf{f}(s)q(Y) \tag{112}$$

We then obtain:

$$q([X,Y]) = \left.\frac{d\mathsf{q}\left(Ad_{e^{tX}}Y\right)}{dt}\right|_{t=0} = f(X)q(Y)\mathsf{q}(e) + \mathsf{f}(e)f(Y)(-q(X)) + f(X)q(Y) \tag{113}$$

where e is neutral element of G. Since $\mathsf{f}(e)$ is identity map and $\mathsf{q}(e) = 0$, we have the equality:

$$f(X)q(Y) - f(Y)q(X) = q([X,Y]) \tag{114}$$

A pair (f, q) of linear representation of f of a Lie algebra g on E and a linear map q from g in E is an affine representation of g in E, if it satisfy:

$$f(X)q(Y) - f(Y)q(X) = q([X,Y]) \tag{115}$$

Inversely, if we assume that g has an affine representation (f, q) on E, by using the coordinate systems $\{x^1, \ldots, x^n\}$ on E, we can express the affine map $v \mapsto f(X)v + q(Y)$ by a matrix representation of size $(n+1) \times (n+1)$:

$$aff(X) = \begin{bmatrix} f(X) & q(X) \\ 0 & 0 \end{bmatrix} \tag{116}$$

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

$X \mapsto aff(X)$ is an injective homomorphism of Lie algebra g in Lie algebra of matrices $(n+1) \times (n+1)$, $gl(n+1, R)$:

$$\left| \begin{array}{l} \mathsf{g} \to gl(n+1, R) \\ X \mapsto aff(X) \end{array} \right. \tag{117}$$

If we note $\mathsf{g}_{aff} = aff(\mathsf{g})$, we write G_{aff} linear Lie sub-group of $GL(n+1, R)$ generated by g_{aff} . One element of $s \in G_{aff}$ could be expressed by:

$$Aff(s) = \begin{bmatrix} \mathsf{f}(s) & \mathsf{q}(s) \\ 0 & 1 \end{bmatrix} \tag{118}$$

Let M_{aff} the orbit of G_{aff} from the origin o, then $M_{aff} = \mathsf{q}(G_{aff}) = G_{aff}/K_{aff}$ where $K_{aff} = \{s \in G_{aff}/\mathsf{q}(s) = 0\} = Ker(\mathsf{q})$.

We can give as example the following case. Let Ω a convex domain in R^n without any straight line, we define the cone $V(\Omega)$ in $R^{n+1} = R^n \times R$ by $V(\Omega) = \{(\lambda x, x) \in R^n \times R / x \in \Omega, \lambda \in R^+\}$. Then, there is an affine embedding:

$$\ell : x \in \Omega \mapsto \begin{bmatrix} x \\ 1 \end{bmatrix} \in V(\Omega) \tag{119}$$

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

$$s \in A(n, R) \mapsto \begin{bmatrix} \mathrm{f}(s) & \mathrm{q}(s) \\ 0 & 1 \end{bmatrix} \in GL(n+1, R) \tag{120}$$

with $A(n, R)$ the group of all affine transformations in R^n. We have $\eta(G(\Omega)) \subset G(V(\Omega))$ and the pair (η, ℓ) of homomorphism $\eta : G(\Omega) \to G(V(\Omega))$ and the map $\ell : \Omega \to V(\Omega)$ are equivariant:

$$\ell \circ s = \eta(s) \circ \ell \textbf{ and } d\ell \circ s = \eta(s) \circ d\ell \tag{121}$$

6 Koszul Lecture on Geometric and Analytics Mechanics, Related to Geometric Theory of Heat (Souriau's Lie Group Thermodynamics) and Theory of Information (Information Geometry)

Before that Professor Koszul passed away in January 2018, he gave his agreement to his book "Introduction to Symplectic Geometry" translation from Chinese to English by SPRINGER [18]. This Koszul's book translation genesis dates back to 2013. We had contacted Professor Jean-Louis Koszul, to deeper understand his work in the field of homogeneous bounded domains within the framework of Information Geometry. Professor Michel Boyom succeeded to convince Jean-Louis Koszul to answer positively to our invitation to attend the 1st GSI "*Geometric Science of Information*" conference in August 2013 at Ecole des Mines ParisTech in Paris, and more especialy to attend the talk of Hirohiko Shima, given for his honor on the topic "*Geometry of Hessian Structures*" (Fig. 9).

I was more particularly interested by Koszul's work developed in the paper «***Domaines bornées homogènes et orbites de groupes de transformations affines*** » [45] of 1961, written by Koszul at the Institute for Advanced Studies at Princeton during a stay funded by the National Science Foundation. Koszul proved in this paper that on a complex homogeneous space, an invariant volume defines with the complex structure the canonical invariant Hermitian form introduced in [15]. It is in this article that Koszul uses the ***affine representation of Lie groups and Lie algebras***.

Fig. 9 Jean-Louis Koszul and Hirihiko Shima at GSI'13 "*Geometric Science of Information*" conference in Ecole des Mines ParisTech in Paris, October 2013

The use by Koszul of the affine representation of Lie groups and Lie algebras drew our attention, especially on the links of his approach with the similar one used by Jean-Marie Souriau in geometric mechanics in the framework of homogeneous symplectic manifolds. I have then looked for links between Koszul and Souriau works. I finally discovered, that in 1986, Koszul published this book "Introduction to symplectic geometry" following a Chinese course in China. I also observed that this book takes up and develops works of Jean-Marie Souriau on homogeneous symplectic manifolds and ***the affine representation of Lie algebras and Lie groups in geometric mechanics***.

I have then exchanged e-mails with Professor Koszul on Souriau works and on genesis of this Book. In May 2015, questioning Koszul on Souriau work on Geometric Mechanics and on Lie Group Thermodynamics, Koszul answered me "[*A l'époque où Souriau développait sa théorie, l'establishment avait tendance à ne pas y voir des avancées importantes. Je l'ai entendu exposer ses idées sur la thermodynamique mais je n'ai pas du tout réalisé à l'époque que la géométrie hessienne était en jeu.*] *At the time when Souriau was developing his theory, the establishment tended not to see significant progress. I heard him explaining his ideas on thermodynamics but I did not realize at the time that Hessian geometry was at stake*". In September 2016, I asked him the origins of Lie Group and Lie Algebra Affine representation. Koszul informed me that he attended Elie Cartan Lecture, where he presented seminal work on this topic: "*[Il y a là bien des choses que je voudrais comprendre (trop peut-être !), ne serait-ce que la relation entre ce que j'ai fait et les travaux de Souriau. Détecter l'origine d'une notion ou la première apparition d'un résultat est souvent difficile. Je ne suis certainement pas le premier à avoir utilisé des représentations affines de groupes ou d'algèbres de Lie. On peut effectivement imaginer que cela se trouve chez Elie Cartan, mais je ne puis rien dire de précis. A propos d'Elie Cartan: je n'ai pas été son élève. C'est Henri Cartan qui a été mon maître pendant mes années de thèse. En 1941 ou 42 j'ai entendu une brève série de conférences données par Elie à l'Ecole Normale et ce sont des travaux d'Elie qui ont été le point de départ de mon travail de thèse.] There are many things that I would like to understand (too much perhaps!), If only the relationship between what I did and the work of Souriau. Detecting the origin of a notion or the first appearance of a result is often difficult. I am certainly not the first to have used affine representations of Lie groups or Lie algebras. We can imagine that it is at Elie Cartan, but I cannot say anything specific. About Elie Cartan: I was not his student. It was Henri Cartan who was my master during my years of thesis. In 1941 or 42, I heard a brief series of lectures given by Elie at the Ecole Normale and it was Elie's work that was the starting point of my thesis work*".

After discovering the existence of this Koszul's book, written in Chinese based on a course given at Nankin, on "Introduction to Symplectic Geometry", where he made reference to Souriau's book and developed his main tools, I started to discuss its content. In January 2017, Koszul wrote me, with the usual humility "*[Ce petit fascicule d'introduction à la géométrie symplectique a été rédigé par un assistant de Nankin qui avait suivi mon cours. Il n'y a pas eu de version initiale en français.] This small introductory booklet on symplectic geometry was written by a Nanjing assistant who had taken my course. There was no initial version in French* ". I asked him if he had personal archive of this course, he answered "*[Je n'ai pas conservé de notes préparatoires à ce cours. Dites-moi à quelle adresse je puis vous envoyer un exemplaire du texte chinois.] I have not kept any preparatory notes for this course. Tell me where I can send you a copy of the Chinese text. "*. Professor Koszul then sent me his last copy of this book in Chinese, a small green book (Fig. 10).

I was not able to read the Chinese text, but I have observed in Chap. 4 "*Symplectic G-spaces*" and in Chap. 5 "*Poisson Manifolds*", that their equations content new original developments of Souriau work on moment map and affine representation of Lie Group and Lie Algebra. More especially, Koszul considered equivariance of moment map, where I recover Souriau theorem. Koszul shows that when $(M;\omega)$ is a connected Hamiltonian G-space and μ a moment map of the action of G, there exists an affine action of G on g^* (dual Lie algebra), whose linear part is the coadjoint action, for which the moment μ is equivariant. Koszul developed Souriau idea that this affine action is obtained by modifying the coadjoint action by means of a closed cochain (called cocycle by Souriau), and that $(M;\omega)$ is a G-Poisson space making reference to Souriau's book for more details.

About collaboration between Koszul and Souriau and another potential Lecture on Symplectic Geometry in Toulouse, Koszul informed me in February 2017 that: "*[J'ai plus d'une fois rencontré Souriau lors de colloques, mais nous n'avons jamais*

Fig. 10 Original small green Koszul's book "Introduction to symplectic geometry" in Chinese

collaboré. Pour ce qui est de cette allusion à un "cours" donné à Toulouse, il y erreur. J'y ai peut être fait un exposé en 81, mais rien d'autre.] I have met Souriau more than once at conferences, but we have never collaborated. As for this allusion to a "course" given in Toulouse, there is error. I could have made a presentation in 81, but nothing else. ". Koszul admitted that he had no direct collaboration with Souriau: "*[Je ne crois pas avoir jamais parlé de ses travaux avec Souriau. Du reste j'avoue ne pas en avoir bien mesuré l'importance à l'époque] I do not think I ever talked about his work with Souriau. For the rest, I admit that I did not have a good idea of the importance at the time*".

Considering the importance of this book for different communities, I tried to find an editor for its translation in English. By chance, I met Catriona Byrne from SPRINGER, when I gave a talk at IHES, invited by Pierre Cartier, on Koszul and Souriau works application in Radar. With help of Michel Boyom, we have convinced Professor Koszul to translate this book, proposing to contextualize this book with regard to the contemporary research trends in Geometric Mechanics, Lie Groups Thermodynamics and Geometric Science of Information. Professors Marle and Boyom accepted to check the translation and help me to write the forewords.

In the historical Foreword of this book, Koszul write "*The development of analytical mechanics provided the basic concepts of symplectic structures. The term symplectic structure is due largely to analytical mechanics. But in this book, the applications of symplectic structure theory to mechanics is not discussed in any detail*". Koszul considers in this book purely algebraic and geometric developments of Geometric/Analytic Mechanics developed during the 60th, more especially Jean-Marie Souriau works detailed in Chaps. 4 and 5. ***The originality of this book lies in the fact that Koszul develops new points of view, and demonstrations not considered initially by Souriau and Geometrical Mechanics community***.

Jean-Marie Souriau was the Creator of a new discipline called "*Mécanique Géométrique (Geometric Mechanics)*". Souriau observed that the collection of motions of a dynamical system is a manifold with an antisymmetric flat tensor that is a symplectic form where the structure contains all the pertinent information on the state of the system (positions, velocities, forces, etc.). Souriau said: "*[Ce que Lagrange a vu, que n'a pas vu Laplace, c'était la structure symplectique] What Lagrange saw, that Laplace didn't see, was the symplectic structure*". Using the symmetries of a symplectic manifold, Souriau introduced a mapping which he called the "*moment map*", which takes its values in a space attached to the group of symmetries (in the dual space of its Lie algebra). Souriau associated to this moment map, the notion of symplectic cohomology, linked to the fact that such a moment is defined up to an additive constant that brings into play an algebraic mechanism (called cohomology). Souriau proved that the moment map is a constant of the motion, and provided geometric generalization of Emmy Noether invariant theorem (invariants of E. Noether theorem are the components of the moment map). Souriau has defined in a geometrically way the Noetherian symmetries using the Lagrange-Souriau 2 form with the application map. Influenced by François Gallissot (Souriau and Galissot both attended ICM'54 in Moscow, and should have exchanged during this conference), Souriau has introduced in Mechanics the Lagrange 2-form, recovering seminal Lagrange ideas. Motivated by variational principles in a coordinate free formulation, inspired by Henri Poincaré and Elie Cartan who introduced a differential 1-form instead of the Lagrangian, Souriau introduced the Lagrange 2-form as the exterior differential of the Poincaré-Cartan 1-form, and obtained the phase space as a symplectic manifold. Souriau proposed to consider this Lagrange 2-form as the fundamental structure for Lagrangian system and not the classical Lagrangian function or the Poincaré-Cartan 1-form. This 2-form is called Lagrange-Souriau 2 form, and is the exterior derivative of the Lepage form (the Poincaré-Cartan form is a first order Lepage form). This structure is developed in Koszul book, where the authors shows that when $(M; \omega)$ is an exact symplectic manifold (when there exists a 1-form α on M such that $\omega = -d\alpha$), and that a symplectic action leaves not only ω, but α invariant, this action is strongly Hamiltonian ($(M; \omega)$ is a g-Poisson space). Koszul shows that a symplectic action of a Lie algebra g on an exact symplectic manifold $(M; \omega = -d\alpha)$ that leaves invariant not only ω, but also α, is strongly Hamiltionian.

In this Book in Chap. 4, Koszul calls symplectic G-space a symplectic manifold $(M; \omega)$ on which a Lie group G acts by a symplectic action (an action which leaves unchanged the symplectic form ω). Koszul then introduces and develop properties of the moment map μ (Souriau's invention) of a Hamiltonian action of the Lie algebra g. Koszul also defines the Souriau 2-cocycle, considering that the difference of two moments of the same Hamiltonian action is a locally constant application on M, showing that when μ is a moment map, for every pair *(a;b)* of elements of g, the function $c_\mu(a, b) = \{\langle\mu, a\rangle, \langle\mu, b\rangle\} - \langle\mu, \{a, b\}\rangle$ is locally constant on M, defining an antisymmetric bilinear application of gxg in $H^0(M; R)$ which verifies Jacobi's identity. ***This is the 2-cocycle introduced by Jean-Marie Souriau in Geometric Mechanics, that will play a fundamental role in Souriau Lie Groups Thermodynamics to define an extension of the Fisher Metric from Information Geometry (what I will call Fisher-Souriau metric in the following).***

To highlight the importance of this Koszul book, we will illustrate the links of the detailed tools, including demonstrations or original Koszul extensions, with Souriau's Lie Groups Thermodynamics, whose applications range from statistical physics to machine learning in Artificial Intelligence. In 1970, Souriau introduced the concept of co-adjoint action of a group on its momentum space, based on the orbit method works, that allows to define physical observables like energy, heat and momentum or moment as pure geometrical objects. In a first step to establish new foundations of thermodynamics, Souriau has defined a Gibbs canonical ensemble on a symplectic manifold M for a Lie group action on M. 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*),* Souriau has proved that when statistical states are Gibbs states (as generalized by Souriau), they are the product of the Liouville measure by the scalar function given by the generalized partition function $e^{\Phi(\beta)-\langle\beta,U(\xi)\rangle}$ defined by the energy U (defined in the dual of the Lie algebra of this dynamical group) and the geometric temperature β, where Φ is a normalizing constant such the mass of probability is equal to 1, $\Phi(\beta) = -\log\int\limits_M e^{-\langle\beta,U(\xi)\rangle}d\lambda$. Jean-Marie Souriau then generalizes the Gibbs equilibrium state to all symplectic manifolds that have a dynamical group. Souriau has observed that if we apply this theory for Galileo, the symmetry will be broken. For each temperature β, element of the Lie algebra $\mathfrak{g}$, Souriau has introduced a tensor $\tilde{\Theta}_\beta$, equal to the sum of the cocycle $\tilde{\Theta}$ and the heat coboundary (with [.,.] Lie bracket):

$$\tilde{\Theta}_\beta(Z_1, Z_2) = \tilde{\Theta}(Z_1, Z_2) + \langle Q, ad_{Z1}(Z_2)\rangle \tag{122}$$

This tensor $\tilde{\Theta}_\beta$ has the following properties: $\tilde{\Theta}(X, Y) = \langle\Theta(X), Y\rangle$ where the map Θ is the symplectic one-cocycle of the Lie algebra $\mathfrak{g}$ with values in $\mathfrak{g}^*$, with $\Theta(X) = T_e\theta(X(e))$ where θ the one-cocycle of the Lie group G. $\tilde{\Theta}(X, Y)$ is constant on M and the map $\tilde{\Theta}(X, Y) : \mathfrak{g}\times\mathfrak{g} \to \Re$ is a skew-symmetric bilinear form, and is called the ***symplectic two-cocycle of Lie algebra*** $\mathfrak{g}$ associated to the *moment map* J, with the following properties:

$$\tilde{\Theta}(X, Y) = J_{[X,Y]} - \{J_X, J_Y\}\text{with } J \text{ the Moment Map} \tag{123}$$

$$\tilde{\Theta}([X,Y],Z)+\tilde{\Theta}([Y,Z],X)+\tilde{\Theta}([Z,X],Y)=0 \tag{124}$$

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

$$J:M\to\mathfrak{g}^*, x\mapsto J(x)\quad \text{such that } J_X(x)=\langle J(x),X\rangle,\ X\in\mathfrak{g} \tag{125}$$

The geometric temperature, element of the algebra $\mathfrak{g}$, is in the kernel of the tensor $\tilde{\Theta}_\beta$:

$$\beta\in Ker\,\tilde{\Theta}_\beta \text{ such that } \tilde{\Theta}_\beta(\beta,\beta)=0,\quad \forall\beta\in\mathfrak{g} \tag{126}$$

The following symmetric tensor $g_\beta([\beta,Z_1],[\beta,Z_2])=\tilde{\Theta}_\beta(Z_1,[\beta,Z_2])$, defined on all values of $ad_\beta(.)=[\beta,.]$ is positive definite, and defines extension of classical Fisher metric in Information Geometry (as hessian of the logarithm of partition function):

$$g_\beta([\beta,Z_1],Z_2)=\tilde{\Theta}_\beta(Z_1,Z_2),\quad \forall Z_1\in\mathfrak{g},\forall Z_2\in \mathrm{Im}\big(ad_\beta(.)\big) \tag{127}$$

$$\text{With } g_\beta(Z_1,Z_2)\geq 0,\quad \forall Z_1,Z_2\in \mathrm{Im}\big(ad_\beta(.)\big) \tag{128}$$

***These equations are universal*, because they are not dependent on the symplectic manifold but only on the dynamical group *G*, the symplectic two-cocycle Θ, the temperature β and the heat Q. Souriau called it "*Lie groups thermodynamics*".**

This antisymmetric bilinear map (127) and (128), with definition (122) and (123) is exactly equal to the mathematical object introduced in Chap. 4 of Koszul's book by:

$$c_\mu(a,b)=\{\langle\mu,a\rangle,\langle\mu,b\rangle\}-\{\mu,\langle a,b\rangle\} \tag{129}$$

In this book, Koszul has studied this antisymmetric bilinear map considering the following developments. For any moment map μ, Koszul defines the skew symmetric bilinear form $\mathrm{c}_\mu(\mathrm{a},\mathrm{b})$ on Lie algebra by:

$$\mathrm{c}_\mu(\mathrm{a},\mathrm{b})=\big\langle d\theta_\mu(a),b\big\rangle, a,b\in\mathfrak{g} \tag{130}$$

Koszul observes that if we use:

$$\theta_\mu(st)=\mu(stx)-Ad^*_{st}\mu(x)=\theta_\mu(s)+Ad^*_s\mu(tx)-Ad^*_sAd^*_t\mu(x)=\theta_\mu(s)+Ad^*_s\theta_\mu(t)$$

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

$$\langle d\mu(ax),b\rangle=\langle\mu(x),[a,b]\rangle+\big\langle \mathrm{d}\theta_\mu(a),b\big\rangle=\{\langle\mu,a\rangle,\langle\mu,b\rangle\}(x),\ x\in M,a,b\in\mathfrak{g} \tag{131}$$

We have then:

$$c_\mu(\mathrm{a, b}) = \{\langle\mu, a\rangle, \langle\mu, b\rangle\} - \langle\mu, [a, b]\rangle = \left\langle \mathrm{d}\theta_\mu(a), b\right\rangle, a, b \in \mathfrak{g} \quad (132)$$

and the property:

$$c_\mu([\mathrm{a, b}], \mathrm{c}) + c_\mu([b, c], a) + c_\mu([c, a], b) = 0,\ a, b, c \in \mathfrak{g} \quad (133)$$

Koszul concludes by observing that if the moment map is transform as $\mu' = \mu + \phi$ then we have:

$$c_{\mu'}(a, b) = c_\mu(a, b) - \langle\phi, [a, b]\rangle \quad (134)$$

Finally using $c_\mu(\mathrm{a, b}) = \{\langle\mu, a\rangle, \langle\mu, b\rangle\} - \langle\mu, [a, b]\rangle = \left\langle \mathrm{d}\theta_\mu(a), b\right\rangle, a, b \in \mathfrak{g}$, koszul highlights the property that:

$$\left\{\mu^*(a), \mu^*(b)\right\} = \{\langle\mu, a\rangle, \langle\mu, b\rangle\} = \mu^*\big([a, b] + c_\mu(a, b)\big) = \mu^*\{a, b\}_{c_\mu} \quad (135)$$

In Chap. 4, Koszul introduces the equivariance of the moment map μ. Based on the definitions of the adjoint and coadjoint representations of a Lie group or a Lie algebra, Koszul proves that when *(M; ω)* is a connected Hamiltonian *G*-space and $\mu : M \to \mathfrak{g}^*$ a moment of the action of *G*, there exists an affine action of *G* on *g**, whose linear part is the coadjoint action, for which the moment μ is equivariant. This affine action is obtained by modifying the coadjoint action by means of a cocycle. This notion is also developed in Chap. 5 for Poisson manifolds. Defining classical operation $Ad_s a = sas^{-1}, s \in G, a \in \mathfrak{g}, ad_a b = [a, b], a \in \mathfrak{g}, b \in \mathfrak{g}$ and coadjoint action given by $Ad_s^* = {}^t\ Ad_{s^{-1}}, s \in G$ with classical properties:

$$Ad_{\exp a} = \exp(-ad_a), a \in \mathfrak{g} \text{ or } Ad^*_{\exp a} = \exp^t(ad_a), a \in \mathfrak{g} \quad (136)$$

Koszul considers:

$$x \mapsto sx, x \in M, \mu : M \to \mathfrak{g}^* \quad (137)$$

From which, he obtains:

$$\langle d\mu(v), a\rangle = \omega(ax, v) \quad (138)$$

Koszul then study $\mu \circ s_M - Ad_s^* \circ \mu : M \to \mathfrak{g}^*$, and develops:

$$d\left\langle Ad_s^* \circ \mu, a\right\rangle = \left\langle Ad_s^* d\mu, a\right\rangle = \langle d\mu, Ad_{s^{-1}} a\rangle \quad (139)$$

$$\langle d\mu(v), Ad_{s^{-1}} a\rangle = \omega(s^{-1}asx, v) = \omega(asx, sv) = \langle d\mu(sv), a\rangle = (d\langle\mu \circ s_M, a\rangle)(v) \quad (140)$$

$d\langle Ad_s^* \circ \mu, a\rangle = d\langle \mu \circ s_M, a\rangle$ and then proves that $d\langle \mu \circ s_M - Ad_s^* \circ \mu, a\rangle = 0$ (141)

Koszul considers the cocycle given by $\theta_\mu(s) = \mu(sx) - Ad_s^* \mu(x), s \in G$, and observes that:

$$\theta_\mu(st) = \theta_\mu(s) - Ad_s^* \theta_\mu(t), s, t \in G \tag{142}$$

From this action of the group on dual Lie algebra:

$$G \times \mathfrak{g}^* \to \mathfrak{g}^*, (s, \xi) \mapsto s\xi = Ad_s^* \xi + \theta_\mu(s) \tag{143}$$

Koszul introduces the following properties:

$$\mu(sx) = s\mu(x) = Ad_s^* \mu(x) + \theta_\mu(s), \forall s \in G, x \in M \tag{144}$$

$$G \times \mathfrak{g}^* \to \mathfrak{g}^*, (e, \xi) \mapsto e\xi = Ad_e^* \xi + \theta_\mu(e) = \xi + \mu(x) - \mu(x) = \xi \tag{145}$$

$$(s_1 s_2)\xi = Ad_{s_1 s_2}^* \xi + \theta_\mu(s_1 s_2) = Ad_{s_1}^* Ad_{s_2}^* \xi + \theta_\mu(s_1) + Ad_{s_1}^* \theta_\mu(s_2)$$
$$(s_1 s_2)\xi = Ad_{s_1}^* (Ad_{s_2}^* \xi + \theta_\mu(s_2)) + \theta_\mu(s_1) = s_1(s_2 \xi), \forall s_1, s_2 \in G, \xi \in \mathfrak{g}^* \tag{146}$$

This Koszul study of the moment map μ equivariance, and the existence of an affine action of G on g^*, whose linear part is the coadjoint action, for which the moment μ is equivariant, is at the cornerstone of Souriau Theory of Geometric Mechanics and Lie Groups Thermodynamics. I illustrate this importance by giving Souriau theorem for Lie Groups Thermodynamics, and the link with, what I call, Souriau-Fisher metric (a covariant definition of Fisher metric):

Theorem (**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 $Ad_g(.)$. Then, the fundamental equations of Lie group thermodynamics are given by the action of the group:*

$$Action\ of\ Lie\ group\ on\ Lie\ algebra : \beta \to Ad_g(\beta) \tag{147}$$

$$Characteristic\ function\ after\ Lie\ group\ action : \Phi \to \Phi - \langle \theta(g^{-1}), \beta \rangle \tag{148}$$

$$Invariance\ of\ entropy\ with\ respect\ to\ action\ of\ Lie\ group : s \to s \tag{149}$$

$$Action\ of\ Lie\ group\ on\ geometric\ heat : Q \to a(g, Q) = Ad_g^*(Q) + \theta(g) \tag{150}$$

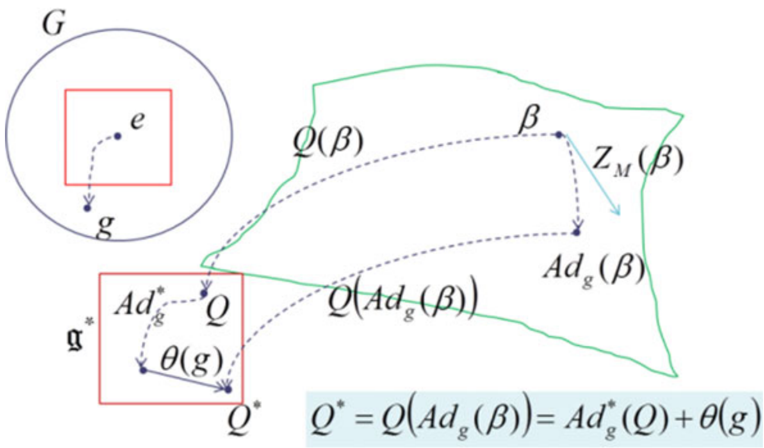


Fig. 11 Broken symmetry on geometric heat Q due to adjoint action of the group on temperature β as an element of the Lie algebra

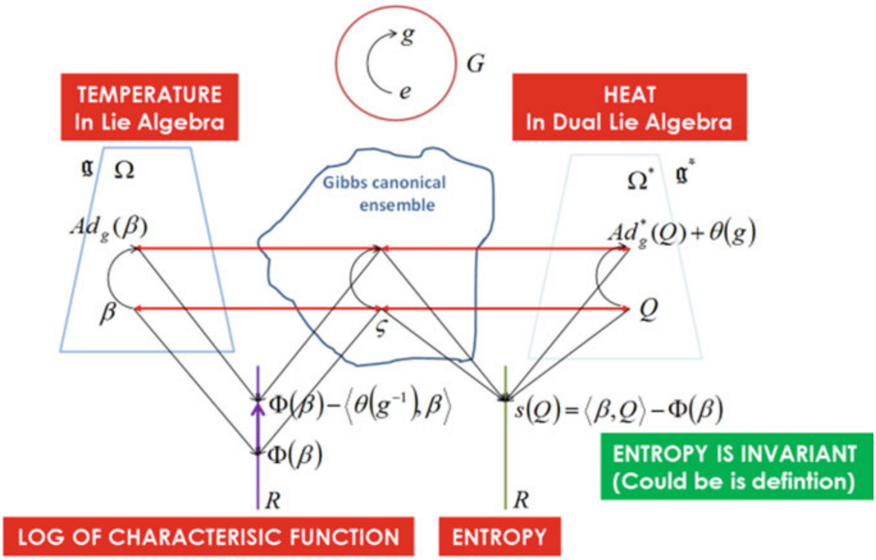


Fig. 12 Global Souriau scheme of Lie group thermodynamics, with entropy *s(Q)*, geometric heat Q element of dual Lie algebra and geometric temperature β element of Lie algebra

Souriau equations of Lie group thermodynamics, related to the moment map μ equivariance, and the existence of an affine action of G on g^*, whose linear part is the coadjoint action, for which the moment μ is equivariant, are summarized in the following figures (Figs. 11 and 12).

I finally observe that ***the Koszul antisymmetric bilinear map*** $c_\mu(a,b) = \{\langle\mu, a\rangle, \langle\mu, b\rangle\} - \langle\mu, \{a, b\}\rangle$ ***is equal to Souriau Riemannian metric***, introduced by mean of symplectic cocycle. I have observed that this metric is a generalization of the Fisher metric from Information Geometry, that I call the Souriau-Fisher metric, defined as a hessian of the partition function logarithm $g_\beta = -\frac{\partial^2\Phi}{\partial\beta^2} = \frac{\partial^2 \log \psi_\Omega}{\partial\beta^2}$ as in classical information geometry. This new definition of Fisher metric has the property to be covariant under the action of the group G. I have established the equality of two terms, between Souriau definition based on Lie group cocycle Θ and parameterized by "*geometric heat*" Q (element of dual Lie algebra) and "*geometric temperature*" β (element of Lie algebra) and hessian of characteristic function $\Phi(\beta) = -\log\Psi_\Omega(\beta)$ with respect to the variable β:

$$g_\beta([\beta, Z_1], [\beta, Z_2]) = \langle\Theta(Z_1), [\beta, Z_2]\rangle + \langle Q, [Z_1, [\beta, Z_2]]\rangle = \frac{\partial^2 \log \psi_\Omega}{\partial\beta^2} \tag{151}$$

If we differentiate this relation of Souriau theorem $Q(Ad_g(\beta)) = Ad_g^*(Q) + \theta(g)$, this relation occurs:

$$\frac{\partial Q}{\partial\beta}(-[Z_1, \beta], .) = \tilde{\Theta}(Z_1[\beta, .]) + \left\langle Q, Ad_{z_1}([\beta, .])\right\rangle = \tilde{\Theta}_\beta(Z_1, [\beta, .]) \tag{152}$$

$$-\frac{\partial Q}{\partial\beta}([Z_1, \beta], Z_2.) = \tilde{\Theta}(Z_1, [\beta, Z_2]) + \left\langle Q, Ad_{.z_1}([\beta, Z_2])\right\rangle = \tilde{\Theta}_\beta(Z_1, [\beta, Z_2]) \tag{153}$$

$$\Rightarrow -\frac{\partial Q}{\partial\beta} = g_\beta([\beta, Z_1], [\beta, Z_2]) \tag{154}$$

The Souriau Fisher metric $I(\beta) = -\frac{\partial^2\Phi(\beta)}{\partial\beta^2} = -\frac{\partial Q}{\partial\beta}$ has been considered by Souriau as a *generalization of* "*heat capacity*". Souriau called it the "*geometric capacity*" and is also equal to "*geometric susceptibility*".

7 Conclusion

The community of "*Geometric Science of Information*" (GSI) has lost a mathematician of great value, who informed his views by the depth of his knowledge of the elementary structures of hessian geometry and bounded homogeneous domains. His modesty was inversely proportional to his talent. Professor Koszul built in over 60 years of mathematical career, in the silence of his passions, an immense work, which makes him one of the great mathematicians of the XX's century, whose importance will only affirm with the time. In this troubled time and rapid transformation of society and science, the example of Professor Koszul must be regarded as a model for future generations, to avoid them the trap of fleeting glories and recognitions too fast acquired. The work of Professor Koszul is also a proof of fidelity to his masters

Fig. 13 (on the left) Jean-Louis Koszul at Grenoble in December 1993, (on the right) last interview of Jean-Louis Koszul in 2016 for 50th birthday of Institut Joseph Fourier in Grenoble

and in the first place to Prof. Elie Cartan, who inspired him throughout his life. Henri Cartan writes on this subject "*I do not forget the homage he paid to Elie Cartan's work in Differential Geometry during the celebration, in Bucharest, in 1969, of the centenary of his birth. It is not a coincidence that this centenary was also celebrated in Grenoble the same year. As always, Koszul spoke with the discretion and tact that we know him, and that we love so much at home*". I will conclude by quoting Jorge Luis Borges "*Forgetfulness and memory are also inventive*" (Brodie's report). Our generation and previous one have forgotten or misunderstood the depth of the work of Jean-Louis Koszul and Elie Cartan on the study of bounded homogeneous domains. It is our responsibility to correct this omission, and to make it the new inspiration for the Geometric Science of Information. I will conclude by requesting you to listen to the last interview of Jean-Louis Koszul for 50th birthday of Joseph Fourier Institute [72], especially when Koszul he is passionate by "*conifers and cedars trees planted by Claude Chabauty*", or by the "*pretty catalpa tree*" which was at the Fourier Institute and destroyed by wind, "*the tree with parentheses*" he says, to which he seemed to be sentimentally attached. He also regrets that the Institute did not use the 1% artistic fund for the art mosaic project in the library. In this Koszul family of mathematicians, musicians, and Scientifics, there was a constant recollection of "beauty" and "truth". Our society no longer cares about timeless "beauty". We have then to extase ourself with Jean-Louis Koszul by observing beautiful "*Catalpa tree*" with "*Parenthese Mushroom*", before there is no longer people to contemplate them (Fig. 13).

> "*Seul la nuit avec un livre éclairé par une chandelle – livre et chandelle, double îlot de lumière, contre les doubles ténèbres de l'esprit et de la nuit. J'étudie ! Je ne suis que le sujet du verbe étudier. Penser je n'ose. Avant de penser, il faut étudier. Seuls les philosophes pensent avant d'étudier.* » - **Gaston Bachelard, La flamme d'une chandelle, 1961**

Appendix

Clairaut(-Legendre) Equation of Maurice Fréchet associated to "distinguished functions" as fundamental equation of Information geometry

Before Rao [4, 124], in 1943, Maurice Fréchet [3] wrote a seminal paper introducing what was then called the Cramer-Rao bound. This paper contains in fact much more that this important discovery. In particular, Maurice Fréchet introduces more general notions relative to "*distinguished functions*", densities with estimator reaching the bound, defined with a function, solution of Clairaut's equation. The solutions "envelope of the Clairaut's equation" are equivalents to standard Legendre transform without convexity constraints but only smoothness assumption. This Fréchet's analysis can be revisited on the basis of Jean-Louis Koszul's works as seminal foundation of "*Information Geometry*".

I will use Maurice Fréchet notations, to consider the estimator:

$$T = H(X_1, \ldots, X_n) \tag{155}$$

and the random variable

$$A(X) = \frac{\partial \log p_\theta(X)}{\partial \theta} \tag{156}$$

that are associated to:

$$U = \sum_i A(X_i) \tag{157}$$

The normalizing constraint $\int_{-\infty}^{+\infty} p_\theta(x)dx = 1$ implies that: $\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \prod_i p_\theta(x_i)dx_i = 1$

If we consider the derivative if this last expression with respect to θ, then $\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \left[\sum_i A(x_i)\right] \prod_i p_\theta(x_i)dx_i = 0$ gives:

$$E_\theta[U] = 0 \tag{158}$$

Similarly, if we assume that $E_\theta[T] = \theta$, then $\int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} H(x_1, \ldots, x_n) \prod_i p_\theta(x_i)dx_i = \theta$, and we obtain by derivation with respect to θ:

$$E[(T - \theta)U] = 1 \tag{159}$$

But as $E[T] = \theta$ and $E[U] = 0$, we immediatly deduce that:

$$E[(T - E[T])(U - E[U])] = 1 \tag{160}$$

From Schwarz inequality, we can develop the following relations:

$$\begin{aligned} &[E(ZT)]^2 \leq E\left[Z^2\right]E\left[T^2\right] \\ &1 \leq E\left[(T - E[T])^2\right]E\left[(U - E[U])^2\right] = (\sigma_T \sigma_U)^2 \end{aligned} \tag{161}$$

U being the summation of independent variables, Bienaymé equality could be applied:

$$(\sigma_U)^2 = \sum_i \left[\sigma_{A(X_i)}\right]^2 = n(\sigma_A)^2 \tag{162}$$

From which, Fréchet deduced the bound, rediscoved by Cramer and Rao 2 years later:

$$(\sigma_T)^2 \geq \frac{1}{n(\sigma_A)^2} \tag{163}$$

Fréchet observed that it is a remarkable inequality where the second member is independent of the choice of the function H defining the "*empirical value*" T, where the first member can be taken to any empirical value $T = H(X_1, \ldots, X_n)$ subject to the unique condition $E_\theta[T] = \theta$ regardless is θ.

The classic condition that the Schwarz inequality becomes an equality helps us to determine when σ_T reaches its lower bound $\frac{1}{\sqrt{n}\sigma_n}$.

The previous inequality becomes an equality if there are two numbers α and β (not random and not both zero) such that $\alpha\left(H' - \theta\right) + \beta U = 0$, with H' particular function among eligible H as we have the equality. This equality is rewritten $H' = \theta + \lambda' U$ with λ' a non-random number.

If we use the previous equation, then:

$$E[(T - E[T])(U - E[U])] = 1 \Rightarrow E\left[\left(H' - \theta\right)U\right] = \lambda' E_\theta\left[U^2\right] = 1 \tag{164}$$

We obtain:

$$U = \sum_i A(X_i) \Rightarrow \lambda' n E_\theta\left[A^2\right] = 1 \tag{165}$$

From which we obtain λ' and the form of the associated estimator H':

$$\lambda' = \frac{1}{nE\left[A^2\right]} \Rightarrow H' = \theta + \frac{1}{nE\left[A^2\right]} \sum_i \frac{\partial \log p_\theta(X_i)}{\partial \theta} \tag{166}$$

It is therefore deduced that the estimator that reaches the terminal is of the form:

$$H' = \theta + \frac{\sum_i \frac{\partial \log p_\theta(X_i)}{\partial \theta}}{n \int\limits_{-\infty}^{+\infty} \left[\frac{\partial p_\theta(x)}{\partial \theta} \right]^2 \frac{dx}{p_\theta(x)}} \tag{167}$$

with $E\left[H'\right] = \theta + \lambda' E[U] = \theta$.

H' would be one of the eligible functions, if H' would be independent of θ. Indeed, if we consider:

$$E_{\theta_0}\left[H'\right] = \theta_0,\ E\left[\left(H' - \theta_0\right)^2\right] \leq E_{\theta_0}\left[(H - \theta_0)^2\right] \forall H \text{ such that } E_{\theta_0}[H] = \theta_0 \tag{168}$$

$H = \theta_0$ satisfies the equation and inequality shows that it is almost certainly equal to θ_0. So to look for θ_0, we should know beforehand θ_0.

At this stage, Fréchet looked for "*distinguished functions*" ("*densités distinguées*" in French), as any probability density $p_\theta(x)$ such that the function:

$$h(x) = \theta + \frac{\frac{\partial \log p_\theta(x)}{\partial \theta}}{\int\limits_{-\infty}^{+\infty} \left[\frac{\partial p_\theta(x)}{\partial \theta} \right]^2 \frac{dx}{p_\theta(x)}} \tag{169}$$

is independant of θ. The objective of Fréchet is then to determine the minimizing function $T = H'(X_1, \ldots, X_n)$ that reaches the bound. By deduction from previous relations, we have:

$$\lambda(\theta) \frac{\partial \log p_\theta(x)}{\partial \theta} = h(x) - \theta \tag{170}$$

But as $\lambda(\theta) > 0$, ***we can consider $\frac{1}{\lambda(\theta)}$ as the second derivative of a function*** $\Phi(\theta)$ such that:

$$\frac{\partial \log p_\theta(x)}{\partial \theta} = \frac{\partial^2 \Phi(\theta)}{\partial \theta^2} [h(x) - \theta] \tag{171}$$

from which we deduce that:

$$\ell(x) = \log p_\theta(x) - \frac{\partial \Phi(\theta)}{\partial \theta} [h(x) - \theta] - \Phi(\theta) \tag{172}$$

Is an independant quantity of θ. ***A distinguished function*** will be then given by:

$$p_\theta(x) = e^{\frac{\partial \Phi(\theta)}{\partial \theta}[h(x) - \theta] + \Phi(\theta) + \ell(x)} \tag{173}$$

with the normalizing constraint $\int\limits_{-\infty}^{+\infty} p_\theta(x) dx = 1$.

These two conditions are sufficient. Indeed, reciprocally, let three functions $\Phi(\theta)$, $h(x)$ et $\ell(x)$ that we have, for any θ:

$$\int_{-\infty}^{+\infty} e^{\frac{\partial \Phi(\theta)}{\partial \theta}[h(x)-\theta]+\Phi(\theta)+\ell(x)} dx = 1 \tag{174}$$

Then the function is distinguished:

$$\theta + \frac{\frac{\partial \log p_\theta(x)}{\partial \theta}}{\int_{-\infty}^{+\infty} \left[\frac{\partial p_\theta(x)}{\partial \theta}\right]^2 \frac{dx}{p_\theta(x)}} = \theta + \lambda(x)\frac{\partial^2 \Phi(\theta)}{\partial \theta^2}[h(x)-\theta] \tag{175}$$

If $\lambda(x)\frac{\partial^2 \Phi(\theta)}{\partial \theta^2} = 1$, when

$$\frac{1}{\lambda(x)} = \int_{-\infty}^{+\infty} \left[\frac{\partial \log p_\theta(x)}{\partial \theta}\right]^2 p_\theta(x) dx = (\sigma_A)^2 \tag{176}$$

The function is reduced to $h(x)$ and then is not dependent of θ.
We have then the following relation:

$$\frac{1}{\lambda(x)} = \int_{-\infty}^{+\infty} \left(\frac{\partial^2 \Phi(\theta)}{\partial \theta^2}\right)^2 [h(x)-\theta]^2 e^{\frac{\partial \Phi(\theta)}{\partial \theta}(h(x)-\theta)+\Phi(\theta)+\ell(x)} dx \tag{177}$$

The relation is valid for any θ, we can derive the previous expression from θ:

$$\int_{-\infty}^{+\infty} e^{\frac{\partial \Phi(\theta)}{\partial \theta}(h(x)-\theta)+\Phi(\theta)+\ell(x)} \left(\frac{\partial^2 \Phi(\theta)}{\partial \theta^2}\right)[h(x)-\theta] dx = 0 \tag{178}$$

We can divide by $\frac{\partial^2 \Phi(\theta)}{\partial \theta^2}$ because it doesn't depend on x.
If we derive again with respect to θ, we will have:

$$\int_{-\infty}^{+\infty} e^{\frac{\partial \Phi(\theta)}{\partial \theta}(h(x)-\theta)+\Phi(\theta)+\ell(x)} \left(\frac{\partial^2 \Phi(\theta)}{\partial \theta^2}\right)[h(x)-\theta]^2 dx = \int_{-\infty}^{+\infty} e^{\frac{\partial \Phi(\theta)}{\partial \theta}(h(x)-\theta)+\Phi(\theta)+\ell(x)} dx = 1 \tag{179}$$

Combining this relation with that of $\frac{1}{\lambda(x)}$, we can deduce that $\lambda(x)\frac{\partial^2 \Phi(\theta)}{\partial \theta^2} = 1$ and as $\lambda(x) > 0$ then $\frac{\partial^2 \Phi(\theta)}{\partial \theta^2} > 0$.

Fréchet emphasizes at this step, another way to approach the problem. We can select arbitrarily $h(x)$ and $l(x)$ and then $\Phi(\theta)$ is determined by:

$$\int_{-\infty}^{+\infty} e^{\frac{\partial\Phi(\theta)}{\partial\theta}[h(x)-\theta]+\Phi(\theta)+\ell(x)}dx = 1 \tag{180}$$

That could be rewritten:

$$e^{\theta.\frac{\partial\Phi(\theta)}{\partial\theta}-\Phi(\theta)} = \int_{-\infty}^{+\infty} e^{\frac{\partial\Phi(\theta)}{\partial\theta}h(x)+\ell(x)}dx \tag{181}$$

If we then fixed arbitrarily $h(x)$ and $l(x)$ and let s an arbitrary variable, the following function will be an explicit positive function given by $e^{\Psi(s)}$:

$$\int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)}dx = e^{\Psi(s)} \tag{182}$$

Fréchet obtained finally the function $\Phi(\theta)$ as solution of the equation:

$$\Phi(\theta) = \theta.\frac{\partial\Phi(\theta)}{\partial\theta} - \Psi\left(\frac{\partial\Phi(\theta)}{\partial\theta}\right) \tag{183}$$

Fréchet noted that this is the Alexis Clairaut Equation.

The case $\frac{\partial\Phi(\theta)}{\partial\theta} = cste$ would reduce the density to a function that would be independent of θ, and so $\Phi(\theta)$ is given by a singular solution of this Clairaut equation, that is unique and could be computed by eliminating the variable s between:

$$\Phi = \theta.s - \Psi(s) \text{ and } \theta = \frac{\partial\Psi(s)}{\partial s} \tag{184}$$

Or between:

$$e^{\theta.s-\Phi(\theta)} = \int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)}dx \text{ and } \int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)}[h(x)-\theta]dx = 0 \tag{185}$$

$\Phi(\theta) = -\log \int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)}dx + \theta.s$ where s is induced implicitly through the constraint given by the integral $\int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)}[h(x)-\theta]dx = 0$.

When we known the distinguished function, H' is among functions $H(X_1, \ldots, X_n)$ verifying $E_\theta[H] = \theta$ and such that σ_H reaches for each value of θ, an absolute minimum, equal to $\frac{1}{\sqrt{n}\sigma_A}$. For the previous equation:

$$h(x) = \theta + \frac{\frac{\partial \log p_\theta(x)}{\partial \theta}}{\int\limits_{-\infty}^{+\infty} \left[\frac{\partial p_\theta(x)}{\partial \theta} \right]^2 \frac{dx}{p_\theta(x)}} \tag{186}$$

We can rewrite the estimator as:

$$H'(X_1, \ldots, X_n) = \frac{1}{n}[h(X_1) + \ldots + h(X_n)] \tag{187}$$

and compute the associated empirical value:

$$t = H'(x_1, \ldots, x_n) = \frac{1}{n} \sum_i h(x_i) = \theta + \lambda(\theta) \sum_i \frac{\partial \log p_\theta(x_i)}{\partial \theta} \tag{188}$$

If we take $\theta = t$, we have as $\lambda(\theta) > 0$:

$$\sum_i \frac{\partial \log p_t(x_i)}{\partial t} = 0 \tag{189}$$

When $p_\theta(x)$ is a distinguished function, the empirical value t of θ corresponding to a sample $x_1, \ldots, x_n$ is a root of previous equation in t. This equation has a root and only one when X is a distinguished variable. Indeed, as we have:

$$p_\theta(x) = e^{\frac{\partial \Phi(\theta)}{\partial \theta}[h(x) - \theta] + \Phi(\theta) + \ell(x)} \tag{190}$$

$$\sum_i \frac{\partial \log p_t(x_i)}{\partial t} = \frac{\partial^2 \Phi(t)}{\partial t^2} \left[\frac{\sum\limits_i h(x_i)}{n} - t \right] \text{with } \frac{\partial^2 \Phi(t)}{\partial t^2} > 0 \tag{191}$$

We can then recover the unique root: $t = \frac{\sum\limits_i h(x_i)}{n}$.

This function $T \equiv H'(X_1, \ldots, X_n) = \frac{1}{n} \sum\limits_i h(X_i)$ can have an arbitrary form, that is a sum of functions of each only one of the quantities and it is even the arithmetic average of N values of a same auxiliary random variable $Y = h(X)$. The dispersion is given by:

$$\left(\sigma_{T_n}\right)^2 = \frac{1}{n(\sigma_A)^2} = \frac{1}{n \int\limits_{-\infty}^{+\infty} \left[\frac{\partial p_\theta(x)}{\partial \theta} \right]^2 \frac{dx}{p_\theta(x)}} = \frac{1}{n \frac{\partial^2 \Phi(\theta)}{\partial \theta^2}} \tag{192}$$

and T_n follows the probability density:

$$p_\theta(t) = \sqrt{n} \frac{1}{\sigma_A \sqrt{2\pi}} e^{-\frac{n(t-\theta)^2}{2.\sigma_A^2}} \quad \text{with} \quad (\sigma_A)^2 = \frac{\partial^2 \Phi(\theta)}{\partial \theta^2} \tag{193}$$

- **Clairaut Equation and Legendre Transform**

To summarize, Fréchet paper novelty, I have just observed that Fréchet introduced distinguished functions depending on a function $\Phi(\theta)$, solution of the Clairaut equation:

$$\Phi(\theta) = \theta.\frac{\partial \Phi(\theta)}{\partial \theta} - \Psi\left(\frac{\partial \Phi(\theta)}{\partial \theta}\right) \tag{194}$$

Or given by the Legendre Transform:

$$\Phi = \theta.s - \Psi(s) \text{ and } \theta = \frac{\partial \Psi(s)}{\partial s} \tag{195}$$

Fréchet also observed that this function $\Phi(\theta)$ could be rewritten: $\Phi(\theta) = -\log \int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)}dx + \theta.s$ where s is induced implicitly by the constraints given by integral $\int_{-\infty}^{+\infty} e^{s.h(x)+\ell(x)}[h(x) - \theta]dx = 0$.

This equation is the fundamental equation of Information Geometry.

The "Legendre" transform was introduced by Adrien-Marie Legendre in 1787 to solve a minimal surface problem Gaspard Monge in 1784. Using a result of Jean Baptiste Meusnier, a student of Monge, it solves the problem by a change of variable corresponding to the transform which now entitled with his name. Legendre wrote: "*I have just arrived by a change of variables that can be useful in other occasions.*" About this transformation, Darboux in his book gives an interpretation of Chasles: "*This comes after a comment by Mr. Chasles, to substitute its polar reciprocal on the surface compared to a paraboloïd.*" The equation of Clairaut was introduced 40 years earlier in 1734 by Alexis Clairaut. Solutions "envelope of the Clairaut equation" are equivalent to the Legendre transform with unconditional convexity, but only under differentiability constraint. Indeed, for a non-convex function, Legendre transformation is not defined where the Hessian of the function is canceled, so that the equation of Clairaut only make the hypothesis of differentiability. The portion of the strictly convex function g in Clairaut equation $y = px - g(p)$ to the function f giving the envelope solutions by the formula $y = f(x)$ is precisely the Legendre transformation. The approach of Fréchet may be reconsidered in a more general context on the basis of the work of Jean-Louis Koszul.

References

1. Ollivier, Y., Marceau-Caron, G.: Natural Langevin dynamics for neural networks. In: Nielsen, F., Barbaresco, F. (eds.) Proceedings of the Conference on Geometric Science of Information (GSI 2017). Lecture Notes in Computer Science, vol. 10589, pp. 451–459. Springer, Berlin (2017) (Best paper award.)
2. Ollivier, Y.: True asymptotic natural gradient optimization manuscript disponible sur. http://arxiv.org/abs/1712.08449
3. Fréchet, M.: Sur l'extension de certaines évaluations statistiques au cas de petits échantillons. Rev. l'Institut Int. Stat. **11**(3/4), 182–205 (1943)
4. Rao, C.R.: Information and the accuracy attainable in the estimation of statistical parameters. Bull. Calcutta Math. Soc. **37**, 81–89 (1945)
5. Chentsov, N.N.: Statistical Decision Rules and Optimal Inferences, Transactions of Mathematics Monograph, vol. 53. American Mathematical Society, Providence, RI, USA, (1982)
6. Amari, S.I.: Differential Geometry of Statistical Models. SPRINGER Series Lecture Notes in Statistics, vol. 28 (1985)
7. Amari, S.I.: Information Geometry and Its Applications. SPRINGER Series Applied Mathematical Sciences, vol. 194 (2016)
8. Weinhold, F.: Thermodynamics and geometry. Phys. Today (1976)
9. Ruppeiner, G.: Thermodynamics: a Riemannian geometric model. Phys. Rev. A **20**, 1608 (1979)
10. Ingarden, R.S.: Information geometry in functional spaces of classical and quantum finite statistical systems. Int. J. Eng. Sci. **19**(12), 1609–1633 (1981)
11. Ingarden, R.S., Kossakowski, A., Ohya, M.: Information Dynamics and Open Systems: Classical and Quantum Approach. Springer, Berlin (1997). Accessed 31 March 1997
12. Mrugala, R.: Geometrical formulation of equilibrium phenomenological thermodynamics. Rep. Math. Phys. **14**, 419 (1978)
13. Mrugala, R., Nulton, J.D., Schön, J.C., Salamon, P.: Statistical approach to the geometric structure of thermodynamics. Phys. Rev. A **41**, 6, 3156 (1990)
14. Janyszek, H., Mrugala, R.: Riemannian geometry and the thermodynamics of model magnetic systems. Phys. Rev. A **39**(12), :6515–6523 (1989)
15. Koszul, J.L.: Sur la forme hermitienne canonique des espaces homogènes complexes. Can. J. Math. **7**, 562–576 (1955)
16. Koszul, J.L.: Exposés sur les Espaces Homogènes Symétriques. Publicação da Sociedade de Matematica de São Paulo, São Paulo, Brazil (1959)
17. Souriau, J.-M.: Structures des Systèmes Dynamiques. Dunod, Paris (1969)
18. Koszul, J.L.: Introduction to Symplectic Geometry. Science Press, Beijing (1986) (in chinese); translated by SPRINGER, with F.Barbaresco, C.M. Marle and M. Boyom forewords, SPRINGER, 2018
19. Balian, R., Alhassid, Y., Reinhardt, H.: Dissipation in many-body systems: a geometric approach based on information theory. Phys. Rep. **131**, 1–146 (1986)
20. Balian, R.: The entropy-based quantum metric. Entropy **16**, 3878–3888 (2014)
21. Selected Papers of J L Koszul, Series in Pure Mathematics, Volume 17, World Scientific Publishing, 1994
22. Koszul, J.L.: L'œuvre d'Élie Cartan en géométrie différentielle, in Élie Cartan, 1869–1951. Hommage de l'Académie de la République Socialiste de Roumanie à l'occasion du centenaire de sa naissance. Comprenant les communications faites aux séances du 4e Congrès du Groupement des Mathématiciens d'Expression Latine, tenu à Bucarest en 1969 (Editura Academiei Republicii Socialiste Romania, Bucharest, 1975) pp. 39–45
23. Vinberg, E.B.: Homogeneous cones. Dokl. Akad. Nauk SSSR. **133**, 9–12 (1960); Sov. Math. Dokl. **1**, 787–790 (1961)
24. Vinberg, E.B.: The Morozov-Borel theorem for real Lie groups. Dokl. Akad. Nauk SSSR. **141**, 270–273 (1961); Sov. Math. Dokl. **2**, 1416–1419 (1962)

25. Vinberg E.B., Convex homogeneous domains, Dokl. Akad. Nauk SSSR., 141 1961, 521–524; Soviet Math. Dokl., n°2, pp. 1470–1473, 1962
26. Vinberg, E.B.: Automorphisms of homogeneous convex cones. Dokl. Akad. Nauk SSSR. **143**, 265–268 (1962); Sov. Math. Dokl. **3**, 371–374 (1963)
27. Vinberg, E.B.: The Theory of Homogeneous Convex Cones, Trudy Moskovskogo Matematicheskogo Obshchestva, vol. 12, pp. 303–358 (1963); Trans. Moscow Math. Soc. **12**, 340–403 (1963)
28. Vinberg, E.B., Gindikin, S.G., Pyatetskii-Shapiro, I.I.: On the classification and canonical realization of bounded homogeneous domains. Trudy Moskov. Mat. Obshch. **12**, 359–388 (1963); Trans. Moscow Math. Soc. **12**, 404–437 (1963)
29. Vinberg, E.B.: The structure of the group of automorphisms of a convex cone. Trudy Moscov. Mat. Obshch. **13**, 56–83 (1964); Trans. Moscow Math. Soc. **13** (1964)
30. Vinberg, E.B.: Structure of the group of automorphisms of a homogeneous convex cone. Tr. Mosk. Mat. Obshchestva **13**, 56–83 (1965)
31. Pyatetskii-Shapiro, I.I.: Certain problems of harmonic analysis in homogeneous cones. Dokl. Akad. Nauk SSSR. 181–184 (1957)
32. Pyatetskii-Shapiro, I.I.: On a problem of E. Cartan Dokl. Akad. Nauk SSSR. **124**, 272–273 (1959)
33. Pyatetskii-Shapiro, I.I.: The geometry of homogeneous domains and the theory of automorphic functions, the solution of a problem of E. Cartan. Uspekhi Mat. Nauk. **14**(3), 190–192 (1959)
34. Pyatetskii-Shapiro, I.I.: On the classification of bounded homogeneous domains in n-dimensional complex space. Dokl. Akad. Nauk SSSR. **141**, 316–319 (1961); Sov. Math. Dokl. **141**, 1460–1463 (1962)
35. Pyatetskii-Shapiro, I.I.: On bounded homogeneous domains in n-dimensional complex space. Izv. Akad. Nauk SSSR. Ser. Mat. **26**, 107–124 (1962)
36. Gindikin, S.G.: Analysis in homogeneous domains. Uspekhi Mat. Nauk. **19**(4), 3–92 (1964); Russ. Math. Surv. **19**(4), 1–89 (1964)
37. Cartan, E.: Sur les domaines bornés de l'espace de n variables complexes. Abh. Math. Semin. Hambg. **1**, 116–162 (1935)
38. Berger, M.: Les espaces symétriques noncompacts. Ann. Sci. l'École Norm. Supérieure Sér. 3 Tome **74**(2), 85–177 (1957)
39. Vey, J.: Sur une notion d'hyperbolicité des variétés localement plates, Thèse de troisième cycle de mathématiques pures, Faculté des sciences de l'université de Grenoble (1969)
40. Vey, J.: Sur les automorphismes affines des ouverts convexes saillants, Annali della Scuola Normale Superiore di Pisa, Classe di Science, 3e série, tome **24**(4), 641–665 (1970)
41. Koszul, J.L.: Variétés localement plates et convexité. Osaka. J. Math. **2**, 285–290 (1965)
42. Alekseevsky, D.: Vinberg's Theory of Homogeneous Convex Cones: Developments and Applications, Transformation Groups 2017. Conference Dedicated to Prof. Ernest B. Vinberg on the occasion of his 80th birthday, Moscou, December (2017). https://www.mccme.ru/tg2017/slides/alexeevsky.pdf, http://www.mathnet.ru/present19121
43. Lichnerowicz, A.: Espaces homogènes Kähleriens. In: Colloque de Géométrie Différentielle. Publication du CNRSP, Paris, France, pp. 171–184 (1953)
44. Siegel, C.L.: Symplectic geometry. Am. J. Math. **65**, 1–86 (1943)
45. Koszul, J.L.: Domaines bornées homogènes et orbites de groupes de transformations affines. Bull. Soc. Math. France **89**, 515–533 (1961)
46. Koszul, J.L.: Ouverts convexes homogènes des espaces affines. Math. Z. **79**, 254–259 (1962)
47. Koszul, J.L.: Lectures on Groups of Transformations. Tata Institute of Fundamental Research, Bombay (1965)
48. Siegel, C.L.: Über der analytische Theorie der quadratischen Formen. Ann. Math. **36**, 527–606 (1935)
49. Koszul, J.L.: Déformations des variétés localement plates. Ann Inst Fourier **18**, 103–114 (1968)

50. Koszul, J.L.: Trajectoires convexes de groupes affines unimodulaires. In: Essays on Topology and Related Topics, pp. 105–110. Springer, Berlin (1970)
51. Shima, H.: Symmetric spaces with invariant locally Hessian structures. J. Math. Soc. Jpn. 581–589 (1977)
52. Shima, H.: Homogeneous Hessian manifolds. Ann. Inst. Fourier 91–128 (1980)
53. Shima, H.: Vanishing theorems for compact Hessian manifolds. Ann. Inst. Fourier 183–205 (1986)
54. Shima, H.: Harmonicity of gradient mappings of level surfaces in a real affine space. Geom. Dedicata 177–184 (1995)
55. Shima, H.: Hessian manifolds of constant Hessian sectional curvature. J. Math. Soc. Jpn. 735–753 (1995)
56. Shima, H.: Homogeneous spaces with invariant projectively flat affine connections. Trans. Am. Math. Soc. 4713–4726 (1999)
57. Shima, H.: The Geometry of Hessian Structures. World Scientific, Singapore (2007)
58. Shima, H.: In: Nielsen, F., Frederic, B. (eds.) Geometry of Hessian Structures. Springer Lecture Notes in Computer Science, vol. 8085, pp. 37–55 (2013) (planches: https://www.see.asso.fr/file/5104/download/25050), (vidéos GSI'13: https://www.youtube.com/watch?time_continue=139&v=6pyXxdIzDNQ, https://www.youtube.com/watch?time_continue=182&v=jG2tUjniOUs, https://www.youtube.com/watch?time_continue=6&v=I5kdMJvuNHA)
59. Boyom, M.N.: Sur les structures affines homotopes à zéro des groupes de Lie. J. Differ. Geom. **31**, 859–911 (1990)
60. Boyom, M.N.: Structures localement plates dans certaines variétés symplectiques. Math. Scand. **76**, 61–84 (1995)
61. Boyom, M.N.: Métriques kählériennes affinement plates de certaines variétés symplectiques. I Proc. Lond. Math. Soc. 3 **66**(2), 358–380 (1993)
62. Boyom, M.N.: The cohomology of Koszul-Vinberg algebras. Pac. J. Math. **225**, 119–153 (2006)
63. Boyom, M.N.: Some Lagrangian Invariants of Symplectic Manifolds, Geometry and Topology of Manifolds; Banach Center Institute of Mathematics, Polish Academy of Sciences, Warsaw, vol. 76, pp. 515–525 (2007)
64. Boyom, M.N., Byande, P.M.: KV Cohomology in Information Geometry Matrix Information Geometry, pp. 69–92. Springer, Heidelberg, Germany (2013)
65. Boyom, M.N.: Transversally Hessian foliations and information geometry I. Am. Inst. Phys. Proc. **1641**, 82–89 (2014)
66. Boyom, M.N., Wolak, R.: Transverse Hessian metrics information geometry MaxEnt 2014. In: AIP. Conference Proceedings American Institute of Physics (2015)
67. Byande, P.M., Ngakeu, F., Boyom, M.N., Wolak, R.: KV-cohomology and differential geometry of affinely flat manifolds. Information geometry. Afr. Diaspora J. Math. **14**, 197–226 (2012)
68. Byande, P.M.: Des Structures Affines à la Géométrie de L'information; European University Editions (2012)
69. Lichnerowicz, A.: Groupes de Lie à structures symplectiques ou Kähleriennes invariantes. In: Albert, C. (ed.) Géométrie Symplectique et Mécanique. Lecture Notes in Mathematics, vol 1416. Springer, Berlin (1990)
70. Cartan, H.: Allocution de Monsieur Henri Cartan, colloques Jean-Louis Koszul. Ann. l'Institut Fourier **37**(4), 1–4 (1987)
71. Malgrange, B.: Quelques souvenirs de Jean-Louis KOSZUL. Gazette des Mathématiciens **156**, 63–64 (2018)
72. Koszul, J.L.: Interview for "Institut Joseph Fourier" 50th birthday in 2016. https://www.youtube.com/watch?v=AzK5K7Q05sw
73. Barbaresco, F.: Jean-Louis Koszul et les structures élémentaires de la géométrie de l'information, revue MATAPLI 2018, SMAI (2018)
74. Koszul, J.M.: Homologie et cohomologie des algèbres de Lie. Bull. Soc. Math. Fr. Tome **78**, 65–127 (1950)

75. Cartan, H.: Les travaux de Koszul, I, Séminaire Bourbaki, Tome 1, Exposé **1**, 7–12 (1948–1951)
76. Cartan, H.: Les travaux de Koszul, II, Séminaire Bourbaki, Tome 1, Exposé **8**, 45–52 (1948–1951)
77. Cartan, H.: Les travaux de Koszul, III, Séminaire Bourbaki, Tome 1, Exposé **12**, 71–74 (1948–1951)
78. Haefliger, A.: Des espaces homogènes à la résolution de Koszul. Ann. l'inst. Fourier Tome **37**(4), 5–13 (1987)
79. Souriau, J.-M.: Mécanique statistique, groupes de Lie et cosmologie, Colloques int. du CNRS numéro 237, Géométrie symplectique et physique mathématique, pp. 59–113 (1974)
80. Barbaresco F., Koszul information geometry and Souriau Lie group thermodynamics. AIP Conference Proceedings 1641, 74, 2015, Proceedings of MaxEnt'14 Conference, Amboise, Septembre 2014
81. Barbaresco, F.: Koszul Information Geometry and Souriau Geometric Temperature/Capacity of Lie Group Thermodynamics, Entropy, vol. 16, pp. 4521–4565 (2014); Published in the book Information, Entropy and Their Geometric Structures, MDPI Publisher, September 2015
82. 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, pp. 529–540. Springer, Berlin (2015)
83. Barbaresco, F.: Geometric theory of heat from souriau lie groups thermodynamics and koszul hessian geometry: applications in information geometry for exponential families. Entropy **18**, 386 (2016)
84. Barbaresco, F.: Poly-symplectic Model of Higher Order Souriau Lie Groups Thermodynamics for Small Data Analytics, GSI 2017. Springer LNCS, vol. 10589, pp. 432–441 (2017)
85. Cartan, E.: Sur les invariants intégraux de certains espaces homogènes clos et les propriétés topologiques de ces espaces. Ann. Soc. Pol. De Math. **8**, 181–225 (1929)
86. Hua, L.K.: Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains. American Mathematical Society, Providence, USA (1963)
87. Berezin, F.: Quantization in complex symmetric spaces. Izv. Akad. Nauk SSSR Ser. Math. **9**, 363–402 (1975)
88. Maliavin, P.: Invariant or quasi-invariant probability measures for infinite dimensional groups, Part II: unitarizing measures or Berezinian measures. Jpn. J. Math. **3**, 19–47 (2008)
89. Rothaus, O.S.: The Construction of Homogeneous Convex Cones. Annals of Mathematics, Series 2, vol. 83, pp. 358–376 (1966)
90. Vesentini, E.: Geometry of Homogeneous Bounded Domains. Springer, Berlin (2011); reprint of the 1st edn. C.IM.E., Ed. Cremonese, Roma 1968
91. Sampieri, U.: A generalized exponential map for an affinely homogeneous cone. Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti Lincei. Matematica e Applicazioni, Serie 8. **75(**6), 320–330 (1983)
92. Sampieri, U.: Lie group structures and reproducing kernels on homogeneous siegel domains. Ann. Mat. **152**(1), 1–19 (1988)
93. Nesterov, Y., Nemirovskii, A.: Interior-Point Polynomial Algorithms in Convex Programming. SIAM Series: Studies in Applied and Numerical Mathematics, pp. ix + 396 (1994)
94. Gromov, M.: Convex sets and kähler manifolds. In: Tricerri, F. (ed.) Advances in Differential Geometry and Topology, pp. 1–38. World Scientific, Singapore (1990)
95. Gromov, M.: In a Search for a Structure, Part 1: On Entropy. http://www.ihes.fr/~gromov/PDF/structre-serch-entropy-july5-2012.pdf (2012). Accessed 23 June 2012
96. Gromov, M.: Gromov Six Lectures on Probability, Symmetry, Linearity. October 2014, Jussieu, November 6th, 2014; Lecture Slides & video of Gromov lectures on youtube: http://www.ihes.fr/~gromov/PDF/probability-huge-Lecture-Nov-2014.pdf, https://www.youtube.com/watch?v=hb4D8yMdov4
97. Gromov, M.: Gromov Four Lectures on Mathematical Structures Arising from Genetics and Molecular Biology, IHES, October 2013; video of Lectures on youtube: https://www.youtube.com/watch?v=v7QuYuoyLQc&t=5935s

98. Barbaresco, F.: Les densités de probabilité « distinguées » et l'équation d'Alexis Clairaut: regards croisés de Maurice Fréchet et de Jean-Louis Koszul, GRETSI'17, Juan-Les-Pins (2017)
99. Massieu, F.: Sur les Fonctions caractéristiques des divers fluides. C. R. Acad. Sci. **69**, 858–862 (1869)
100. Massieu, F.: Addition au précédent Mémoire sur les Fonctions caractéristiques. C. R. Acad. Sci. **69**, 1057–1061 (1869)
101. 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); Académie des Sciences, Paris, 31 p. (1873)
102. Massieu, F.: Thermodynamique: Mémoire sur les Fonctions Caractéristiques des Divers Fluides et sur la Théorie des Vapeurs. Académie des Sciences, Paris, France, p. 92 (1876)
103. Poincaré, H.: Calcul des Probabilités. Gauthier-Villars, Paris, France (1896)
104. Balian, R., Balazs, N.: Equiprobability, inference and entropy in quantum theory. Ann. Phys. **179**, 97–144 (1987)
105. Balian, R.: On the principles of quantum mechanics. Am. J. Phys. **57**, 1019–1027 (1989)
106. Balian, R.: From Microphysics to Macrophysics: Methods and Applications of Statistical Physics, Vols. I and II. Springer, Heidelberg, Germany (1991, 1992)
107. Balian, R.: Incomplete descriptions and relevant entropies. Am. J. Phys. **67**, 1078–1090 (1999)
108. Balian, R., Valentin, P.: Hamiltonian structure of thermodynamics with gauge. Eur. Phys. J. B **21**, 269–282 (2001)
109. Balian, R.: Entropy, a protean concept. In: Dalibard, J., Duplantier, B., Rivasseau, V. (eds.) Poincaré Seminar 2003, pp. 119–144. Birkhauser, Basel, Switzerland (2004)
110. Balian, R.: Information in statistical physics. In: Studies in History and Philosophy of Modern Physics, Part B. Elsevier, Amsterdam (2005)
111. Balian, R.: François Massieu et les Potentiels Thermodynamiques, Évolution des Disciplines et Histoire des Découvertes. Académie des Sciences, Avril, France (2015)
112. Leray, J.: Un prolongement de la transformation de Laplace qui transforme la solution unitaire d'un opérateur hyperbolique en sa solution élémentaire. (Problème de Cauchy. IV.). Bulletin de la Société Mathématique de France **90**, 39–156 (1962)
113. Legendre, A.M.: Mémoire Sur L'intégration de Quelques Equations aux Différences Partielles; Mémoires de l'Académie des Sciences: Paris, France, pp. 309–351 (1787)
114. Casalis, M.: Familles exponentielles naturelles invariantes par un groupe de translations. C. R. Acad. Sci. Ser. I Math. **307**, 621–623 (1988)
115. Casalis, M.: Familles exponentielles naturelles invariantes par un groupe. Ph.D. Thesis, Thèse de l'Université Paul Sabatier, Toulouse, France (1990)
116. Letac, G.: Lectures on Natural Exponential Families and Their Variance Functions, Volume 50 of Monografias de Matematica (Mathematical Monographs); Instituto de Matematica Pura e Aplicada (IMPA). Rio de Janeiro, Brazil (1992)
117. Jaynes, E.T.: Information theory and statistical mechanics. Phys. Rev. Ser. II **106**, 620–630 (1957)
118. Jaynes, E.T.: Information theory and statistical mechanics II. Phys. Rev. Ser. **108**, 171–190 (1957)
119. Jaynes, E.T.: Prior probabilities. IEEE Trans. Syst. Sci. Cybern. **4**, 227–241 (1968)
120. Moreau, J.J.: Fonctions convexes duales et points proximaux dans un espace hilbertien. C. R. Acad. Sci. **255**, 2897–2899 (1962)
121. Dacunha-Castelle, D., Gamboa, F.: Maximum d'entropie et problèmes des moments. Ann. Inst. H. Poincaré Prob. Stat. **26**, 567–596 (1990)
122. Gamboa, F., Gassiat, E.: Maximum d'entropie et problème des moments: Cas multidimensionnel. Probab. Math. Stat. **12**, 67–83 (1991)
123. Dacunha-Castelle, D., Gamboa, F.: Maximum de l'entropie sous contraintes non linéaires. Ann. Inst H. Poincaré Prob. Stat. **4**, 567–596 (1990)
124. Burbea, J., Rao, C.R.: Entropy differential metric, distance and divergence measures in probability spaces: a unified approach. J. Multivar. Anal. **12**, 575–596 (1982)

125. Crouzeix, J.P.: A relationship between the second derivatives of a convex function and of its conjugate. Math. Progr. **3**, 364–365 (1977)
126. Hiriart-Urruty, J.B.: A new set-valued second-order derivative for convex functions. In: Mathematics for Optimization; Mathematical Studies Series, vol. 129. Elsevier: Amsterdam, The Netherlands (1986)
127. Eriksen, P.: Geodesics Connected with the Fisher Metric on the Multivariate Normal Manifold; Technical Report, pp. 86–13; Institute of Electronic Systems, Aalborg University, Aalborg, Denmark (1986)
128. Eriksen, P.S.: Geodesics connected with the Fisher metric on the multivariate normal manifold. In: Proceedings of the GST Workshop, Lancaster, UK (1987). Accessed 28–31 Oct 1987
129. Koszul, J.L.: Formes hermitiennes canoniques des espaces homogènes complexes. Sémin. Bourbaki Tome 3, Exposé **108**, 69–75 (1954–1956)
130. Bourguignon, J.P.: La géométrie kählérienne, domaines de la géométrie différentielle, séminaire Histoires de géométries, FMSH (2005). https://youtu.be/SDmMo4a1vbk
131. Gauduchon, P.: Calabi's extremal Kähler metrics: an elementary introduction. https://germanio.math.unifi.it/wp-content/uploads/2015/03/dercalabi.pdf
132. Cartier, P.: In memoriam Jean-Louis KOSZUL. Gazette des Mathématiciens **156**, 64–66 (2018)
133. Kosmanek, E.-E.: Hommage à mon directeur de thèse, Gazette des Mathématiciens **156**, 67, (2018)