Entropy **2014**, *16*, 4521-4565; doi:10.3390/e16084521

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

Article

Koszul Information Geometry and Souriau Geometric Temperature/Capacity of Lie Group Thermodynamics

Frédéric Barbaresco

Thales Air Systems, Advanced Radar Concepts Business Unit, Voie Pierre-Gilles de Gennes, Limours F-91470, France; E-Mail: frederic.barbaresco@thalesgroup.com; Tel.: +33-(0)630071419

Received: 30 March 2014; in revised form: 11 June 2014 / Accepted: 23 June 2014/ Published: 12 August 2014

Abstract: The François Massieu 1869 idea to derive some mechanical and thermal properties of physical systems from "Characteristic Functions", was developed by Gibbs and Duhem in thermodynamics with the concept of potentials, and introduced by Poincaré in probability. This paper deals with generalization of this Characteristic Function concept by Jean-Louis Koszul in Mathematics and by Jean-Marie Souriau in Statistical Physics. The Koszul-Vinberg Characteristic Function (KVCF) on convex cones will be presented as cornerstone of "Information Geometry" theory, defining Koszul Entropy as Legendre transform of minus the logarithm of KVCF, and Fisher Information Metrics as hessian of these dual functions, invariant by their automorphisms. In parallel, Souriau has extended the Characteristic Function in Statistical Physics looking for other kinds of invariances through co-adjoint action of a group on its momentum space, defining physical observables like energy, heat and momentum as pure geometrical objects. In covariant Souriau model, Gibbs equilibriums states are indexed by a geometric parameter, the Geometric (Planck) Temperature, with values in the Lie algebra of the dynamical Galileo/Poincaré groups, interpreted as a space-time vector, giving to the metric tensor a null Lie derivative. Fisher Information metric appears as the opposite of the derivative of Mean "Moment map" by geometric temperature, equivalent to a Geometric Capacity or Specific Heat. We will synthetize the analogies between both Koszul and Souriau models, and will reduce their definitions to the exclusive Cartan "Inner Product". Interpreting Legendre transform as Fourier transform in $(Min,+)$ algebra, we conclude with a definition of Entropy given by a relation mixing Fourier/Laplace transforms: $Entropy = (minus)\ Fourier_{(Min,+)}\ o\ Log\ o\ Laplace_{(+,X)}$.

Keywords: Koszul-Vinberg characteristic function; Koszul forms; Koszul entropy; temperature vector; covariant thermodynamics; Souriau-Gibbs equilibrium state

PACS Codes: 02 (Mathematical methods in physics), 05 (Statistical physics, thermodynamics, and nonlinear dynamical systems)

1. Introduction

The Koszul-Vinberg Characteristic Function (KVCF) is a dense knot in important mathematical fields such as Hessian Geometry, Kählerian Geometry and Affine Differential Geometry. As essence of Information Geometry, this paper develops KVCF as a transverse concept in Thermodynamics, in Statistical Physics and in Probability. From general KVCF definition, the paper introduces Koszul Entropy as the Legendre transform of minus the logarithm of KVCF, and compares both functions by analogy with the Dual Massieu-Duhem potentials in thermodynamics. This paper will also explore close inter-relations between these domains through geometric tools developed by Jean-Louis Koszul and Jean-Marie Souriau. The cornerstone of "Information Geometry" Theory will appear to be based on the fundamental property that derivatives of the Koszul-Vinberg Characteristic Function Logarithm (KVCFL) $\log \psi_\Omega(x) = \log \int_{\Omega^*} e^{-\langle \xi, x \rangle} d\xi$, defined on convex dual cone of Ω , are invariant by the automorphisms of Ω , and that its hessian defines a non-arbitrary Riemannian metric.

In thermodynamics, François Massieu [1–3] was the first to introduce the concept of characteristic function ϕ . This characteristic function or thermodynamic potential is able to provide all the body properties from their derivatives. In thermodynamics, Entropy S is one of the Massieu-Duhem potentials [4–8], derived from the Legendre-Moreau transform of the characteristic function logarithm ϕ : $S = \phi - \beta . \frac{\partial \phi}{\partial \beta}$ with $\beta = \frac{1}{kT}$ being the thermodynamic temperature. The most popular notion of "*characteristic function*" was introduced in a second step by Henri Poincaré in his lecture on probability [9,10], using the property that all moments of statistical laws could be deduced from its derivatives. Paul Levy then made systematic use of this concept in his 1925 book. We assume that Poincaré was influenced by his school fellow at Ecole des Mines de Paris, François Massieu, and his work on thermodynamic potentials (generalized by Pierre Duhem in an Energetic Theory). This assertion is corroborated by the observation that Poincaré added in his lecture on thermodynamics in the 2nd edition [9,10] in 1892, a chapter on the "Massieu characteristic function" with many developments and applications, before developing the concept in Probability [9,10], see Figure 1.

In Thermodynamics, Statistical Physics and Probability, we can observe that the "*characteristic function*" and its derivatives capture all information of system or physical model and random variable. Furthermore, the general notion of Entropy could be naturally defined by the Legendre Transform of minus the Koszul characteristic function logarithm. In the general case, Legendre transform of minus the logarithm of the KVCF will be designated in the following as "*Koszul Entropy*".

Figure 1. Text of Poincaré Lecture on Thermodynamic with development of the concept of "Massieu Characteristic Function".

M. Massieu a montré que, si l'on fait choix pour variables indépendantes de v et de T ou de p et de T, il existe une fonction, d'ailleurs inconnue, de laquelle les trois fonctions des variables, p, U et S dans le premier cas, v, U et S dans le second, peuvent se déduire facilement. M. Massieu à donné à cette fonction, dont la forme dépend du choix des variables, le nom de *fonction caractéristique.*

[M. Massieu showed that, if we make choice for independent variables of v and T or of p and T, there is a function, moreover unknown, of which three functions of variables, p, U and S in the first case, v, U and S in the second, can be deducted easily. M. Massieu gave to this function, the form of which depends on the choice of variables, name of characteristic function.]

Puisque des fonctions de M. Massieu on peut déduire les autres fonctions des variables, toutes les équations de la Thermodynamique pourront s'écrire de manière à ne plus renfermer que ces fonctions et leurs dérivées; il en résultera donc, dans certains cas, une grande simplification. Nous verrons bientôt une application importante de ces fonctions.

[Because functions of M. Massieu, we can deduct the other functions of variables, all the equations of the Thermodynamics can be written not so as to contain more than these functions and their derivatives; it will thus result from it, in certain cases, a large simplification. We shall see soon an important application of these functions.]

This general notion of "*characteristic function*" has been generalized by the French physicist Jean-Marie Souriau. In 1970, Souriau, that had followed the Elie Cartan Lecture at ENS Ulm in 1946 (one year after his aggregation), introduced the concept of co-adjoint action of a group on its momentum space (or "*moment map*": mapping induced by symplectic manifold symmetries), based on the orbit method works, that allows to define physical observables like energy, heat and momentum as pure geometrical objects (the moment map takes its values in a space attached to the group of symmetries in the dual space of its Lie algebra). The moment map is a constant of the motion and is associated to symplectic cohomology (assignment of algebraic invariants to a topological space that arises from the algebraic dualization of the homology construction). For Souriau, equilibrium states are indexed by a geometric parameter β with values in the Lie algebra of the dynamical group (Galileo or Poincaré group). The Souriau approach generalizes the Gibbs equilibrium states, β playing the role of temperature. The invariance with respect to the group, and the fact that the entropy S is a convex function of β, imposes very strict conditions, that allow Souriau to interpret β as a space-time vector (the vector-valued temperature of Planck), giving to the metric tensor a null Lie derivative. For Souriau, all the details of classical mechanics appear as geometric necessities (e.g., mass is the measure of the symplectic cohomology of the action of a Galileo group). We will synthetize the analogies between the Koszul and Souriau models in tables (the Information Geometry case being a particular case of Koszul Hessian geometry).

The Koszul-Vinberg characteristic function is a dense knot in mathematics and could be introduced in the framework of different geometries: Hessian Geometry (Jean-Louis Koszul's work), Homogeneous convex cones geometry (Ernest Vinberg's work [11]), Homogeneous Symmetric Bounded Domains Geometry [12,13] (Elie Cartan [14] and Carl Ludwig Siegel's works [15,16]), Symplectic Geometry [17,18] (Thomas von Friedrich [19] & Jean-Marie Souriau's work), Affine Geometry (Takeshi Sasaki and Eugenio Calabi's works) and Information Geometry (Calyampudi Rao and Nikolai Chentsov' works). Through Legendre duality, Contact Geometry (Vladimir Arnold's work) is considered as the odd-dimensional twin of symplectic geometry and could be used to understand Legendre mapping in Information Geometry. Fisher metrics of Information Geometry could be introduced as hessian metrics from minus Koszul-Vinberg characteristic function logarithm or from Koszul Entropy (Legendre transform of minus Koszul-Vinberg characteristic function logarithm). In a more general context, we can consider Information Geometry in the framework of "*Geometric Science of Information*", a new "corpus" that also covers probability in metric space (Maurice Fréchet's work), probability/geometry on structures (Yann Ollivier and Misha Gromov's works [20–23]) and probability on Riemannian manifold (Michel Emery and Marc Arnaudon's works). This link between "*Information Theory*" and "Geometry" is also deeply developed and influenced by fundamental works of Yann Ollivier [24,25] (initially described in his HDR report "*Randomness and Curvature*" in 2009 and more recent papers on IGO flow).

2. Legendre Duality and Projective Duality

In following chapters, we will see that the minus Logarithm of the Characteristic Function and Entropy will be related by the Legendre transform, that can be considered in the context of projective duality. Duality is an old and very fruitful idea in mathematics that has been constantly generalized [26–38]. A duality translates concepts, theorems or mathematical structures into other concepts, theorems or structures, in a one-to-one fashion, often by means of an involution operation and sometimes with fixed points.

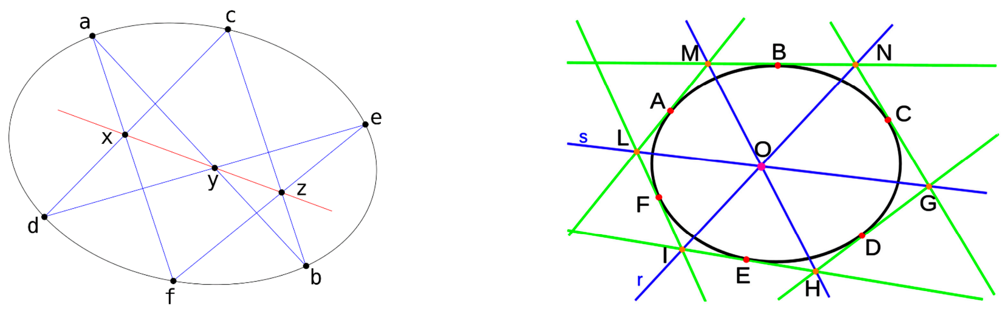
The simplest duality is linear duality in the plane with points and lines (two different points can be joined by a unique line. Two different lines meet in one point unless they are parallel). By adding some points at infinity (to avoid particular case of parallel lines) then we obtain the projective plane in which the duality is given symmetrical relationship between points and lines, and led to the classical principle of projective duality, where the dual theorem is also a theorem.

Most Famous example is given by *Pascal's theorem* (the Hexagrammum Mysticum Theorem) stating that:

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

The dual of Pascal's Theorem is known as *Brianchon's Theorem*, as illustrated in Figure 2:

Figure 2. (**a**) Pascal's theorem, (**b**) Brianchon's theorem.



(**a**) (**b**)

- If the sides of a simple hexagon are lines of a line conic, then the diagonal lines are concurrent.

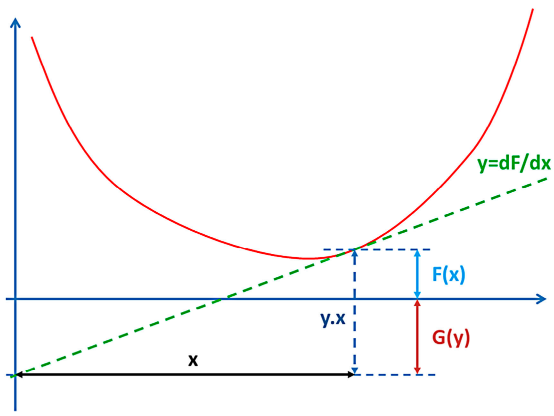
The Legendre(-Moreau) transform [39,40] is an operation from convex functions on a vector space to functions on the dual space. The Legendre transform is related to projective duality and tangential coordinates in algebraic geometry, and to the construction of dual Banach spaces in analysis. Classical Legendre transform in Euclidean space is given by fixing a scalar product $\langle .,. \rangle$ on R^n. For a function $F : R^n \to R \cup \{\pm\infty\}$, let:

$$G(y) = LF(y) = \underset{x}{Sup}\left\{\langle y, x \rangle - F(x)\right\} \tag{1}$$

The Legendre transform is illustrated in Figure 3.

This is an involution on the class of convex lower semi-continuous functions on R^n. There are two dual possibilities to describe a function. We can either use a function, or we may regard the curve as the envelope of its tangent planes. We give in Appendix A1 the historical context of Legendre Transform introduction on a Minimal Surface problem considered initially by Gaspard Monge.

Figure 3. Legendre Transform *G(y)* of *F(x)*.



The Legendre Transform is very important in Information Geometry [39], which uses mutually dual (conjugate) affine connections, dual potentials in dual coordinates systems and dual metrics that are studied in the framework of Hessian or affine differential geometry.

To illustrate the role of Legendre transform in Information Geometry, we provide a canonical example, with the relations for the Multivariate Normal Gaussian Law $N(m, R)$:

- Dual Coordinates systems:

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

- Dual potential functions:

$$\begin{cases} \tilde{\Psi}\left(\tilde{\Theta}\right) = 2^{-1}Tr\left(\Theta^{-1}\theta\theta^T\right) - 2^{-1}\log\left(\det\Theta\right) + 2^{-1}n\log(2\pi e) \\ \tilde{\Phi}\left(\tilde{H}\right) = -2^{-1}\log\left(1 + \eta^T H^{-1}\eta\right) - 2^{-1}\log\left(\det(-H)\right) - 2^{-1}n\log\left(2\pi e\right) \end{cases} \tag{3}$$

related by Legendre transform:

$$\tilde{\Phi}\left(\tilde{H}\right)=\left\langle \tilde{\Theta},\tilde{H}\right\rangle -\tilde{\Psi}\left(\tilde{\Theta}\right) \text{ with } \left\langle \tilde{\Theta},\tilde{H}\right\rangle =Tr\left(\theta\eta^{T}+\Theta H^{T}\right) \tag{4}$$

where dual coordinate systems are given by derivatives of dual potential functions:

$$\begin{cases} \dfrac{\partial\tilde{\Psi}}{\partial\theta}=\eta \\ \dfrac{\partial\tilde{\Psi}}{\partial\Theta}=H \end{cases} \quad and \quad \begin{cases} \dfrac{\partial\tilde{\Phi}}{\partial\eta}=\theta \\ \dfrac{\partial\tilde{\Phi}}{\partial H}=\Theta \end{cases} \tag{5}$$

with $\tilde{\Phi}\left(\tilde{H}\right)=E[\log p]$ being the Entropy.

In the theory of Information Geometry introduced by Rao and Chentsov, a Riemannian manifold is then defined by a metric tensor given by hessian of these dual potential functions:

$$g_{ij}=\frac{\partial^2\tilde{\Psi}}{\partial\tilde{\Theta}_i\partial\tilde{\Theta}_j} \text{ and } g_{ij}^{*}=\frac{\partial^2\tilde{\Phi}}{\partial\tilde{H}_i\partial\tilde{H}_j} \tag{6}$$

In this paper, we will develop the concept of "*Hessian Manifolds*" theory that was initially studied by Koszul in a more general framework. In the next section, we will expose the theory of the Koszul-Vinberg characteristic function on convex sharp cones that will be presented as a general framework of Information Geometry.

3. Koszul Characteristic Function/Entropy by Legendre Duality

We define the Koszul-Vinberg Hessian metric on a convex sharp cone, and observe that the Fisher information metric of Information Geometry coincides with the canonical Koszul Hessian metric (given by Koszul forms) [41–47]. We also observe, by Legendre duality (Legendre transform of minus Koszul characteristic function logarithm), that we are able to introduce a *Koszul Entropy*, that plays the role of the generalized Shannon Entropy.

3.1. Koszul-Vinberg Characteristic Function and Metric for Convex Sharp Cone

Jean-Louis Koszul [41,42,47] and Ernest B. Vinberg [48,49] have introduced an affinely invariant hessian metric on a sharp convex cone Ω^* through its characteristic function ψ. In the following, Ω^* is a sharp open convex cone in a vector space E of finite dimension on R (a convex cone is sharp if it does not contain any full straight line). In dual space E^* of E, Ω^* is the set of linear strictly positive forms on $\bar{\Omega}-\{0\}$. Ω^* is the dual cone of Ω and is a sharp open convex cone. If $\xi\in\Omega^*$, then the intersection $\Omega\cap\left\{x\in E/\left\langle x,\xi\right\rangle =1\right\}$ is bounded. $G=Aut(\Omega)$ is the group of linear transform of E that preserves Ω. $G=Aut(\Omega)$ operates on Ω^* by $\forall g\in G=Aut(\Omega),\forall\xi\in E^*$ then $\tilde{g}.\xi=\xi\circ g^{-1}$.

Koszul-Vinberg Characteristic Function Definition:

*L*et $d\xi$ be the Lebesgue measure on E^*, the following integral:

$$\psi_\Omega(x)=\int_{\Omega^*}e^{-\left\langle \xi,x\right\rangle }d\xi \quad \forall x\in\Omega \tag{7}$$

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

The Koszul-Vinberg Characteristic Function has the following properties:

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

$$K_\Omega(x) = \int_{\Omega^*} e^{-\langle \xi, x\rangle} \psi_{\Omega^*}(\xi)^{-1} d\xi \tag{8}$$

- ψ_Ω is analytic function defined on the interior of Ω and $\psi_\Omega(x) \to +\infty$ as $x \to \partial\Omega$

If $g \in Aut(\Omega)$ then $\psi_\Omega(gx) = \left|\det g\right|^{-1} \psi_\Omega(x)$ and since $tI \in G = Aut(\Omega)$ for any $t > 0$, we have

$$\psi_\Omega(tx) = \psi_\Omega(x)/t^n \tag{9}$$

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

From the KVCF, could be introduced two forms defined by Koszul:

Koszul 1-form α : The differential 1-form

$$\alpha = d\phi_\Omega = d\log\psi_\Omega = d\psi_\Omega / \psi_\Omega \tag{10}$$

is invariant by all automorphisms $G = Aut(\Omega)$ of Ω. If and $u \in E$ then

$$\langle \alpha_x, u\rangle = -\int_{\Omega^*} \langle \xi, u\rangle . e^{-\langle \xi, x\rangle} d\xi \ \ and \ \ \alpha_x \in -\Omega^* \tag{11}$$

and:

Koszul 2-form β: The symmetric differential 2-form:

$$\beta = D\alpha = d^2 \log \psi_\Omega \tag{12}$$

is a positive definite symmetric bilinear form on E invariant under $G = Aut(\Omega)$. $D\alpha > 0$

This positivity is given by Schwarz inequality and:

$$d^2 \log\psi_\Omega(u,v) = \int_{\Omega^*} \langle \xi, u\rangle \langle \xi, v\rangle e^{-\langle \xi, u\rangle} d\xi \tag{13}$$

We can then introduce the Koszul metric based on previous definitions:

Koszul Metric: $D\alpha$ defines a Riemanian structure invariant by $Aut(\Omega)$, and then the Riemanian metric is given by $g = d^2 \log \psi_\Omega$

$$\left(d^2 \log\psi(x)\right)(u) = \frac{1}{\psi(u)^2}\left[\int_{\Omega^*} F(\xi)^2 d\xi . \int_{\Omega^*} G(\xi)^2 d\xi - \left(\int_{\Omega^*} F(\xi).G(\xi) d\xi\right)^2\right] > 0 \tag{14}$$

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

This result is obtained using Schwarz inequality, $d\log\psi = \frac{d\psi}{\psi}$ and $d^2\log\psi = \frac{d^2\psi}{\psi} - \left(\frac{d\psi}{\psi}\right)^2$ where $(d\psi(x))(u) = -\int_{\Omega^*} e^{-\langle x,\xi\rangle}\langle u,\xi\rangle d\xi$ and $(d^2\psi(x))(u) = -\int_{\Omega^*} e^{-\langle x,\xi\rangle}\langle u,\xi\rangle^2 d\xi$

A diffeomorphism is used to define dual coordinate:

$$x^* = -\alpha_x = -d\log\psi_\Omega(x) \tag{15}$$

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

$$(x^*)^* = x \ , \ \langle x, x^*\rangle = n \text{ and } \psi_\Omega(x)\psi_{\Omega^*}(x^*) = cste \tag{16}$$

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

$$x^* = \int_{\Omega^*} \xi . e^{-\langle \xi,x\rangle} d\xi \Big/ \int_{\Omega^*} e^{-\langle \xi,x\rangle} d\xi \quad \text{and} \quad \langle -x^*, h\rangle = d_h \log\psi_\Omega(x) = -\int_{\Omega^*} \langle \xi,h\rangle e^{-\langle \xi,x\rangle} d\xi \Big/ \int_{\Omega^*} e^{-\langle \xi,x\rangle} d\xi \tag{17}$$

If we set $\Phi(x) = -\log\psi_\Omega(x)$, Misha Gromov [20,21] has observed that $x^* = d\Phi(x)$ is an injection where the closure of the image equals the convex hull of the support and the volume of this hull is the the n-dimensional volume defined by the integral of the determinant of the hessian of this function $\Phi(x)$, where the map $\Phi \mapsto M(\Phi) = \int_\Omega \det(Hess(\Phi(x))).dx$ obeys non-trivial convexity relation given by the Brunn-Minkowsky inequality $[M(\Phi_1 + \Phi_2)]^{1/n} \geq [M(\Phi_1)]^{1/n} + [M(\Phi_2)]^{1/n}$.

3.2. Koszul Entropy and Its Barycenter

From this last equation, we can deduce the "*Koszul Entropy*" defined as the Legendre Transform of $\Phi(x)$ minus logarithm of Koszul-Vinberg characteristic function:

$$\begin{gathered} \Phi^*(x^*) = \langle x, x^*\rangle - \Phi(x) \text{ with } x^* = D_x\Phi \text{ and } x = D_{x^*}\Phi^* \\ \text{where } \Phi(x) = -\log\psi_\Omega(x) \end{gathered} \tag{18}$$

$$\Phi^*(x^*) = \left\langle (D_x\Phi)^{-1}(x^*), x^*\right\rangle - \Phi\left[(D_x\Phi)^{-1}(x^*)\right] \quad \forall x^* \in \{D_x\Phi(x) / x \in \Omega\} \tag{19}$$

By the definition of the Koszul-Vinberg Characteristic function, and by using $-\langle \xi, x\rangle = \log e^{-\langle \xi,x\rangle}$, we can write:

$$-\langle x^*, x\rangle = \int_{\Omega^*} \log e^{-\langle \xi,x\rangle} . e^{-\langle \xi,x\rangle} d\xi \Big/ \int_{\Omega^*} e^{-\langle \xi,x\rangle} d\xi \tag{20}$$

and:

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

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

$$\Phi^*(x^*)=\left[\log\int\limits_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi-\int\limits_{\Omega^*}\log e^{-\langle\xi,x\rangle}.\frac{e^{-\langle\xi,x\rangle}}{\int\limits_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi}d\xi\right]\tag{21}$$

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

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

In this last equation, $p_x(\xi)=e^{-\langle\xi,x\rangle}\Big/\int\limits_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi$ appears as a density, and the Legendre transform $\Phi^*()$ looks like the classical Shannon Entropy, named in the following *Koszul Entropy*:

$$\Phi^*=-\int\limits_{\Omega^*}p_x(\xi)\log p_x(\xi)d\xi\tag{22}$$

with:

$$p_x(\xi)=e^{-\langle\xi,x\rangle}\Big/\int\limits_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi=e^{-\langle x,\xi\rangle-\log\int\limits_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi}=e^{-\langle x,\xi\rangle+\Phi(x)}\text{ and }x^*=\int\limits_{\Omega^*}\xi.p_x(\xi)d\xi\tag{23}$$

We will call $p_x(\xi)=\dfrac{e^{-\langle\xi,x\rangle}}{\int\limits_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi}$ the Koszul Density, with the property that:

$$\log p_x(\xi)=-\langle x,\xi\rangle-\log\int\limits_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi=-\langle x,\xi\rangle+\Phi(x)\tag{24}$$

and:

$$E_\xi\left[-\log p_x(\xi)\right]=\left\langle x,x^*\right\rangle-\Phi(x)\tag{25}$$

We can observe that:

$$\begin{gathered}\Phi(x)=-\log\int\limits_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi=-\log\int\limits_{\Omega^*}e^{-\left[\Phi^*(\xi)+\Phi(x)\right]}d\xi=\Phi(x)-\log\int\limits_{\Omega^*}e^{-\Phi^*(\xi)}d\xi\\ \Rightarrow\int\limits_{\Omega^*}e^{-\Phi^*(\xi)}d\xi=1\end{gathered}\tag{26}$$

But the development is not achieved and we have to make appear x^* in $\Phi^*(x^*)$. For this objective, we have to write:

$$\begin{aligned}&\log p_x(\xi)=\log e^{-\langle x,\xi\rangle+\Phi(x)}=\log e^{-\Phi^*(\xi)}=-\Phi^*(\xi)\\&\Rightarrow\Phi^*=-\int_{\Omega^*}p_x(\xi)\log p_x(\xi)d\xi=\int_{\Omega^*}\Phi^*(\xi)p_x(\xi)d\xi=\Phi^*(x^*)\end{aligned}\tag{27}$$

The last equality is true if and only if we have the following relation:

$$\int_{\Omega^*}\Phi^*(\xi)p_x(\xi)d\xi=\Phi^*\left(\int_{\Omega^*}\xi.p_x(\xi)d\xi\right)\text{ as }x^*=\int_{\Omega^*}\xi.p_x(\xi)d\xi\tag{28}$$

This condition could be written more synthetically [50,51]:

$$E\left[\Phi^*(\xi)\right]=\Phi^*\left(E[\xi]\right)\ ,\ \xi\in\Omega^*\tag{29}$$

The meaning of this relation is that "*the Barycenter of Koszul Entropy is the Koszul Entropy of Barycenter*".

This condition is achieved for $x^*=D_x\Phi$ taking into account Legendre Transform property:

$$\begin{aligned}&\text{Legendre Transform: }\ \Phi^*(x^*)=\underset{x}{Sup}\left[\left\langle x,x^*\right\rangle-\Phi(x)\right]\\&\Rightarrow\begin{cases}\Phi^*(x^*)\geq\left\langle x,x^*\right\rangle-\Phi(x)\\\Phi^*(x^*)\geq\int_{\Omega^*}\Phi^*(\xi)p_x(\xi)d\xi\end{cases}\Rightarrow\begin{cases}\Phi^*(x^*)\geq E\left[\Phi^*(\xi)\right]\\\text{equality for }x^*=\dfrac{d\Phi}{dx}\end{cases}\end{aligned}\tag{30}$$

3.3. Relation of Koszul Density with the Maximum Entropy Principle

We will observe in this section that Koszul density is a solution of the Maximum Entropy. Classically, the density given by the Maximum Entropy Principle [52–58] is given by:

$$\underset{p_x(.)}{Max}\left[-\int_{\Omega^*}p_x(\xi)\log p_x(\xi)d\xi\right]\ \text{ such }\ \begin{cases}\int_{\Omega^*}p_x(\xi)d\xi=1\\\int_{\Omega^*}\xi.p_x(\xi)d\xi=x^*\end{cases}\tag{31}$$

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

$$\begin{cases}\int_{\Omega^*}q_x(\xi).d\xi=\int_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi/\int_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi=1\\\log q_x(\xi)=\log e^{-\langle x,\xi\rangle-\log\int_{\Omega^*}e^{-\langle\xi,x\rangle}d\xi}=-\langle x,\xi\rangle-\log\int_{\Omega^*}e^{-\langle x,\xi\rangle}d\xi\end{cases}\tag{32}$$

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

$$-\int_{\Omega^*}p_x(\xi)\log\frac{p_x(\xi)}{q_x(\xi)}d\xi\leq-\int_{\Omega^*}p_x(\xi)\left(1-\frac{q_x(\xi)}{p_x(\xi)}\right)d\xi\tag{33}$$

We can then observe that:

$$\int_{\Omega^*} p_x(\xi)\left(1-\frac{q_x(\xi)}{p_x(\xi)}\right)d\xi = \int_{\Omega^*} p_x(\xi)d\xi - \int_{\Omega^*} q_x(\xi)d\xi = 0 \tag{34}$$

because $\int_{\Omega^*} p_x(\xi)d\xi = \int_{\Omega^*} q_x(\xi)d\xi = 1$

We can then deduce that:

$$-\int_{\Omega^*} p_x(\xi)\log\frac{p_x(\xi)}{q_x(\xi)}d\xi \le 0 \Rightarrow -\int_{\Omega^*} p_x(\xi)\log p_x(\xi)d\xi \le -\int_{\Omega^*} p_x(\xi)\log q_x(\xi)d\xi \tag{35}$$

If we develop the last inequality, using expression of $q_x(\xi)$:

$$-\int_{\Omega^*} p_x(\xi)\log p_x(\xi)d\xi \le -\int_{\Omega^*} p_x(\xi)\left[-\langle x,\xi\rangle - \log\int_{\Omega^*} e^{-\langle x,\xi\rangle}d\xi\right]d\xi \tag{36}$$

$$-\int_{\Omega^*} p_x(\xi)\log p_x(\xi)d\xi \le \left\langle x, \int_{\Omega^*} \xi.p_x(\xi)d\xi\right\rangle + \log\int_{\Omega^*} e^{-\langle x,\xi\rangle}d\xi \tag{37}$$

If we take $x^* = \int_{\Omega^*} \xi.p_x(\xi)d\xi$ and $\Phi(x) = -\log\int_{\Omega^*} e^{-\langle x,\xi\rangle}d\xi$, then we deduce that the Koszul density $q_x(\xi) = e^{-\langle \xi,x\rangle} / \int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi = e^{-\langle x,\xi\rangle - \log\int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi}$ is the Maximum Entropy solution constrained by $\int_{\Omega^*} p_x(\xi)d\xi = 1$ and $\int_{\Omega^*} \xi.p_x(\xi)d\xi = x^*$:

$$-\int_{\Omega^*} p_x(\xi)\log p_x(\xi)d\xi \le \langle x,x^*\rangle - \Phi(x) \tag{38}$$

$$-\int_{\Omega^*} p_x(\xi)\log p_x(\xi)d\xi \le \Phi^*(x^*) \tag{39}$$

We have then observed that Koszul Entropy provides density of Maximum Entropy:

$$p_{\bar{\xi}}(\xi) = \frac{e^{-\langle \xi,\Theta^{-1}(\bar{\xi})\rangle}}{\int_{\Omega^*} e^{-\langle \xi,\Theta^{-1}(\bar{\xi})\rangle}d\xi} \text{ with } x = \Theta^{-1}(\bar{\xi}) \text{ and } \bar{\xi} = \Theta(x) = \frac{d\Phi(x)}{dx} \tag{40}$$

where:

$$\bar{\xi} = \int_{\Omega^*} \xi.p_{\bar{\xi}}(\xi)d\xi \text{ and } \Phi(x) = -\log\int_{\Omega^*} e^{-\langle x,\xi\rangle}d\xi \tag{41}$$

We can then deduce the Maximum Entropy solution without solving the classical variational problem with Lagrangian hyperparameters, but only by inversing function $\bar{\xi} = \Theta(x) = \frac{d\Phi(x)}{dx}$. This remark was made by Jean-Souriau in the paper [59]. If we take vector with tensor components $\xi = \begin{pmatrix} z \\ z\otimes z \end{pmatrix}$, components of $\bar{\xi}$ will provide moments of 1st and 2nd order of the density of probability $p_{\bar{\xi}}(\xi)$, that is defined by Gaussian law. In this particular case, we can write:

$$\langle \xi, x \rangle = a^T z + \frac{1}{2} z^T H z \tag{42}$$

with $a \in R^n$ and $H \in Sym(n)$. By the change of variables given by $z' = H^{1/2} z + H^{-1/2} a$, we can then compute the logarithm of the Koszul characteristic function:

$$\Phi(x) = -\frac{1}{2}\left[a^T H^{-1} a + \log\det\left[H^{-1}\right] + n\log(2\pi)\right] \tag{43}$$

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

$$\Phi^*(\bar{\xi}) == \frac{1}{2}\left[\log\det\left[H^{-1}\right] + n\log(2\pi.e)\right] \tag{44}$$

3.4. Crouzeix Relation on Hessian of Dual Potentials and Its Consequences

In previous sections, we have used the duality between dual potential functions that is recovered by this relation:

$$\Phi^*(x^*) + \Phi(x) = \langle x, x^* \rangle \text{ with } x^* = \frac{d\Phi}{dx} \text{ and } x = \frac{d\Phi^*}{dx^*} \text{ where } \Phi(x) = -\log\psi_\Omega(x) \tag{45}$$

If we develop relations, we can deduce that the hessian of one potential function is the inverse of the hessian of the dual potential function, then the Information Geometry metric could be given in two systems of dual coordinates:

$$\begin{gathered}\begin{cases}\dfrac{d\Phi}{dx} = x^* \\ \dfrac{d\Phi^*}{dx^*} = x\end{cases} \Rightarrow \begin{cases}\dfrac{d^2\Phi}{dx^2} = \dfrac{dx^*}{dx} \\ \dfrac{d^2\Phi^*}{dx^{*2}} = \dfrac{dx}{dx^*}\end{cases} \Rightarrow \frac{d^2\Phi}{dx^2}.\frac{d^2\Phi^*}{dx^{*2}} = 1 \Rightarrow \frac{d^2\Phi}{dx^2} = \left[\frac{d^2\Phi^*}{dx^{*2}}\right]^{-1} \\ \Rightarrow ds^2 = -\frac{d^2\Phi}{dx^2}dx^2 = -\left[\frac{d^2\Phi^*}{dx^{*2}}\right]^{-1}.\left[\frac{d^2\Phi^*}{dx^{*2}}.dx^*\right]^2 = -\frac{d^2\Phi^*}{dx^{*2}}.dx^{*2}\end{gathered} \tag{46}$$

Gromov [22] observed that the hessian of the entropy Φ^* on the space of probability measure is positive definite by the Shannon inequality and defines a (non-complete) Riemannian metric, and that this metric is called the Fisher-Rao-Kramer, Antonelli-Strobeck, Svirezhev-Shahshahani, Karquist metric.

The relation $\frac{d^2\Phi}{dx^2} = \left[\frac{d^2\Phi^*}{dx^{*2}}\right]^{-1}$ has been established first by Crouzeix in 1977 in a short communication [60] for convex smooth functions and their Legendre transforms. This result has been extended for non-smooth function by Seeger [61] and Hiriart-Urruty [62], using a polarity relationship between the second-order sub-differentials. This relation was mentioned in texts of calculus of variations and theory of elastic materials (with work potentials) [62].

This last relation has also been used in the framework of the Monge-Ampere measure associated to a convex function, to prove equality with Lebesgue measure λ:

$$m_{\Phi}(\Lambda) = \int_{\Lambda} \varphi(x)dx = \lambda\left(\left\{\nabla\phi(\mathrm{x})/\mathrm{x} \in \Lambda\right\}\right)$$
$$\forall \Lambda \in B_{\Omega} \text{ (Borel set in } \Omega\text{) and } \varphi(x) = \det\left[\nabla^2\Phi(x)\right] \tag{47}$$

That is proved using the Crouzeix relation $\nabla^2\Phi(x) = \nabla^2\Phi\left(\nabla\Phi^*(y)\right) = \left[\nabla^2\Phi^*(y)\right]^{-1}$:

$$m_{\Phi}(\Lambda) = \int_{\Lambda} \varphi(x)dx = \int_{\Lambda} \det\left[\nabla^2\Phi(x)\right]dx$$
$$m_{\Phi}(\Lambda) = \int_{\left(\nabla\Phi^*\right)^{-1}(A)} \det\left[\nabla^2\Phi\left(\nabla\Phi^*(y)\right)\right]\det\left[\nabla^2\Phi^*(y)\right]dy = \int_{\nabla\Phi(\Lambda)} 1.dy = \lambda\left(\left\{\nabla\phi(\mathrm{x})/\mathrm{x} \in \Lambda\right\}\right) \tag{48}$$

3.5. Fisher Information Geometry Metric as a Particular Case of Koszul Metric

To make the link with the classical Fisher metric given by Fisher Information matrix $I(x)$ in Information Geometry, we can observe that the second derivative of $\log p_x(\xi)$ is given by:

$$p_x(\xi) = e^{-\langle \xi,x\rangle} \Big/ \int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi = e^{-\langle x,\xi\rangle - \log \int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi} \Rightarrow \log p_x(\xi) = -\langle x,\xi\rangle - \log\int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi$$
$$\text{with } \Phi(x) = -\log\int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi = -\log\psi_{\Omega}(x) \tag{49}$$

$$\frac{\partial^2 \log p_x(\xi)}{\partial x^2} = \frac{\partial^2\Phi(x)}{\partial x^2} \tag{50}$$

$$\Rightarrow I(x) = -E_{\xi}\left[\frac{\partial^2 \log p_x(\xi)}{\partial x^2}\right] = -\frac{\partial^2\Phi(x)}{\partial x^2} = \frac{\partial^2 \log\psi_{\Omega}(x)}{\partial x^2} \tag{51}$$

We could then deduce the close interrelation between Fisher metric and hessian of minus Koszul-Vinberg characteristic logarithm, that are totally equivalent. Information Geometry then appears as a particular case of Koszul Hessian Geometry.

We can also observed that the Fisher metric or hessian of KVCF logarithm is related to the variance of ξ:

$$\log\Psi_{\Omega}(x) = \log\int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi \Rightarrow \frac{\partial \log\Psi_{\Omega}(x)}{\partial x} = -\frac{1}{\int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi}\int_{\Omega^*} \xi.e^{-\langle \xi,x\rangle}d\xi \tag{52}$$

$$\frac{\partial^2 \log\Psi_{\Omega}(x)}{\partial x^2} = -\frac{1}{\left(\int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi\right)^2}\left[-\int_{\Omega^*}\xi^2.e^{-\langle \xi,x\rangle}d\xi.\int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi + \left(\int_{\Omega^*}\xi.e^{-\langle \xi,x\rangle}d\xi\right)^2\right] \tag{53}$$

$$\frac{\partial^2 \log\Psi_{\Omega}(x)}{\partial x^2} = \int_{\Omega^*}\xi^2.\frac{e^{-\langle \xi,x\rangle}}{\int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi}d\xi - \left(\int_{\Omega^*}\xi.\frac{e^{-\langle \xi,x\rangle}}{\int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi}d\xi\right)^2 = \int_{\Omega^*}\xi^2.p_x(\xi)d\xi - \left(\int_{\Omega^*}\xi.p_x(\xi)d\xi\right)^2 \tag{54}$$

$$I(x) = -E_\xi\left[\frac{\partial^2 \log p_x(\xi)}{\partial x^2}\right] = \frac{\partial^2 \log \psi_\Omega(x)}{\partial x^2} = E_\xi\left[\xi^2\right] - E_\xi\left[\xi\right]^2 = Var(\xi) \quad (55)$$

The Inverse of the Fisher/Information Matrix $I(x)$ defines the lower bound of statistical estimators. Classically, this Lower bound is called Cramer-Rao Bound because it was described in the Rao's paper of 1945 [63]. Historically, this bound has been published first by Maurice Fréchet in 1939 in his winter "Mathematical Statistics" Lecture at the Institut Henri Poincaré during winter 1939–1940. Maurice Fréchet has published these elements in a paper as early as 1943 [64]. We can read at the bottom of the first page of his paper [64]:

> "*Le contenu de ce mémoire a formé une partie de notre cours de statistique mathématique a l'Institut Henri Poincaré pendant l'hiver 1939–1940. Il constitue l'un des chapitres du deuxième cahier (en préparation) de nos «Leçons de Statistique Mathématique», dont le premier cahier, «Introduction: Exposé préliminaire de Calcul des Probabilités" (119 pages in-quarto, dactylographiées) vient de paraitre au «Centre de Documentation Universitaire, Tournois et Constans. Paris».*"
>
> [*The contents of this report formed a part of our lecture of mathematical statistics at the Henri Poincaré institute during winter 1939–1940. It constitutes one of the chapters of the second exercise book (in preparation) of our "Lessons of Mathematical Statistics", the first exercise book of which, "Introduction: preliminary Presentation of Probability theory" (119 pages quarto, typed) has just been published in the "Centre de Documentation Universitaire, Tournois et Constans. Paris".*]

3.6. Extended Results by Koszul, Vey and Sasaki

Koszul [41,65] and Vey [66,67] have developed extended results with the following theorem for connected hessian manifolds:

Koszul-Vey Theorem*:* Let M be a connected hessian manifold with hessian metric g. Suppose that M admits a closed 1-form α such that $D\alpha = g$ and there exists a group G of affine automorphisms of M preserving α:

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

On this basis, Koszul has given a Lie Group construction of a homogeneous cone that has been developed and applied in Information Geometry by Shima [68,69] and Boyom [70] in the framework of Hessian Geometry.

After the pioneering work of Koszul, Sasaki has developed the study of hessian manifolds in Affine Geometry [71,72]. He has denoted by S_c the level surface of ψ_Ω : $S_c = \{\psi_\Omega(x) = c\}$ which is a non-compact sub-manifold in Ω, and by ω_c the induced metric of $d^2 \log \psi_\Omega$ on S_c, then assuming that the cone Ω is homogeneous under $G(\Omega)$, he proved that S_c is a homogeneous hyperbolic affine hypersphere and every such hyperspheres can be obtained in this way .Sasaki also remarks that ω_c is identified with the affine metric and S_c is a global Riemannian symmetric space when Ω is a self-dual cone. He concludes that, let Ω be a regular convex cone and let $g = d^2 \log \psi_\Omega$ be the canonical

Hessian metric, then each level surface of the characteristic function ψ_Ω is a minimal surface of the Riemannian manifold (Ω, g).

3.7. Geodesics Equation for the Koszul Hessian Metric

The last contribution has been given by Rothaus [73] who studied the construction of geodesics for this hessian metric geometry, using the following property:

$$\Gamma^i_{jk} = \frac{1}{2} g^{il}\left(\frac{\partial g_{lj}}{\partial x_k} + \frac{\partial g_{lk}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_l}\right) = \frac{1}{2} g^{il} \frac{\partial^3 \log \psi_\Omega(x)}{\partial x_j \partial x_k \partial x_l} \text{ with } g_{ij} = \frac{\partial^2 \log \psi_\Omega(x)}{\partial x_i \partial x_j} \tag{56}$$

or expressed also according the Christoffel symbol of the first kind:

$$\Gamma_{ijk} = \frac{1}{2}\left(\frac{\partial g_{jk}}{\partial x_i} + \frac{\partial g_{ki}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_k}\right) = \frac{1}{2} \frac{\partial^3 \log \psi_\Omega(x)}{\partial x_j \partial x_k \partial x_l} \tag{57}$$

Then geodesic is given by:

$$\frac{d^2 x_k}{ds^2} + \Gamma^k_{ij} \frac{dx_i}{ds}\frac{dx_j}{ds} = g_{kl} \frac{d^2 x_k}{ds^2} + \Gamma_{ijk} \frac{dx_i}{ds}\frac{dx_j}{ds} = 0 \tag{58}$$

that could be developed with previous relation:

$$\frac{d^2 x_k}{ds^2} \frac{\partial^2 \log \psi_\Omega}{\partial x_k \partial x_l} + \frac{1}{2} \frac{dx_i}{ds}\frac{dx_j}{ds} \frac{\partial^3 \log \psi_\Omega}{\partial x_l \partial x_i \partial x_j} = 0 \tag{59}$$

We can then observe that:

$$\frac{d^2}{ds^2}\left[\frac{\partial \log \psi_\Omega}{\partial x_l}\right] = \frac{dx_i}{ds}\frac{dx_j}{ds} \frac{\partial^3 \log \psi_\Omega}{\partial x_i \partial x_j \partial x_l} + \frac{d^2 x_k}{ds^2} \frac{\partial^2 \log \psi_\Omega}{\partial x_k \partial x_l} \tag{60}$$

The geodesic equation can then be rewritten:

$$\frac{d^2 x_k}{ds^2} \frac{\partial^2 \log \psi_\Omega}{\partial x_k \partial x_l} + \frac{d^2}{ds^2}\left[\frac{\partial \log \psi_\Omega}{\partial x_l}\right] = 0 \tag{61}$$

That we can put in vector form using notations $x^* = -d \log \psi_\Omega$ and Fisher matrix $I(x) = d^2 \log \psi_\Omega$:

$$I(x)\frac{d^2 x}{ds^2} - \frac{d^2 x^*}{ds^2} = 0 \text{ or } I(x) = \left[\frac{d^2 x}{ds^2}\right]^{-1} \frac{d^2 x^*}{ds^2} \tag{62}$$

3.8. Koszul Metric for Siegel Homogeneous Domains

Koszul [42] has developed his previously described theory for Homogenous Siegel Domains *SD*. He has proved that there is a subgroup *G* in the group of the complex affine automorphisms of these domains (Iwasawa subgroup), such that *G* acts on *SD* simply transitively. The Lie algebra g of *G* has a structure that is an algebraic translation of the Kähler structure of *SD*. There is an integrable almost complex structure J on, g and there exists $\eta \in g^*$ such that $\langle X,Y\rangle_\eta = \langle [JX,Y],\eta\rangle$ defines a *J-invariant positive definite inner product on* g. Koszul has proposed as admissible form $\eta \in g^*$, the form ξ:

$$\Psi(X) = \langle X,\xi\rangle = Tr[ad(JX) - J.ad(X)] \quad \forall X \in g \tag{63}$$

Koszul has proved that $\langle X,Y\rangle_\xi$ coincides, up to a positive number multiple with the real part of the Hermitian inner product obtained by the Bergman metric of *SD* by identifying g with the tangent space of *SD*. The First Koszul form is then given by:

$$\alpha = -\frac{1}{4}d\Psi(X) \tag{64}$$

We can illustrate this new Koszul expression for Poincaré's Upper Half Plane $V = \{z = x + iy \,/\, y > 0\}$ (most simple symmetric homogeneous bounded domain).

Define vector fields $X = y\dfrac{d}{dx}$ and $Y = y\dfrac{d}{dy}$, and J an almost complex structure on V defined by $JX = Y$

As:

$$[X,Y] = -Y \text{ and } ad(Y).Z = [Y,Z] \text{ then } \begin{cases} Tr[ad(JX) - Jad(X)] = 2 \\ Tr[ad(JY) - Jad(Y)] = 0 \end{cases} \tag{65}$$

The Koszul 1-form and then the Koszul/Poincaré metric is given by:

$$\Psi(X) = 2\frac{dx}{y} \Rightarrow \alpha = -\frac{1}{4}d\Psi = -\frac{1}{2}\frac{dx \wedge dy}{y^2} \Rightarrow ds^2 = \frac{dx^2 + dy^2}{2y^2} \tag{66}$$

This could be also applied for Siegel's Upper Half Space $V = \{Z = X + iY \,/\, X,Y \in Sym(p), Y > 0\}$ (more natural extension of Poincaré Upper-half plane, and general notion of symmetric bounded homogeneous domains studied by Elie Cartan and Carl-Ludwig Siegel):

$$\begin{cases} SZ = (AZ + B)D^{-1} \\ A^T D = I, B^T D = D^T B \end{cases} \text{ with } S = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \text{ and } J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{67}$$

$$\Psi(dX + idY) = \frac{3p+1}{2}Tr\left(Y^{-1}dX\right) \Rightarrow \begin{cases} \alpha = -\dfrac{1}{4}d\Psi = \dfrac{3p+1}{8}Tr\left(Y^{-1}dZ \wedge Y^{-1}d\bar{Z}\right) \\ ds^2 = \dfrac{(3p+1)}{8}Tr\left(Y^{-1}dZY^{-1}d\bar{Z}\right) \end{cases} \tag{68}$$

To recover the metric of the space of Symmetric Positive Definite (HPD) matrices, we take $Z = iR$ (with $X = 0$), and obtain the metric $ds^2 = Tr\left[\left(R^{-1}dR\right)^2\right]$. In the context of Information Geometry, this metric is the metric for multivariate Gaussian law of covariance matrix R and zero mean. For more development and application for Radar signal processing, we give reference to author papers [74–77].

4. Souriau Geometric Temperature and Covariant Definition of Thermodynamic Equilibriums

Souriau, a student of Elie Cartan [78] at ENS Ulm in 1946, has given in [59,79–87] a covariant definition of thermodynamic equilibriums and has formulated statistical mechanics [88–90] and thermodynamics in the framework of Symplectic Geometry [59] by use of symplectic moments and distribution-tensor concepts, giving a geometric status for temperature, heat and entropy. This work has been extended by Vallée and de Saxcé [91–94], Iglésias [95,96] and Dubois [97]. Other recent works address equilibrium states on manifolds of negative curvature and could be analyzed in the framework of Information Geometry [98–103].

Other directions related to polarized surface have been developed by Donaldson, Guillemin and Abreu, in which invariant Kähler metrics correspond to convex functions on the moment polytope of a toric variety [104–108] based on precursor work of Atiyah and Bott [109] on moment map and its convexity by Bruguières [110], Condevaux [111], Delzant [112], Guillemin and Sternberg [113] and Kirwan [114]. More recently, Mikhail Kapranov has also given a thermodynamical interpretation of the moment map for toric varieties [115]. Readers may consult the tutorial paper of Biquard [116].

The first general definition of the "*moment map"* (constant of the motion for dynamical systems) was introduced by Souriau during 1970s, with geometric generalization of such earlier notions as the Hamiltonian and the invariant theorem of Noether describing the connection between symmetries and invariants (it is the moment map for a one-dimensional Lie group of symmetries). In symplectic geometry the analog of Noether's theorem is the statement that the moment map of a Hamiltonian action which preserves a given time evolution is itself conserved by this time evolution. The conservation of the moment of a Hamilotnian action was called by Souriau the "*Symplectic or Geometric Noether theorem*" (considering phases space as symplectic manifold, cotangent fiber of configuration space with canonical symplectic form, if Hamiltonian has Lie algebra, moment map is constant along system integral curves. Noether theorem is obtained by considering independently each component of moment map).

In previous approach based on Koszul's work, we have defined two convex functions $\Phi(x)$ and $\Phi^*(x^*)$ with dual system of coordinates x and x^* on dual cones Ω and Ω^*:

$$\Phi(x) = -\log \int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi \quad \forall x \in \Omega \text{ and } \Phi^*(x^*) = \left\langle x, x^* \right\rangle - \Phi(x) = -\int_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi \tag{69}$$

where:

$$x^* = \int_{\Omega^*} \xi . p_x(\xi) d\xi \text{ and } p_x(\xi) = e^{-\langle \xi, x\rangle} / \int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi = e^{-\langle x, \xi\rangle - \log \int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi} = e^{-\langle x, \xi\rangle + \Phi(x)} \tag{70}$$

with

$$x^* = \frac{\partial \Phi(x)}{\partial x} \text{ and } x = \frac{\partial \Phi^*(x^*)}{\partial x^*} \tag{71}$$

Souriau introduced these relations in the framework of variational problems to extend them with a covariant definition. Let M be a differentiable manifold with a continuous positive density $d\omega$ and let E a finite vector space and $U(\xi)$ a continuous function defined on M with values in E. A continuous positive function $p(\xi)$ solution of this problem with respect to calculus of variations:

$$\underset{p(\xi)}{ArgMin}\left[s=-\int_M p(\xi)\log p(\xi)d\omega\right] \text{ such that } \begin{cases}\int_M p(\xi)d\omega=1\\ \int_M U(\xi)p(\xi)d\omega=Q\end{cases} \tag{72}$$

is given by:

$$p(\xi)=e^{\Phi(\beta)-\beta.U(\xi)} \text{ with } \Phi(\beta)=-\log\int_M e^{-\beta.U(\xi)}d\omega \text{ and } Q=\frac{\int_M U(\xi)e^{-\beta.U(\xi)}d\omega}{\int_M e^{-\beta.U(\xi)}d\omega} \tag{73}$$

Entropy $s=-\int_M p(\xi)\log p(\xi)d\omega$ can be stationary only if there exist a scalar Φ and an element β belonging to the dual of *E*, where Φ and β are Lagrange parameters associated to the previous constraints. Entropy appears naturally as Legendre transform of Φ:

$$s(Q)=\beta.Q-\Phi(\beta) \tag{74}$$

This value is a strict minimum of *s*, and the equation $Q=\frac{\int_M U(\xi)e^{-\beta.U(\xi)}d\omega}{\int_M e^{-\beta.U(\xi)}d\omega}$ has a maximum of one solution for each value of *Q*. The function $\Phi(\beta)$ is differentiable and we can write $d\Phi=d\beta.Q$ and identifying *E* with its dual:

$$Q=\frac{\partial\Phi}{\partial\beta} \tag{75}$$

Uniform convergence of $\int_M U(\xi)\otimes U(\xi)e^{-\beta.U(\xi)}d\omega$ proves that $-\frac{\partial^2\Phi}{\partial\beta^2}>0$ and that $-\Phi(\beta)$ is convex. Then, $Q(\beta)$ and $\beta(Q)$ are mutually inverse and differentiable, where $ds=\beta.dQ$.

Identifying *E* with its bidual:

$$\beta=\frac{\partial s}{\partial Q} \tag{76}$$

Classically, if we take $U(\xi)=\begin{pmatrix}\xi\\ \xi\otimes\xi\end{pmatrix}$, components of *Q* will provide moments of first and second order of the density of probability $p(\xi)$, that is defined by Gaussian law.

Souriau has applied this approach for classical statistical mechanic systems. Considering a mechanical system with *n* parameters $q_1,\cdots,q_n$, its movement could be defined by its phase at arbitrary time *t* on a manifold of dimension *2n*: $q_1,\cdots,q_n,p_1,\cdots,p_n$.

The Liouville theorem shows that coordinate changes have a Jacobian equal to unity, and a Liouville density could be defined on manifold M: $d\omega=dq_1\cdots dq_n dp_1\cdots dp_n$ that will not depend on choice to *t*.

A system state is one point on *2n*-Manifold *M* and a statistical state is a law of probability defined on M such that $\int_M p(\xi)d\omega(\xi)=1$, and its time evolution is driven by:

$$\frac{\partial p}{\partial t}=\sum\frac{\partial p}{\partial p_j}\frac{\partial H}{\partial q_j}-\frac{\partial p}{\partial q_j}\frac{\partial H}{\partial p_j} \tag{77}$$

where H is the Hamiltonian.

A thermodynamic equilibrium is a statistical state that maximizes the entropy:

$$s=-\int_M p(\xi)\log p(\xi)d\omega \tag{78}$$

among all states giving the mean value of energy Q:

$$\int_M H(\xi).p(\xi)d\omega=Q \tag{79}$$

Applying this for free particles, for an ideal gas, equilibrium is given for $\beta=\frac{1}{kT}$ (with k being the Boltzmann constant) and if we set $S=k.s$, the previous relation $dS=\frac{dQ}{T}$ provides: $S=\int\frac{dQ}{T}$ and $S=\int\frac{dQ}{T}$ and $\Phi(\beta)$ is identified with the Massieu-Duhem Potential. We recover also the Maxwell Speed law:

$$p(\xi)=cste.e^{-\frac{H}{kT}} \tag{80}$$

The main discovery of Jean-Marie Souriau is that *previous thermodynamic equilibrium is not covariant on a relativity point of view*. Then, he has proposed a covariant definition of thermodynamic equilibrium where the previous definition is a particular case. In previous formalization, manifold M was solution of the calculus of variations problem:

$$d\int_{t_0}^{t_1} l\left(t,q_j,\frac{dq_j}{dt}\right)dt=0 \text{ with } p_j=\frac{\partial l}{\partial q_j} \tag{81}$$

We can then consider the time variable *t* like other variables q_j through an arbitrary parameter τ, and define the new calculus of variations problem by:

$$d\int_{t_0}^{t_1} L(q_J,\dot{q}_J)d\tau=0 \text{ with } t=q_{n+1},\ \dot{q}_J=\frac{dq_J}{d\tau} \text{ and } J=1,2,...,n+1 \tag{82}$$

where:

$$L(q_J,\dot{q}_J)=l\left(t,q_j,\frac{\dot{q}_j}{\dot{t}}\right)\dot{t} \tag{83}$$

Variables p_j are not changed and we have the relation:

$$p_{n+1}=l-\sum_j p_j.\frac{dq_j}{dt} \tag{84}$$

If we compare with classical mechanic, we have:

$$p_{n+1} = -H \text{ with } H = \sum_j p_j.\frac{dq_j}{dt} - l \text{ (}H\text{ is Legendre transform of } l\text{)} \tag{85}$$

H is the energy of the system that is conservative if the Lagrangian doesn't depend explicitly of time *t*. It is a particular case of Noether Theorem:

If Lagrangian L is invariant by an infinitesimal transform $dQ_J = F_J(Q_K)$,then $u = \sum_J p_J dQ_J$ is first integral of variations equations.

As energy is not the conjugate variable of time *t*, or the value provided by Noether theorem by system invariance to time translation, the thermodynamic equilibrium is not covariant. Then, Souriau proposes a new covariant definition of thermodynamic equilibrium:

Let a mechanical system with a Lagrangian invariant by a Lie Group G. Equilibrium states by Group G are statistical states that maximizes the Entropy, while providing given mean values to all variables associated by Noether theorem to infinitesimal transforms of group G.

Neither theorem allows associating to all system movement ξ , a value $U(\xi)$ belonging to the vector space dual of Lie Algebra $\mathfrak{g}$ of group *G*. $U(\xi)$ is called *the moment* of the group.

For each derivation δ of this Lie algebra [83], we take:

$$U(\xi)(\delta) = \sum_J p_J.\delta Q_J \tag{86}$$

With previous development, as $\mathfrak{g}^*$ is dual of $\mathfrak{g}$, value β belongs to this Lie algebra $\mathfrak{g}$, geometric generalization of thermodynamic temperature. Value Q is a geometric generalization of heat and belongs to $\mathfrak{g}^*$, the dual of $\mathfrak{g}$.

An Equilibrium state exists having the largest entropy, with a distribution function $p(\xi)$ that is the exponential of an affine function of *U* [83]:

$$p(\xi) = e^{\Phi(\beta)-\beta.U(\xi)} \text{ with } \Phi(\beta) = -\log \int_M e^{-\beta.U(\xi)} d\omega \text{ and } Q = \frac{\int_M U(\xi) e^{-\beta.U(\xi)} d\omega}{\int_M e^{-\beta.U(\xi)} d\omega} \tag{87}$$

with:

$$s(Q) = \beta.Q - \Phi(\beta),\ d\Phi = d\beta.Q \text{ and } ds = \beta.dQ \tag{88}$$

A statistical state $p(\xi)$ is invariant by δ if $\delta[p(\xi)] = 0$ for all ξ (then $p(\xi)$ is invariant by finite transform of *G* generated by δ).

Jean-Marie Souriau gave the following theorem:

Souriau Theorem 1: An equilibrium state allowed by a group *G* is invariant by an element δ of Lie Algebra $\mathfrak{g}$, if and only if $[\delta, \beta] = 0$ (with [.], the Lie Bracket), with β the generalized equilibrium temperature.

For classical thermodynamic, where *G* is an Abelian group of translation with respect to time *t*, all equilibrium states are invariant under *G*. For Group of transformation of Space-Time, elements of Lie Algebra of *G* could be defined as vector fields in Space-Time. The generalized temperature β previously defined, would be also defined as a vector field. For each point of manifold *M*, we could then define:

- Temperature Vector:

$$\beta_M = \frac{V}{kT} \tag{89}$$

with:

- Unitary Mean Speed:

$$\text{Unitary Mean Speed: } V = \frac{\beta_M}{\|\beta_M\|} \text{ with } \|V\| = 1 \tag{90}$$

- Eigen Absolute Temperature:

$$T = \frac{1}{k.\|\beta_M\|} \tag{91}$$

Classical formula of thermodynamics are thus generalized, but variables are defined with a geometrical status, like the geometrical temperature β_M an element of the Lie algebra of the Galileo or Poincaré groups, interpreted as the field of space-time vectors. Souriau proved that in relativistic version β_M is a *time like vector* with an orientation that characterizes *the arrow of time*. The temperature vector and entropy flux are in duality. Souriau said "β, *c'est la flèche qui nous indique dans quel sens coule le temps*" [β, it is the arrow that informs about the flow of time direction].

5. Souriau-Gibbs Canonical Ensemble of Dynamical Group and Lie Group Thermodynamics

In statistical mechanics, a canonical ensemble [117–121] is the statistical ensemble that is used to represent the possible states of a mechanical system that is being maintained in thermodynamic equilibrium. Souriau has defined this Gibbs canonical ensemble on 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. The seminal idea of Lagrange was to consider that a statistical state is simply a probability measure on the manifold of motions, as in the Souriau approach, where one movement of a dynamical system (classical state) is a point on manifold of movements. For statistical mechanics, the movement variable is replaced by a random variable where a statistical state is probability law on this manifold. As symplectic manifolds have a completely continuous measure, invariant by diffeomorphisms, the Liouville measure λ, all statistical states will be the product of Liouville measure by the scalar function given by the generalized partition function $e^{\Phi-\beta.U}$ defined by the generalized energy *U* (the moment that is defined in dual of 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, $\Phi = -\log \int_M e^{-\beta.U} d\omega$. Souriau then generalizes the Gibbs equilibrium state to all symplectic manifolds that have a dynamical group. To ensure that all integrals, that will be defined, could converge, *the canonical Gibbs ensemble is the largest open proper subset (in Lie algebra) where these integrals are convergent. This canonical Gibbs ensemble is convex*. The derivative of Φ, $Q = \frac{\partial \Phi}{\partial \beta}$ is equal to the mean value of the energy U (heat in thermodynamic). The minus derivative of this generalized heat Q, $-\frac{\partial Q}{\partial \beta}$ is symmetric and positive (it is a generalization of heat capacity). Entropy s is then defined by Legendre transform of Φ, $s = \beta.Q - \Phi$. If this approach is applied for the group of time translation, this is the classical thermodynamic theory. But Souriau has observed that if we apply this theory for non-commutative group (Galileo or Poincaré groups), the symmetry has been broken. Classical Gibbs equilibrium states are no longer invariant by this group. This symmetry breaking provides new equations, discovered by Souriau.

For each temperature β, Souriau has introduced a tensor f_β, equal to the sum of cocycle f and Heat coboundary (with [.,.] Lie bracket):

$$f_\beta(Z_1, Z_2) = f(Z_1, Z_2) + Q.Ad_{Z_1}(Z_2) \quad \text{with} \quad Ad_{Z_1}(Z_2) = [Z_1, Z_2] \tag{92}$$

This tensor f_β has the following properties:

- f is a symplectic cocycle (we refer to books of Sympectic geometry for cocycle definition)
- $\beta \in Ker\ f_\beta$
- The following symmetric tensor g_β, defined on all values of $Ad_\beta(.)$ is positive definite:

$$g_\beta([\beta, Z_1], [\beta, Z_2]) = f_\beta(Z_1, [\beta, Z_2]) \tag{93}$$

These equations are universal, because they are not dependent of the symplectic manifold but only of the dynamical group G, its symplectic cocycle f, the temperature β and the heat Q. Souriau called this model "*Lie Groups Thermodynamics*". We can read in his paper this prophetical sentence "*Peut-être cette thermodynamique des groups de Lie a-t-elle un intérêt mathématique"[Maybe this thermodynamics of Lie groups has a mathematical interest]*. He explains that for dynamic Galileo group (rotation and translation) with only one axe of rotation, this thermodynamic theory is the theory of centrifuge where the temperature vector dimension is equal to 2 (sub-group of invariance of size 2), used to make "butter", "uranium 235" and "ribonucleic acid". The physical meaning of these 2 dimensions for vector-valued temperature are "thermic conduction" and "viscosity". Souriau said that the model unifies "heat conduction" and "viscosity" (Fourier and Navier equations) in the same theory of irreversible process. Souriau has applied this theory in details for relativistic ideal gas with Poincaré group for dynamical group.

We will give in the following the two others main theorems of Souriau on this Lie Group Thermodynamics.

Souriau Theorem 2. Let Ω be the largest open proper subset of $\mathfrak{g}$, Lie algebra of G, such that $\int_M e^{-\beta.U(\xi)} d\omega$ and $\int_M \xi.e^{-\beta.U(\xi)} d\omega$ are convergent integrals, this set Ω is convex and is invariant under every transformation $\overline{a}_{\mathfrak{g}}$, where $a \mapsto \overline{a}_{\mathfrak{g}}$ is the adjoint representation of G. Then, the variables are changed according to:

$$\beta \rightarrow \overline{a}_{\mathfrak{g}}(\beta) \tag{94}$$

$$\Phi \rightarrow \Phi - \theta(a^{-1})\beta = \Phi + \theta(a).\overline{a}_{\mathfrak{g}}(\beta) \tag{95}$$

$$s \rightarrow s \tag{96}$$

$$Q \to \overline{a}_{\mathfrak{g}^*}(Q) + \theta(a) = \overline{a}_{\mathfrak{g}^*_\theta}(Q) \tag{97}$$

$$\varsigma \to \overline{a}_M^+(\varsigma) \tag{98}$$

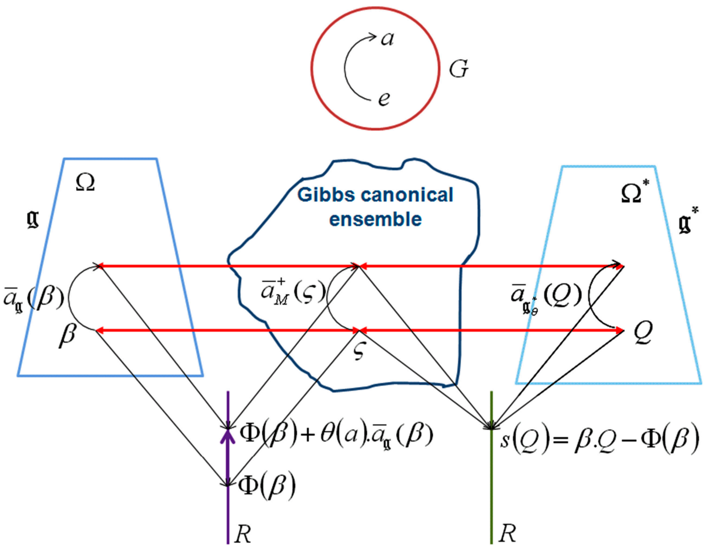
where θ is the cocycle associated with the group G and the moment, and $\overline{a}_M^+(\varsigma)$ is the image under $\overline{a}_M$ of the probability measure ς.

We observe that the entropy s is unchanged, and Φ is changed but with linear dependence to β, with consequence that Fisher Information Geometry metric is unchanged by the dynamical group:

$$I(\overline{a}_{\mathfrak{g}}(\beta)) = -\frac{\partial^2\left[\Phi - \theta\left(a^{-1}\right)\beta\right]}{\partial\beta^2} = -\frac{\partial^2\Phi}{\partial\beta^2} = I(\beta) \tag{99}$$

These transformations have been geometrically interpreted by Souriau in Figure 4:

Figure 4. Souriau figure on Lie Groups Thermodynamics.



In previous notation, $a \mapsto \overline{a}_{\mathfrak{g}}$ the adjoint representation of G can be written:

$$\overline{a}_{\mathfrak{g}}(Z) = \delta\left[a \times b \times a^{-1}\right] \quad \text{with} \quad b = e \ , \ \delta b = Z \quad \text{and} \quad \delta a = 0 \tag{100}$$

$a \mapsto \overline{a}_{\mathfrak{g}}$ defines an action of G on its Lie algebra $\mathfrak{g}$, with $\overline{a}_{\mathfrak{g}}$ is called the adjoint representation, that is a linear representation of G on its Lie algebra $\mathfrak{g}$.

Let a be an arbitrary element of G and $\overline{a}_M$ action of a on the manifold M. Since $\overline{a}_M^{-1}$ is a symplectomorphism, the image under $\overline{a}_M^{-1}$ of the Liouville measure λ is equal to λ. The integral $\int_M e^{-\beta.U(\xi)}.d\omega$ is equal with invariance property of Liouville measure to the integral $\int_M e^{-\beta.U\left(\overline{a}_M^{-1}(\xi)\right)}.d\omega$:

$$\int_M e^{-\beta.U(\xi)}.d\omega = \int_M e^{-\beta.U\left(\overline{a}_M^{-1}(\xi)\right)}.d\omega \tag{101}$$

We can then use the following relation:

$$U\left(\overline{a}_M^{-1}(\xi)\right) = \overline{a}_{\mathfrak{g}^*}^{-1}(U(\xi)) + \theta\left(a^{-1}\right) \tag{102}$$

with θ a symplectic cocycle of *G*. This cocycle is defined for:

$$\begin{gathered} U : M \to \mathfrak{g}^* \\ \xi \mapsto \mu \end{gathered} \tag{103}$$

there exist then a differential map θ defined by:

$$\begin{gathered} \theta : G \to \mathfrak{g}^* \\ a \mapsto U\left(\bar{a}_M(\xi)\right) - \bar{a}_{\mathfrak{g}^*}\left(U(\xi)\right) \end{gathered} \tag{104}$$

This differential map θ satisfy the condition:

$$\theta(a \times b) = \theta(a) + \bar{a}_{\mathfrak{g}^*}\left(\theta(b)\right) \tag{105}$$

and its derivative $f = D(\theta)(e)$ where e is the identity element of *G*, is a 2-form on the Lie algebra $\mathfrak{g}$ of *G* which satisfies:

$$f\left(Z_1,[Z_2,Z_3]\right) + f\left(Z_2,[Z_3,Z_1]\right) + f\left(Z_3,[Z_1,Z_2]\right) = 0 \quad , \quad \forall Z_1,Z_2,Z_3 \in \mathfrak{g} \tag{106}$$

and the following identities:

$$D(U)(\xi, Z_M(\xi)) = U(\xi).Ad_Z(.) + f(Z,Z) \tag{107}$$

where $Z_M(\xi)$ is the fundamental vector field on the manifold *M* associated to $Z \in \mathfrak{g}$:

$$Z_M(\xi) = \delta\left[\bar{a}_M(\xi)\right] \text{ for } a = e \ , \ \delta a = Z \text{ and } \delta\xi = 0 \tag{108}$$

$$\sigma\left(Z_{1,M}(\xi), Z_{2,M}(\xi)\right) = \mu.[Z_1,Z_2] + f\left(Z_1,Z_2\right) \tag{109}$$

with σ the Lagrange form.

If we use previous relation $U\left(\bar{a}_M^{-1}(\xi)\right) = \bar{a}_{\mathfrak{g}^*}^{-1}\left(U(\xi)\right) + \theta\left(a^{-1}\right)$, and the property that $\bar{a}_{\mathfrak{g}^*}\left(U(\xi)\right) = U(\xi).\bar{a}_{\mathfrak{g}}^{-1}$, by defining:

$$\beta' = \bar{a}_{\mathfrak{g}}(\beta) \tag{110}$$

the integral is then defined by:

$$\int_M e^{-\beta'.U(\xi)}.d\omega = \int_M e^{-\bar{a}_{\mathfrak{g}}(\beta).U\left(\bar{a}_M^{-1}(\xi)\right)}.d\omega = \int_M e^{-\bar{a}_{\mathfrak{g}}(\beta).\left[\bar{a}_{\mathfrak{g}^*}^{-1}(U(\xi)) + \theta\left(a^{-1}\right)\right]}.d\omega = e^{\theta\left(a^{-1}\right).\beta}\int_M e^{-\beta.U(\xi)}.d\omega \tag{111}$$

We can then deduce the equation of Souriau theorem on Φ :

$$\Phi' = \Phi(\beta') = \Phi\left(\bar{a}_{\mathfrak{g}}(\beta)\right) = -\log\int_M e^{-\beta'.U(\xi)}.d\omega = -\log\left(e^{\theta\left(a^{-1}\right).\beta}\int_M e^{-\beta.U(\xi)}.d\omega\right) = \Phi(\beta) - \theta\left(a^{-1}\right)\beta \tag{112}$$

The equation of Souriau theorem on Q uses the relation $\bar{a}_{\mathfrak{g}^*}(Q) = Q.\bar{a}_{\mathfrak{g}}^{-1}$:

$$\Phi' = \Phi(\beta') = \Phi\left(\bar{a}_{\mathfrak{g}}(\beta)\right) = -\log\int_M e^{-\beta'.U(\xi)}.d\omega = -\log\left(e^{\theta\left(a^{-1}\right).\beta}\int_M e^{-\beta.U(\xi)}.d\omega\right) = \Phi(\beta) - \theta\left(a^{-1}\right)\beta \tag{113}$$

Finally, using $\bar{a}_{\mathfrak{g}^*}(Q) = Q.\bar{a}_{\mathfrak{g}}^{-1}$, we can prove that the Entropy is invariant:

$$s' = \beta'.Q' - \Phi' = \bar{a}_{\mathfrak{g}}(\beta).\left(\bar{a}_{\mathfrak{g}^*}(Q) + \theta(a)\right) - \left(\Phi + \theta(a)\bar{a}_{\mathfrak{g}}(\beta)\right) = \bar{a}_{\mathfrak{g}}(\beta).\bar{a}_{\mathfrak{g}^*}(Q) - \Phi = \bar{a}_{\mathfrak{g}}^{-1}\bar{a}_{\mathfrak{g}}(\beta).Q - \Phi = \beta.Q - \Phi = s \tag{114}$$

Considering the density of probability $p_\beta(\xi) = e^{-\beta.U(\xi)+\Phi(\beta)}$ with $\beta' = \bar{a}_{\mathfrak{g}}(\beta)$, then: $p_{\beta'}(\xi) = e^{-\bar{a}_{\mathfrak{g}}(\beta).U(\xi)+\Phi(\beta)-\theta\left(a^{-1}\right)\beta}$.

From which, we can recover $\bar{a}_M^+$ the image under $\bar{a}_M$ of the probability measure.

The last Souriau theorem is given by:

Souriau Theorem 3. Let $f = D(\theta)(e)$ be the derivative of θ (symplectic cocycle of G) at the identity element and let us define:

$$\forall \beta \in \Omega,\ \ f_\beta(Z_1, Z_2) = f(Z_1, Z_2) + Q.Ad_{Z_1}(Z_2) \ \text{ with } \ Ad_{Z_1}(Z_2) = [Z_1, Z_2] \tag{115}$$

Then

f_β *is a symplectic cocycle of* $\mathfrak{g}$ *, that is independent of the moment of G*

$$f_\beta(\beta, \beta) = 0 \ \ , \ \ \forall \beta \in \Omega \ \ (\beta \in Ker\ f_\beta) \tag{116}$$

- There exists a symmetric tensor g_β defined on the image of $Ad_\beta(.) = [.,\beta]$ such that:

$$g_\beta([\beta, Z_1], Z_2) = f_\beta(Z_1, Z_2) \ , \ \forall Z_1 \in \mathfrak{g}, \forall Z_2 \in \mathrm{Im}(Ad_\beta(.)) \tag{117}$$

and:

$$g_\beta(Z_1, Z_2) \geq 0 \ \ , \ \ \forall Z_1, Z_2 \in \mathrm{Im}(Ad_\beta(.)) \tag{118}$$

Last equation gives the structure of a positive Euclidean space.

$f_\beta(\beta,\beta) = 0$ could be deduced by differentiating $\Phi(\bar{a}_{\mathfrak{g}}(\beta)) = \Phi + \theta(a).\bar{a}_{\mathfrak{g}}(\beta)$ and taking $a = e$, $\delta a = Z_2$ and $\delta Z_1 = 0$. As $Z_M(\xi) = \delta[\bar{a}_M(\xi)]$ and $Z_{\mathfrak{g}} = -Ad_Z(.)$, we have $Q[Z_1, Z_2] = -f(Z_1, Z_2)$.

If we differentiate $Q(\bar{a}_{\mathfrak{g}}(\beta))\bar{a}_{\mathfrak{g}^*}(Q) + \theta(a)$, *the following relation* $\frac{\partial Q}{\partial \beta}(-[Z_1,\beta]) = f(Z_1,Z_1) + Q.Ad_{Z_1}(.) = f_\beta(Z_1,Z_1)$ appears. Then, writing $\delta\beta = [\beta, Z_1] = Z_2$, we have $\delta Q.\delta\beta \geq 0 \Rightarrow f_\beta(Z_1, Z_2) \geq 0$.

See more details in appendix A.3.

6. Synthesis of Analogies Between the Koszul Information Geometry Model and Souriau Statistical Physics Model

6.1. Comparison of Koszul and Souriau Models

We will synthetize in Table 1 results of previous chapters with Koszul Hessian Structure of Information Geometry and the Souriau model of Statistical Physics with the general concepts of geometric temperature, heat and capacity. Analogies between models will deal with characteristic function, Entropy, Legendre Transform, density of probability, dual coordinate systems, Hessian Metric and Fisher metric.

As $Q = \frac{\partial \Phi}{\partial \beta}$, we observe that the Information Geometry metric $I(\beta) = -\frac{\partial^2 \Phi(\beta)}{\partial \beta^2} = -\frac{\partial Q}{\partial \beta}$ could be considered as a *generalization of "Heat Capacity"*. Souriau called it K the "*Geometric Capacity*".

When $\beta = \frac{1}{kT}$, $K = -\frac{\partial Q}{\partial \beta} = -\frac{\partial Q}{\partial T}\left(\frac{\partial \frac{1}{kT}}{\partial T}\right) = \frac{1}{kT^2}\frac{\partial Q}{\partial T}$, then this geometric capacity is related to calorific capacity. Q is related to the mean, and K is related to the variance of U [122]:

$$I(\beta) = -\frac{\partial Q}{\partial \beta} = \mathrm{var}(U) = \int_M U(\xi)^2 . p_\beta(\xi) d\omega - \left(\int_M U(\xi).p_\beta(\xi) d\omega\right)^2 \tag{119}$$

Table 1. Synthesis of Koszul and Souriau models.

	Koszul Information Geometry Model	**Souriau Lie Groups Thermodynamics Model**
Characteristic function	$\Phi(x) = -\log \int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi \quad \forall x \in \Omega$	$\Phi(\beta) = -\log \int_M e^{-\beta.U(\xi)} d\omega \quad \forall \beta \in \mathfrak{g}$
Entropy	$\Phi^*(x^*) = -\int_{\Omega^*} p_x(\xi) \log p_x(\xi) d\xi$	$s = -\int_M p(\xi) \log p(\xi) d\omega$
Legendre Transform	$\Phi^*(x^*) = \left\langle x, x^* \right\rangle - \Phi(x)$	$s(Q) = \beta.Q - \Phi(\beta)$
Density of probability	$p_x(\xi) = e^{-\langle x, \xi\rangle + \Phi(x)}$ $p_x(\xi) = \frac{e^{-\langle \xi, x\rangle}}{\int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi}$	$p_\beta(\xi) = e^{-\beta.U(\xi) + \Phi(\beta)}$ $p_\beta(\xi) = \frac{e^{-\beta.U(\xi)}}{\int_M e^{-\beta.U(\xi)} d\omega}$
Dual Coordinate Systems	$x \in \Omega$ and $x^* \in \Omega^*$ $x^* = \int_{\Omega^*} \xi.p_x(\xi) d\xi = \frac{\int_{\Omega^*} \xi.e^{-\langle \xi, x\rangle} d\xi}{\int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi}$ $x^* = \frac{\partial \Phi(x)}{\partial x}$ and $x = \frac{\partial \Phi^*(x^*)}{\partial x^*}$	$\beta \in \mathfrak{g}$ and $Q \in \mathfrak{g}^*$ $Q = \int_M U(\xi).p_\beta(\xi) d\omega = \frac{\int_M U(\xi) e^{-\beta.U(\xi)} d\omega}{\int_M e^{-\beta.U(\xi)} d\omega}$ β : Souriau Geometric Temperature U : Souriau Moment map Q : Mean of Souriau Moment Map or Geometric heat $Q = \frac{\partial \Phi}{\partial \beta}$ and $\beta = \frac{\partial s}{\partial Q}$
Hessian Metric	$ds^2 = -d^2\Phi(x)$	$ds^2 = -d^2\Phi(\beta)$
Fisher metric	$I(x) = -E_\xi\left[\frac{\partial^2 \log p_x(\xi)}{\partial x^2}\right]$ $I(x) == -\frac{\partial^2 \Phi(x)}{\partial x^2} = \frac{\partial^2 \log \int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi}{\partial x^2}$	$I(\beta) = -E_\xi\left[\frac{\partial^2 \log p_\beta(\xi)}{\partial \beta^2}\right]$ $I(\beta) = -\frac{\partial^2 \Phi(\beta)}{\partial \beta^2} = \frac{\partial^2 \log \int_M e^{-\beta.U(\xi)} d\omega}{\partial \beta^2}$ $I(\beta) = -\frac{\partial^2 \Phi(\beta)}{\partial \beta^2} = -\frac{\partial Q}{\partial \beta}$ $K = -\frac{\partial Q}{\partial \beta}$: Souriau Geometric Capacity

6.2. Invariances in Koszul and Souriau Models

We have observed in previous chapters the main invariances characterizing the Koszul Model and the Souriau Model. We will synthetize these invariances in Table 2.

In both the Koszul and Souriau models, the Information Geometry Metric and the Entropy are invariant respectively to the automosphisms g of the convex cone Ω and to $\overline{a}_{\mathfrak{g}}$ adjoint representation of Dynamical group G acting on Ω, the convex cone considered as largest open subset of $\mathfrak{g}$, Lie algebra of G, such that $\int_M e^{-\beta.U(\xi)}d\omega$ and $\int_M \xi.e^{-\beta.U(\xi)}d\omega$ are convergent integrals.

6.3. Souriau Thermometer

Souriau has built a thermometer (θερμός) device principle that could measure the Geometric Temperature using "*Relative Ideal Gas Thermometer*" based on a theory of Dynamical Group Thermometry, and has also recovered the Laplace barometric law $p(\vec{r}) \propto e^{-m\beta\langle\vec{g},\vec{r}\rangle}$.

Table 2. Comparison of invariances for the Koszul and Souriau models.

	Koszul Information Geometry Model	**Souriau Lie Groups Thermodynamics Model**
Convex Cone	$x \in \Omega$ Ω *convex cone*	$\beta \in \Omega$ Ω *convex cone: largest open subset of* $\mathfrak{g}$, *Lie algebra of G, such that* $\int_M e^{-\beta.U(\xi)}d\omega$ *and* $\int_M \xi.e^{-\beta.U(\xi)}d\omega$ *are convergent integrals*
Transformation	$x \to gx$ with $g \in Aut(\Omega)$	$\beta \to \overline{a}_{\mathfrak{g}}(\beta)$
Transformation of Potential (non invariant)	$\Phi_\Omega(x) \to \Phi_\Omega(gx) = \Phi_\Omega(x) + \log(\lvert\det g\rvert)$	$\Phi(\beta) \to \Phi(\overline{a}_{\mathfrak{g}}(\beta)) = \Phi(\beta) - \theta(a^{-1})\beta$
Transformation of Entropy (invariant)	$\Phi^*_{\Omega^*}(x^*) \to \Phi^*_{\Omega^*}\left(\frac{\partial \Phi_\Omega(gx)}{\partial x}\right) = \Phi^*_{\Omega^*}(x^*)$ with $x^* = \frac{\partial \Phi_\Omega(x)}{\partial x}$	$s(Q) \to s'(Q') = \beta'.Q' - \Phi' = \beta.Q - \Phi = s(Q)$ .with $\beta' = \overline{a}_{\mathfrak{g}}(\beta)$ $Q' = \frac{\partial \Phi'}{\partial \beta'} = \frac{\partial(\Phi + \theta(a)\overline{a}_{\mathfrak{g}}(\beta))}{\partial \overline{a}_{\mathfrak{g}}(\beta)} = a_{\mathfrak{g}^*}(Q) + \theta(a)$ $\Phi' = \Phi(\beta') = \Phi(\overline{a}_{\mathfrak{g}}(\beta)) = \Phi(\beta) - \theta(a^{-1})\beta$
Information Geometry Metric (invariant)	$I(gx) = -\frac{\partial^2[\Phi_\Omega(x) + \log(\lvert\det g\rvert)]}{\partial x^2}$ $I(gx) = -\frac{\partial^2 \Phi_\Omega(x)}{\partial x^2} = I(x)$	$I(\overline{a}_{\mathfrak{g}}(\beta)) = -\frac{\partial^2[\Phi(\beta) - \theta(a^{-1})\beta]}{\partial \beta^2} = -\frac{\partial^2 \Phi(\beta)}{\partial \beta^2} = I(\beta)$

7. From Characteristic Function to Generative Inner Product

Cartan's works have greatly influenced Koszul (Koszul's PhD thesis extended previous work of Cartan) and Souriau (Souriau was a student of Elie Cartan at ENS, the year after his aggregation). We have shown that "Information Geometry" could be considered as a particular application domain of Hessian Geometry through Koszul's work (Koszul-Vinberg metric deduced from the associated characteristic function having the main property of being invariant to all automorphisms of the convex cone), that could be extended in the framework of Souriau's theory, as an extension towards "Lie Group Thermodynamics" with vector-valued geometric temperature (providing a geometric extension of Noether's theorem). Should we deduce that the "essence" of Information Geometry is limited to the "Koszul Characteristic Function"? This notion seems to not be the more general one, and we will explore the notion of Generative Inner Products. We will reduce Koszul's and Souriau's definitions to exclusive "Inner Product" selection using symmetric bilinear "Cartan-Killing form" introduced by Cartan in 1894.

In Koszul Geometry, we have two convex dual functions $\Phi(x)$ and $\Phi^*(x^*)$ with dual system of coordinates x and x^* defined on dual cones Ω and Ω^*: $\Phi(x) = -\log \int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi \quad \forall x \in \Omega$ and $\Phi^*(x^*) = \langle x, x^* \rangle - \Phi(x)$. We can then remark that if we can define an Inner Product $\langle .,. \rangle$, we will be able to build a convex function $\Phi(x) = -\log \psi_\Omega(x)$ and its dual by Legendre transform because both are only dependent of the Inner product, and dual coordinate is also defined by $x^* = \arg\min \left\{ \psi_\Omega(y) / y \in \Omega^*, \langle x, y \rangle = n \right\} = \int_{\Omega^*} \xi . e^{-\langle \xi, x\rangle} d\xi \Big/ \int_{\Omega^*} e^{-\langle \xi, x\rangle} d\xi$ where x^* is also the center of gravity of the cross section $\left\{ y \in \Omega^*, \langle x, y \rangle = n \right\}$ of Ω^* (with notation: $\Phi(x) = -\log \psi_\Omega(x)$).

It is not possible to define an ad(g)-invariant inner product for any two elements of a Lie Algebra, but a symmetric bilinear form, called "Cartan-Killing form", could be introduced. This form has been introduced first by Cartan in 1894 in his PhD thesis. This form is defined according to the adjoint endomorphism Ad_x of g that is defined for every element x of g with the help of the Lie bracket:

$$Ad_x(y) = [x, y] \tag{120}$$

The trace of the composition of two such endomorphisms defines a bilinear form, the Cartan-Killing form:

$$B(x, y) = Tr\left(Ad_x Ad_y \right) \tag{121}$$

The Cartan-Killing form is symmetric:

$$B(x, y) = B(y, x) \tag{122}$$

and has the associativity property:

$$B([x, y], z) = B(x, [y, z]) \tag{123}$$

given by:

$$B([x, y], z) = Tr\left(Ad_{[x,y]} Ad_z \right) = Tr\left(\left[Ad_x, Ad_y \right] Ad_z \right) = Tr\left(Ad_x \left[Ad_y, Ad_z \right] \right) = B(x, [y, z]) \tag{124}$$

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

$$B(\sigma(x), \sigma(y)) = B(x, y) \tag{125}$$

To prove this invariance, we have to consider:

$$\begin{cases} \sigma[x,y]=[\sigma(x),\sigma(y)] \\ z=\sigma(y) \end{cases} \Rightarrow \sigma\left[x,\sigma^{-1}(z)\right]=[\sigma(x),z] \quad \text{rewritten} \quad Ad_{\sigma(x)}=\sigma\circ Ad_x\circ\sigma^{-1} \tag{126}$$

Then:

$$B\left(\sigma(x),\sigma(y)\right)=Tr\left(Ad_{\sigma(x)}Ad_{\sigma(y)}\right)=Tr\left(\sigma\circ Ad_xAd_y\circ\sigma^{-1}\right)=Tr\left(Ad_xAd_y\right)=B(x,y) \tag{127}$$

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

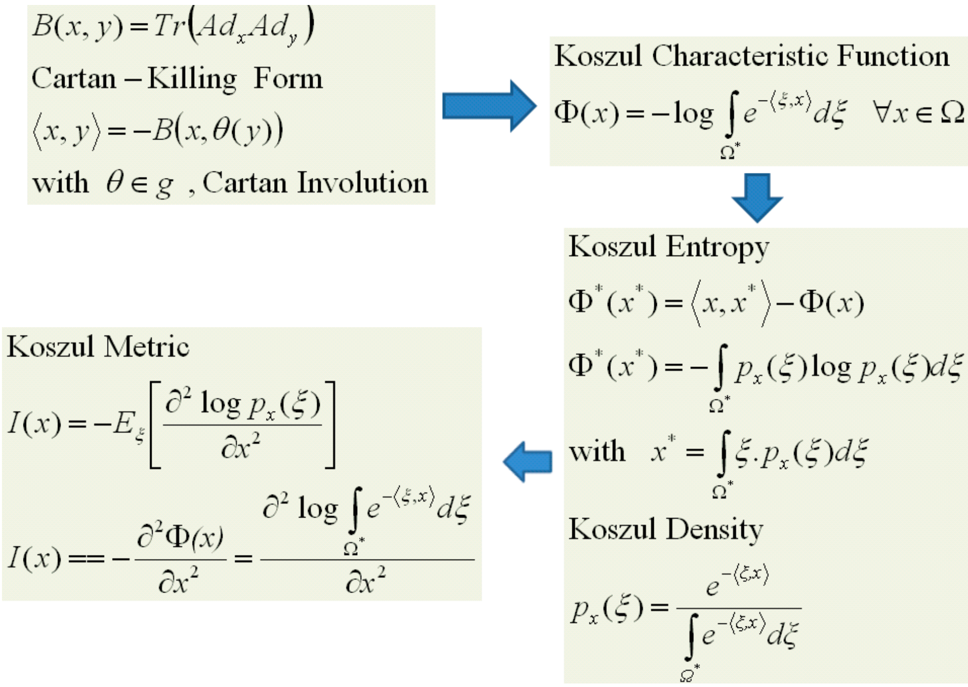
Cartan Generative Inner Product: The following Inner product defined by Cartan-Killing form is invariant by automorphisms of the algebra

$$\langle x,y\rangle=-B\left(x,\theta(y)\right) \tag{128}$$

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

From the Cartan Inner Product, we can generate logarithm of the Koszul Characteristic Function, and its Legendre Transform to define Koszul Entropy, Koszul Density and Koszul Metric, as explained in the following Figure 5:

Figure 5. Generation of Koszul elements from Cartan Inner Product.



In Appendix A2, we give the definition of another inner product, Gromov Inner product, in CAT(−1) space, that could be also used to generalize Koszul definition of Characteristic Function.

On the concept of generative structure, we could also explore the notion of Generative Function [123–126] and come back to seminal paper of Chentsov about axiomatization of Information Geometry [127].

8. Conclusions on General Definition of Entropy by Legendre Transform

Definition of Entropy has been widely debated [128,129]. Based on the cornerstone concept of the Koszul Vinberg Characteristic Function, we have introduced Koszul Entropy as the Legendre transform of its logarithm. This definition of Entropy could be extended by interpretating Legendre transform as Fourier transform in (Min,+) algebra [130,131].

As we have observed previously, Koszul Entropy has a Shannon Entropy structure:

$$\begin{aligned}
&\Phi^*(x^*) = \Phi^*(E[\xi]) = -\int_{\Omega^*} p_x(\xi)\log p_x(\xi)d\xi = \int_{\Omega^*} \Phi^*(\xi)p_x(\xi)d\xi = E\left[\Phi^*(\xi)\right] \\
&\text{with } \ \Phi^*(\xi) = -\log p_x(\xi) \ \text{ and } \ x^* = \int_{\Omega^*} \xi.p_x(\xi)d\xi = E[\xi] \\
&\text{where } \ p_x(\xi) = \frac{e^{-\langle \xi, x\rangle}}{\int_{\Omega^*} e^{-\langle \xi, x\rangle}d\xi} = e^{-\langle x,\xi\rangle - \log \int_{\Omega^*} e^{-\langle \xi,x\rangle}d\xi} = e^{-\langle x,\xi\rangle + \Phi(x)}
\end{aligned} \tag{129}$$

In last equation, variable x could be defined by $\bar{\xi} = E[\xi] = x^*$ if function $\frac{d\Phi(x)}{dx}$ could be inverted:

$$p_{\bar{\xi}}(\xi) = \frac{e^{-\langle \xi, \Theta^{-1}(\bar{\xi})\rangle}}{\int_{\Omega^*} e^{-\langle \xi, \Theta^{-1}(\bar{\xi})\rangle}d\xi} \ \text{ with } \ x = \Theta^{-1}(\bar{\xi}) \ \text{ and } \ \bar{\xi} = \Theta(x) = \frac{d\Phi(x)}{dx} \tag{130}$$

where:

$$\bar{\xi} = \int_{\Omega^*} \xi.p_{\bar{\xi}}(\xi)d\xi \ \text{ and } \ \Phi(x) = -\log \int_{\Omega^*} e^{-\langle x,\xi\rangle}d\xi \tag{131}$$

In previous chapters, a definition of Koszul Entropy $\Phi^*(x^*)$ through Legendre transform of Koszul-Vinberg characteristic function $\Phi(x)$ has been given:

$$\begin{aligned}
&\Phi^*(x^*) = \left\langle x, x^* \right\rangle - \Phi(x) \\
&\text{with } \ \Phi(x) = -\log \int_{\Omega^*} e^{-\langle \xi, x\rangle}d\xi \quad \forall x \in \Omega
\end{aligned} \tag{132}$$

where $\Phi(x)$ could be interpreted as opposite of logarithm of Laplace transform [132,133]:

$$Entropy = Legendre\left[-\log\left[Laplace\right]\right] \tag{133}$$

that we will write synthetically as:

$$Ent = -Leg \circ Log \circ Lapl \tag{134}$$

The function $Leg \circ Log \circ Lapl$ is sometimes called “Cramer transform”.

If we remark that the Legendre transform is closely related to the idempotent analogue of the Fourier transform [130,131,134–136], we could then give a new definition of Entropy.

If we consider the semiring $R_{\min} = R \cup \{+\infty\}$ with the operations $\oplus = Min$ and $\bullet = +$. In $R_{\min} = R \cup \{+\infty\}$ the idempotent analogues of integration on R^N is given by the formula:

$$I(f)=\int_{R^N}^{\oplus} f(x)dx = \underset{x\in R^N}{Inf}\, f(x) \tag{135}$$

Then, the Legendre transform is equivalent to the Fourier transform in $(\oplus,\bullet)=(Min,+)$ algebra [130]:

$$\Phi^*(\xi)=\underset{x\in\Omega}{Sup}\left[\langle x,\xi\rangle-\Phi(x)\right]=-\int_{\Omega}^{\oplus}\left(-\langle x,\xi\rangle\right)\bullet\Phi(x)dx = Four_{(Min,+)}\left[\Phi(x)\right] \tag{136}$$

The Legendre transform generates an idempotent version of harmonic analysis for the space of convex functions. We can then give a general definition of Entropy:

$$Ent=-Four_{(Min,+)}\circ Log\circ Lapl_{(+,\times)} \tag{137}$$

We can also observe the following properties deduced from the Laplace and Legendre transforms' characteristics:

$$Ent(\mu\otimes\gamma)=Ent(\mu)\bullet Ent(\gamma) \tag{138}$$

where * is the convolution operator and $\otimes$ the inf-convolution operator (see [130] for the definition of inf-convolution) defined by:

$$[f\bullet g](z)=\underset{x}{Inf}\left[f(x)+g(y-x)\right] \tag{139}$$

with f and g, two functions $R\to R_{\min}$.

> "*La théorie cinétique des gaz laisse encore subsister bien des points embarrassants pour ceux qui sont accoutumés à la rigueur mathématique… L'un des points qui m'embarrassaient le plus était le suivant: il s'agit de démontrer que l'entropie va en diminuant, mais le raisonnement de Gibbs semble supposer qu'après avoir fait varier les conditions extérieures on attend que le régime soit établi avant de les faire varier à nouveau. Cette supposition est-elle essentielle, ou en d'autres termes, pourrait-on arriver à des résultats contraires au principe de Carnot en faisant varier les conditions extérieures trop vite pour que le régime permanent ait le temps de s'établir?* "

Henri Poincaré « Réflexions sur la théorie cinétique des gaz », 1906

> [*The kinetic theory of gases leaves awkward points for those who are accustomed to mathematical rigor … One of the points which embarrassed me most was the following one: it is a question of demonstrating that the entropy keeps decreasing, but the reasoning of Gibbs seems to suppose that having made vary the outside conditions we wait that the regime is established before making them vary again. Is this supposition essential, or in other words, we could arrive at opposite results to the principle of Carnot by making vary the outside conditions too fast so that the permanent regime has time to become established* ?]

Henri Poincaré "Reflection on The kinetic theory of gases", 1906

"*Quel est l'objet de l'art ? Si la réalité venait frapper directement nos sens et notre conscience, si nous pouvions entrer en communication immédiate avec les choses et avec nous-mêmes, je crois bien que l'art serait inutile, ou plutôt que nous serions tous artistes, car notre âme vibrerait alors continuellement à l'unisson de la nature.*"

Henri Bergson, Le rire, p.115, Éd. P.U.F

[*What is the object of art? Could reality come into direct contact with sense and consciousness, could we enter into immediate communion with things and with ourselves, probably art would be useless, or rather we should all be artists, for then our soul would continually vibrate in perfect accord with nature.*]

Henri Bergson, Laughter

Acknowledgments

Many thanks are due to members of the Leon Brillouin seminar with very fruitful discussions on Geometric Science of Information and Information Geometry, initiated since December 2009. Souriau Models were more clearly understood with the help of Claude Vallée that sent me unpublished chapters of Souriau's book [86]. Jean-Louis Koszul has participated to the first GSI conference at Ecole des Mines in 2013, where Hirohiko Shima gave a keynote lecture on Koszul Hessian Geometry.

Appendix

A1. Legendre Transform and Minimal Surface

Laplace contribution to probability was around 1774 [137]. At almost the same period, in 1787, Adrien-Marie Legendre has introduced the "Legendre Transform" [138] to solve the Minimal Surface Problem equation introduced by Lagrange and partially solved by Gaspard Monge in 1784 [139]. In 1768, Lagrange considered the variational problem of least area surface stretched across a given closed contour. Based on Euler-Lagrange equation, Lagrange has introduced the equation of *Minimal Surface* $z(x,y)$:

$$\left(1+q^2\right)\frac{d^2z}{dx^2}-2pq\frac{d^2z}{dxdy}+\left(1+p^2\right)\frac{d^2z}{dy^2}=0 \quad \text{with} \quad \frac{dz}{dx}=p \quad \text{and} \quad \frac{dz}{dy}=q \tag{140}$$

Lagrange has observed that affine functions $z(x,y)=a.x+b.y+c$ are solutions of this equation and minimal surfaces are planes.

Jean-Baptiste Marie Meusnier de La Place, a student of Monge, has observed that for this surface, two curvature radiuses are everywhere equal but directed in opposite direction, because first equation is equal to two times the mean curvature H_z:

$$2H_z=\frac{d}{dx}\left(\frac{\dfrac{dz}{dx}}{\sqrt{1+\left(\dfrac{dz}{dx}\right)^2+\left(\dfrac{dz}{dy}\right)^2}}\right)+\frac{d}{dy}\left(\frac{\dfrac{dz}{dy}}{\sqrt{1+\left(\dfrac{dz}{dx}\right)^2+\left(\dfrac{dz}{dy}\right)^2}}\right)=\frac{\left(1+q^2\right)\dfrac{d^2z}{dx^2}-2pq\dfrac{d^2z}{dxdy}+\left(1+p^2\right)\dfrac{d^2z}{dy^2}}{\left(1+\left(\dfrac{dz}{dx}\right)^2+\left(\dfrac{dz}{dy}\right)^2\right)^{3/2}} \tag{141}$$

Gaspard Monge integrated this equation in [139] but with a non-rigorous approach and asked Legendre to find a more classical solution. For this task, Legendre has introduced a change of variable that is the nowadays well-known "Legendre transform". Adrien-Marie Legendre said "*J'y suis parvenu simplement par un changement de variables qui peut être utile dans d'autres occasions*"(*"I reached there simply by a change of variables which can be useful in other opportunities"*).

Legendre reduced the problem to solve to determine p and q as functions of x and y such that:

$$p.dx + q.dy \quad \text{and} \quad \frac{p.dy - q.dx}{\sqrt{1+p^2+q^2}} \tag{142}$$

are exact differentials. If we set $1+p^2+q^2=u^2$, then these other expressions are complete differentials:

$$x.dp + y.dq \quad \text{and} \quad y.d\left(\frac{p}{u}\right) + x.d\left(\frac{q}{u}\right) \tag{143}$$

Legendre considered x and y as functions of p and q:

$$x.dp + y.dq = d\omega \quad \text{with} \quad x = \frac{d\omega}{dp} \quad \text{and} \quad y = \frac{d\omega}{dq} \tag{144}$$

If we then develop $y.d\left(\frac{p}{u}\right)+x.d\left(\frac{q}{u}\right)$, we have:

$$\left[\left(1+q^2\right)y + pq.x\right]\frac{dp}{u^3} - \left[\left(1+p^2\right)x + pq.y\right]\frac{dq}{u^3} \tag{145}$$

That should be an exact differential. By replacing x and y, we have a new equation:

$$\left(1+q^2\right)\frac{d^2\omega}{dq^2} + 2pq.\frac{d^2\omega}{dpdq} + \left(1+p^2\right)\frac{d^2\omega}{dp^2} = 0 \tag{146}$$

This new equation is very similar to the previous one, but simpler because it depends on p and q and not their partial differentials of first order. When the function ω will be known, then functions x, y and z will be also defined according to p and q thanks to "*Legendre transform*":

$$\begin{gathered} z(x,y) = p.x + q.y - \omega(p,q) \\ \text{with} \quad x = \frac{d\omega}{dp} \quad \text{and} \quad y = \frac{d\omega}{dq} \end{gathered} \tag{147}$$

About this Legendre transform, Darboux [140] gave an interpretation by Chasles "*Ce qui revient suivant une remarque de M. Chasles, à substituer à la surface sa polaire réciproque par rapport à un paraboloïde*" [*What is equivalent according to M. Chasles's remark, to substitute for the surface its mutual polar with regard to a paraboloid*]. This equation could be also written as classical "*Legendre transform*" with our previous notations:

$$s(Q) = \beta.Q - \Phi(\beta) = \langle \beta, Q \rangle - \Phi(\beta)$$

$$\text{with} \quad \begin{cases} \Phi(\beta) = z(x,y) \\ s(Q) = \omega(p,q) \end{cases}, \quad \begin{cases} Q = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \begin{bmatrix} p \\ q \end{bmatrix} \\ \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \end{cases} \quad \text{and} \quad \begin{cases} Q = \begin{bmatrix} \dfrac{d\Phi}{d\beta_1} \\ \dfrac{d\Phi}{d\beta_2} \end{bmatrix} = \dfrac{d\Phi}{d\beta} \\ \beta = \begin{bmatrix} \dfrac{ds}{dQ_1} \\ \dfrac{ds}{dQ_2} \end{bmatrix} = \dfrac{ds}{dQ} \end{cases} \tag{148}$$

In the following relation, we recover the definition of Entropy $ds = \beta.dQ = \dfrac{dQ}{T}$:

$$\begin{cases} x.dp + y.dq = d\omega \\ x = \dfrac{d\omega}{dp} \ \text{and} \ y = \dfrac{d\omega}{dq} \end{cases} \Rightarrow \begin{cases} \beta.dQ = ds \\ \beta = \dfrac{ds}{dQ} \end{cases} \tag{149}$$

The equation of the surface is characterized by the following equation:

$$\frac{d}{d\beta}\left(\frac{Q}{\sqrt{1+\|Q\|^2}}\right) = \frac{d}{d\beta}\left(\frac{\dfrac{d\Phi}{d\beta}}{\sqrt{1+\left\|\dfrac{d\Phi}{d\beta}\right\|^2}}\right) = 2.H_\Phi \quad \text{or} \quad div_\beta\left(\frac{Q}{\sqrt{1+\|Q\|^2}}\right) = div_\beta\left(\frac{\nabla\Phi}{\sqrt{1+\|\nabla\Phi\|^2}}\right) = 2.H_\Phi \tag{150}$$

We can then observed that when $\|Q\| << 1$, $\dfrac{d}{d\beta}\left(\dfrac{Q}{\sqrt{1+\|Q\|^2}}\right) \approx \dfrac{dQ}{d\beta} = -I(\beta) = 2.H_\Phi$.

We can also characterized Entropy with this 2nd equation:

$$\left(1+Q_2^2\right)\frac{d^2 s(Q)}{dQ_2^2} + 2pq.\frac{d^2 s(Q)}{dQ_1 dQ_2} + \left(1+Q_1^2\right)\frac{d^2 s(Q)}{dQ_1^2} = 0 \tag{151}$$

We can also find direct equations for x, y and z, based on "Legendre transform" and Equation (146):

$$\left(1+q^2\right)\frac{d^2 x}{dq^2} + 2pq.\frac{d^2 x}{dpdq} + \left(1+p^2\right)\frac{d^2 x}{dp^2} + 2q\frac{dx}{dp} + 2p\frac{dx}{dq} = 0 \tag{152}$$

We have exactly same equations for y and z.

Legendre then solved Equations (145) and (148), by determining two constants a and b given by double integral of the equation:

$$\left(1+q^2\right)dp^2 - 2pq.dp.dq + \left(1+p^2\right)dp^2 = 0 \tag{153}$$

By selecting $p = aq + A$ with a and A two constants. Previous equation gives $1 + a^2 + A^2 = 0$. Then a will be let an arbitrary function and $A = \pm\sqrt{-1-a^2}$. Two integrals of Equation (129) will be:

$$\begin{cases} p = aq + \sqrt{-1-a^2} = aq + A \\ p = bq - \sqrt{-1-b^2} = bq + B \end{cases} \tag{154}$$

with *a* and *b* two arbitrary constants, roots of the following Equation:

$$\left(1+q^2\right)v^2 - 2pq..v + \left(1+p^2\right) = 0$$
$$\text{with} \quad \begin{cases} a+b = \dfrac{2pq}{1+q^2} \\ ab = \dfrac{1+p^2}{1+q^2} \end{cases} \tag{155}$$

Equations (145) and (148) could be then simplified:

$$\begin{aligned} &(a-b)\frac{d^2\omega}{dadb} - \frac{A}{B}.\frac{d\omega}{da} + \frac{B}{A}.\frac{d\omega}{db} = 0 \\ &\frac{d^2x}{dadb} = 0 \end{aligned} \tag{156}$$

Then Legendre deduced that three coordinates could be given by two arbitrary functions:

$$\begin{cases} x = \dfrac{d\varphi}{da} + \dfrac{d\psi}{db} \\ y = \varphi - a\dfrac{d\varphi}{da} + \psi - b\dfrac{d\psi}{db} \\ z = -\int A\dfrac{d^2\varphi}{da^2}da + \int B\dfrac{d^2\psi}{db^2}db \end{cases} \tag{157}$$

This is the integral solution of "*Minimal surface*" Lagrange equation (Legendre recovered the solution given by Monge in 1784).

A2. Gromov Inner Product

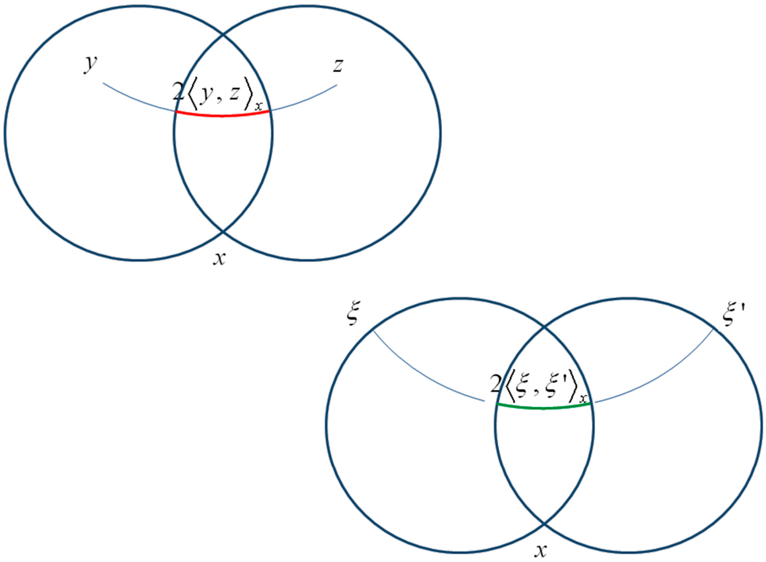
As other generalization of inner product, we can consider for specific case *CAT(-1)-space*[141,142] (generalization of simply connected Riemannian manifold of negative curvature lower than unity) or for an Homogeneous Symmetric Bounded domains, a "generative" Gromov Inner Product between points y–z (relatively to *x*) that is defined by the distance [143,144]:

$$\left\langle y,z\right\rangle_x = \frac{1}{2}\left(d(x,y) + d(x,z) - d(y,z)\right) \tag{158}$$

with *d(.,.)* the distance in CAT(−1). This Gromov Inner Product is illustrated in Figure 6. Intuitively, this inner product measures the distance of *x* to the geodesics between *y* to *z*. This Inner product could be also defined for points on the Shilov Boundary of the domain through Busemann distance:

$$\left\langle \xi,\xi'\right\rangle_x = \frac{1}{2}\left(B_\xi(x,p) + B_{\xi'}(x,p)\right) \tag{159}$$

Figure 6. Gromov Inner product in homogeneous bounded domains and its Shilov boundary.



Independent of p, where $B_\xi(x,y)=\underset{t\to+\infty}{Lim}\left[\left|x-r(t)\right|-\left|y-r(t)\right|\right]$ is the horospheric distance, from x to y relatively to ξ, with $r(t)$ geodesic ray. We have the property that:

$$\langle \xi,\xi'\rangle_x = \underset{\substack{y\to\xi \\ y'\to\xi'}}{Lim}\langle y,y'\rangle_x \tag{160}$$

We can then define a visual metric on the Shilov boundary by:

$$\begin{aligned} &d_x(\xi,\xi')=e^{-\langle \xi,\xi'\rangle_x} \quad \text{if} \quad \xi\neq\xi' \\ &d_x(\xi,\xi')=0 \quad \text{otherwise} \end{aligned} \tag{161}$$

We can then define the characteristic function according to the origin *0*:

$$\Phi(x)=-\log\int_{\Omega^*} e^{-\langle x,\gamma\rangle_0}d\gamma \text{ or } \Phi_\Omega(x)=-\log\int_{\Omega^*} e^{-\frac{1}{2}\left(d(0,x)+d(0,\gamma)-d(x,\gamma)\right)}d\gamma \tag{162}$$

and:

$$\Phi^*(x^*)=\langle x,x^*\rangle_0-\Phi(x)=\frac{1}{2}\left(d(0,x)+d(0,x^*)-d(x,x^*)\right)-\Phi(x) \tag{163}$$

$$d(x,x^*)=\left(d(0,x^*)-2\Phi^*(x^*)\right)+\left(d(0,x)-2\Phi(x)\right) \tag{164}$$

with the center of gravity:

$$x^*=\int_{\Omega^*}\gamma.e^{-\langle x,\gamma\rangle_0}d\gamma \Big/ \int_{\Omega^*}e^{-\langle x,\gamma\rangle_0}d\gamma \tag{165}$$

All these relations are also true on the Shilov Boundary:

$$\Phi(\xi)=-\log\int_{\partial\Omega^*}e^{-\langle \xi,\xi'\rangle_0}d\xi'=-\log\int_{\partial\Omega^*}d_0(\xi,\xi').d\xi' \tag{166}$$

where $\int_{\partial\Omega^*}d_0(\xi,\xi').d\xi'$ is the functional of Busemann barycenter on the Shilov Boundary $\partial\Omega^*$ (existence and unicity of this barycenter have been proved by Cartan [14] for Cartan-Hadamard Spaces).

A3. The Cohomology of a Dynamical Group

In the following, we give some details of Souriau development about the Moment of the G action (see Figure 7) and the Cohomology of a dynamical group (see Figure 8). Other details about Symplectic Geometry could be found in [145] or [146].

Figure 7. Moment of the G action.

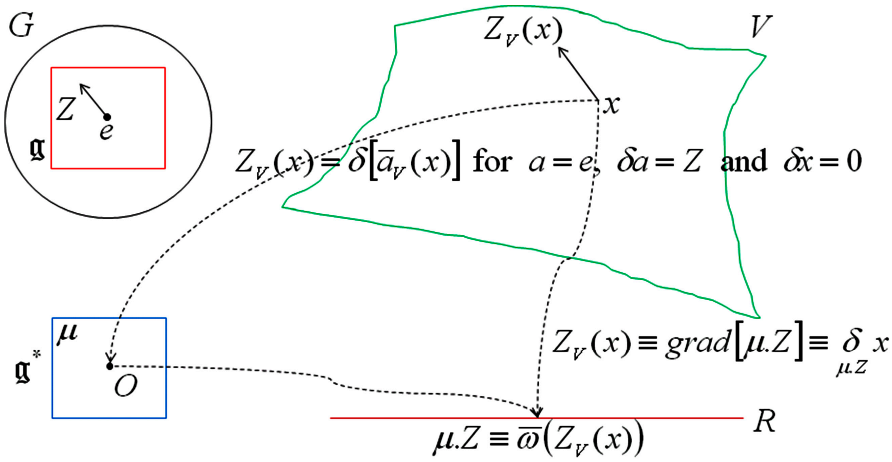
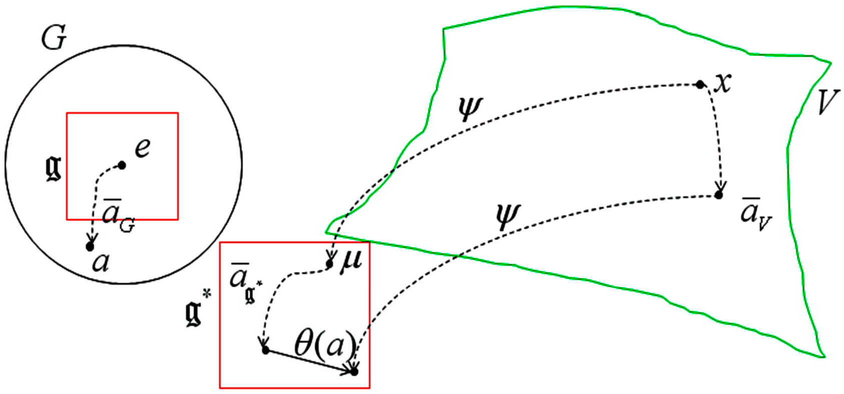


Figure 8. The Cohomology of a dynamical group.



If G is a dynamical group of a symplectic manifold V, torsor μ is called a moment of the G-action, if there is a differential map $x \mapsto \mu$ from V to $\mathfrak{g}^*$ such that:

$$\sigma(Z_V(x)) \equiv -d[\mu.Z] \qquad (167)$$

To every torsor μ, there corresponds a field $[x \mapsto \overline{\omega}]$ of 1-forms (Maurer-Cartan forms) on G which is invariant under right translation and which takes the value μ when x is the identity element:

$$\sigma_V = d\overline{\varpi} \qquad (168)$$

Using the moment of the G action, Souriau has introduced the following theorem on the Cohomology of a dynamical group:

Theorem. Let V be a connected symplectic manifold and let G be a dynamical group of V possessing a moment μ. Finally, let ψ denote the map $x \mapsto \mu$ from V to the space $\mathfrak{g}^*$ of torsor of G:

There exists a differential map $\theta : G \to \mathfrak{g}^*$:

$$\theta(a) \equiv \psi(\bar{a}_V(x)) - \bar{a}_{\mathfrak{g}^*}(\psi(x)) \tag{169}$$

The derivative $f = D(\theta)(e)$ is a 2-form on the Lie algebra $\mathfrak{g}$ of G:

$$f(Z)([Z', Z'']) + f(Z')([Z'', Z]) + f(Z'')([Z, Z']) \equiv 0 \tag{170}$$

Then, the following identities hold:

$$\sigma(Z_V(x))(Z_V'(x)) \equiv \mu.[Z, Z'] + f(Z)(Z') \tag{171}$$

$$D(\psi)(x)(Z_V(x)) \equiv \psi(x).ad(Z) + f(Z) \tag{172}$$

Conflicts of Interest

The author declares no conflict of interest.

References

1. Massieu, F. Sur les Fonctions caractéristiques des divers fluides. *Comptes Rendus de l'Académie des Sciences* **1869**, *69*, 858–862. (In French)
2. Massieu, F. Addition au précédent Mémoire sur les Fonctions caractéristiques. *Comptes Rendus de l'Académie des Sciences* **1869**, *69*, 1057–1061. (In French)
3. Massieu, F. Thermodynamique: Mémoire sur les Fonctions Caractéristiques des Divers Fluides et sur la Théorie des Vapeurs; Académie des Sciences: Paris, France, 1876; p. 92. (In French)
4. Duhem, P. Sur les équations générales de la Thermodynamique. *Annales Scientifiques de l'Ecole Normale Supérieure* **1891**, *8*, 231–266. (In French)
5. Duhem, P. Commentaire aux principes de la Thermodynamique—Première partie. *Journal de Mathématiques pures et appliquées* **1892**, *8*, 269–330. (In French)
6. Duhem, P. Commentaire aux principes de la Thermodynamique—Troisième partie. *Journal de Mathématiques pures et appliquées* **1894**, *10*, 207–286. (In French)
7. Duhem, P. Les théories de la chaleur. *Revue des deux Mondes* **1895**, *130*, 851–868.
8. Gibbs, J.W. *Graphical Methods in the Thermodynamics of Fluids*. In *The Scientific Papers of J. Willard Gibbs*; Bumstead, H.A., van Name, R.G.; Eds.; Dover: New York, NY, USA, 1961.
9. Poincaré, H. *Calcul des Probabilités*; Gauthier-Villars: Paris, France, 1896. (In French)
10. Poincaré, H. *Thermodynamique, Cours de Physique Mathématique*; Carré, G., Ed.; 1892; Available online: http://gallica.bnf.fr/ark:/12148/bpt6k2048983 (accessed on 30 July 2014; In French)
11. Vinberg, E.B. Structure of the Group of Automorphisms of a Homogeneous Convex Cone. *Trudy Moskovskogo Matematicheskogo Obshchestva* **1965**, *13*, 56–83.
12. Faraut, J.; Koranyi, A. *Analysis on Symmetric Cones, Oxford Mathematical Monographs*; The Clarendon Press; Oxford University Press: New York, NY, USA, 1994.
13. Lichnerowicz, A. Espaces homogènes Kähleriens. In *Colloque de Géométrie Différentielle*; Pub. du CNRSP: Paris, France, 1953; pp. 171–184. (In French)

14. Cartan, E. Sur les domaines bornés de l'espace de n variables complexes. *Abh. Math. Seminar Hamburg* **1935**, *1*, 116–162. (In French)
15. Siegel, C.L. Über der analytische Theorie der quadratischen Formen. *Ann. Math.* **1935**, *36*, 527–606. (In German)
16. Siegel, C.L. Symplectic geometry. *Amer. J. Math*. **1943**, *65*, 1–86.
17. Marle, C.M. On mechanical systems with a Lie group as configuration space. In *Jean Leray'99 Conference*, Proceedings of the Karlskrona Conference in the Honor of Jean Leray, Kluwer, Dordrecht, The Netherlands, 2003; de Gosson, M., Ed.; Springer: Berlin/Heidelberg, Germany, 2003; pp. 183–203.
18. Marle, C.M. On Henri Poincaré's note "Sur une forme nouvelle des équations de la mécanique". *J. Geom. Symmetry Phys.* **2013**, *29*, 1–38.
19. Friedrich, T. Die Fisher-Information und symplektische Strukturen. *Math. Nachr*. **1991**, *153*, 273–296. (In German)
20. Gromov, M. In a Search for a Structure, Part 1: On Entropy. Available online: http://www.ihes.fr/~gromov/PDF/structre-serch-entropy-july5-2012.pdf (accessed on 23 June 2013).
21. Gromov, M. Convex Sets and Kähler Manifolds. In *Advances in Differential Geometry and Topology*; Tricerri, F., Ed.; World Scientific: Singapore, Singapore, 1990; pp. 1–38.
22. Gromov, M. Entropy and Isoperimetry for Linear and non-Linear Group Actions. *Groups Geom. Dyn*. **2008**, *2*, 499–593.
23. Gromov, M. Carnot-Carathéodory spaces seen from within. In *Progress in Mathematics*; Springer: Berlin/Heidelberg, Germany, 1996; Volume 144.
24. Ollivier, Y.; Akimoto, Y. Objective Improvement in Information-Geometric Optimization. In *Foundations of Genetic Algorithms XII*; Springer: Berlin/Heidelberg, Germnay, 2013.
25. Bensadon, J. Black-box optimization using geodesics in statistical manifolds. **2013**, arXiv:1309.7168.
26. Bennequin, D. Dualités de champs et de cordes*.* *Séminaire N. Bourbaki* **2003**, *899*, 117–148. (In French)
27. Bennequin, D. Dualité Physique-Géométrie et Arithmétique. Avalable online: http://archive.numdam.org/ARCHIVE/SB/SB_2001-2002__44_/SB_2001-2002__44__117_0/SB_2001-2002__44__117_0.pdf (accessed on 16 June 2014).
28. Chasles, M. Aperçu Historique sur L'origine et le Développement des Méthodes en Géométrie; Gauthier-Villars: Paris, France, 1837. (In French)
29. Gergonne, J.D. Polémique mathématique. Réclamation de M. le capitaine Poncelet (extraite du bulletin universel des annonces et nouvelles scientifiques); avec des notes. *Annales de Mathématiques Pures et Appliquées* **1827–1828**, *18*, 125. (In French)
30. Poncelet, J.V. *Traité des propriétés projectives des figures*; Gauthier-Villars: Paris, France, 1822. (In French)
31. André, Y. Dualités, Sixième séance, ENS, Séminaire MaxMux. In *Leçons de Mathématiques contemporaines à l'IRCAM*; ENS: Paris, France, 2009; Chap.6. (In French)
32. Atiyah, M.F. Duality in mathematics and physics. Available online: https://www.fme.upc.edu/arxius/butlleti-digital/riemann/071218_conferencia_atiyah-d_article.pdf (accessed on 20 June 2014).
33. Von Oettingen, A.J. Das duale System der Harmonie (Part 1). *Annalen der Naturphilosophie* **1902**, *1*, 62–75. (In German)

34. Von Oettingen, A.J. Das duale System der Harmonie (Part 2). *Annalen der Naturphilosophie* **1903**, *2*, 375–403. (In German)
35. Von Oettingen, A.J. Das duale System der Harmonie (Part 3). *Annalen der Naturphilosophie* **1904**, *3*, 241–269. (In German)
36. Von Oettingen, A.J. Das duale System der Harmonie (Part 4). *Annalen der Naturphilosophie* **1905**, *4*, 116–152. (In German)
37. Von Oettingen, A.J. Das duale System der Harmonie (Part 5). *Annalen der Naturphilosophie* **1906**, *5* 449–503. (In German)
38. Von Oettingen, A.J. *Das duale Harmoniesystem; C.F.W. Siegel's musikalienhandlung*; R. Linnemann: Leipzig, Germany, 1913. (In German)
39. Moreau, J.J. Fonctions convexes duales et points proximaux dans un espace hilbertien. *Comptes Rendus de l'Académie des Sciences* **1962**, *255*, 2897–2899. (In French)
40. Zia, R.K.P.; Redish, E.F.; McKay, S.R. Making Sense of the Legendre Transform. **2009**, arXiv:0806.1147.
41. Koszul, J.L. Variétés localement plates et convexité. *Osaka. J. Math.* **1965**, *2*, 285–290. (In French)
42. Koszul, J.L. Exposés sur les Espaces Homogènes Symétriques; Publicação da Sociedade de Matematica de São Paulo: São Paulo, Brazil, 1959. (In French)
43. Koszul, J.L. Sur la forme hermitienne canonique des espaces homogènes complexes. *Can. J. Math.* **1955**, *7*, 562–576. (In French)
44. Koszul, J.L. *Lectures on Groups of Transformations*; Tata Institute of Fundamental Research: Bombay, India, 1965.
45. Koszul, J.L. Domaines bornées homogènes et orbites de groupes de transformations affines. *Bull. Soc. Math. Fr.* **1961**, *89*, 515–533. (In French)
46. Koszul, J.L. Ouverts convexes homogènes des espaces affines. *Math. Z.* **1962**, *79*, 254–259. (In French)
47. Koszul, J.L. Déformations des variétés localement plates. *Ann. Inst. Fourier* **1968**, *18*, 103–114. (In French)
48. Vinberg, E.B. Homogeneous convex cones. *Trans. Mosc. Math. Soc.* **1963**, *12*, 340–363.
49. Vinberg, E.B. The Theory of Homogeneous Convex Cones. *Trudy Moskovskogo Ma tematicheskogo Obshchestva* **1963**, *12*, 303–358.
50. Jensen, J.L.W. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Math*. **1906**, *30*, 175–193. (In French)
51. Needham, T. A Visual Explanation of Jensen's Inequality. *Am. Math. Mon*. **1993**, *8*, 768–777.
52. Jaynes, E.T. Information Theory and Statistical Mechanics. *Phys. Rev. Ser. II* **1957**, *106*, 620–630.
53. Jaynes, E.T. Information Theory and Statistical Mechanics II. *Phys. Rev. Ser*. **1957**, *108*, 171–190.
54. Jaynes, E.T. Prior Probabilities. *IEEE Trans. Syst. Sci. Cybern.* **1968**, *4*, 227–241.
55. Dacunha-Castelle, D.; Gamboa, F. Maximum d'entropie et problèmes des moments. *Ann. Inst. H. Poincaré Prob. Stat.* **1990**, *26*, 567–596. (In French)
56. Gamboa, F.; Gassiat, E. Maximum d'entropie et problème des moments: Cas multidimensionnel. *Probab. Math. Stat*. **1991**, *12*, 67–83. (In French)
57. Dacunha-Castelle, D.; Gamboa, F. Maximum de l'entropie sous contraintes non linéaires. *Ann. Inst H. Poincaré Prob. Stat.* **1990**, *4*, 567–596. (In French)

58. Krein, M.G.; Nudelman, A.A. *The Markov Moment Problem And Extremal Problems*; American Mathematical Society: New York, NY, USA, 1977.
59. Souriau, J.M. D*éfinition covariante* des éq*u*ilibres thermodynamiques. *Suppl. Nuov. Cimento* **1966**, *1*, 203–216. (In French)
60. Crouzeix, J.P. A Relationship Between The Second Derivatives of a Convex Function and of Its Conjugate. *Math. Program*. **1977**, *3*, 364–365.
61. Seeger, A. Second Derivative of a Convex Function and of Its Legendre-Fenchel Transformate. *SIAM J. Optim*. **1992**, *2*, 405–424.
62. Hiriart-Urruty, J.B. A new set-valued second-order derivative for convex functions. In *Mathematics for Optimization*; Mathematical Studies Series 129; Elsevier: Amsterdam, The Netherlands, 1986.
63. Rao, C.R. Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* **1945**, *37*, 81–89.
64. Fréchet, M. Sur l'extension de certaines évaluations statistiques au cas de petits échantillons. *Revue de l'Institut International de Statistique* **1943**, *11*, 182–205. (In French)
65. Koszul, J.L. Trajectoires Convexes de Groupes Affines Unimodulaires. In *Essays on Topology and Related Topics*; Springer: Berlin, Germany, 1970; pp. 105–110.
66. Vey, J. Sur les Automorphismes Affines des Ouverts Convexes Saillants. *Annali della Scuola Normale Superiore di Pisa, Classe di Science* **1970**, *24*, 641–665. (In French)
67. Vey, J. Sur une notion d'hyperbolicité des variables localement plates. *Thèse de Troisième Cycle de Mathématiques Pures*; Faculté des sciences de l'université de Grenoble: Grenoble, France, 1969. (In French)
68. Shima, H. Geometry of Hessian Structures. In *Springer Lecture Notes in Computer Science*; Nielsen, F., Frederic, B., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8085, pp. 37–55.
69. Shima, H. The Geometry of Hessian Structures; World Scientific: Singapore, Singapore, 2007.
70. Byande, P.M.; Ngakeu, F.; Nguiffo Boyom, M.; Wolak, R. KV-Cohomology and Differential Geometry of Affinely Flat Manifolds. Information Geometry. *Afr. Diaspora J. Math*. **2012**, *14*, 197–226.
71. Sasaki, T. A Note on Characteristic Functions and Projectively Invariant Metrics on a Bounded Convex Domain. *Tokyo J. Math.* **1985**, *8*, 49–79.
72. Sasaki, T. Hyperbolic Affine Hyperspheres. *Nagoya Math. J.* **1980**, *77*, 107–123.
73. Rothaus, O.S. The Construction of Homogeneous Convex Cones. *Ann. Math.* **1966**, *83*, 358–376.
74. Barbaresco, F. *Information Geometry of Covariance Matrix: Cartan-Siegel Homogeneous Bounded Domains, Mostow/Berger Fibration and Fréchet Median, Matrix Information Geometry*; Bhatia, R., Nielsen, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 199–256.
75. Arnaudon, M.; Barbaresco, F.; Yang, L. Riemannian medians and means with applications to radar signal processing. *IEEE J. Sel. Top. Signal Process*. **2013**, *7*, 595–604.
76. Yang, L. Médianes de Mesures de Probabilité dans les Variétés Riemanniennes et Applications à la Détection de Cibles Radar. Thales Ph.D, Thèse de l'Université de Poitiers, Poitiers, France, 2012. (In French)

77. Barbaresco, F. Information/Contact Geometries and Koszul Entropy, Geometric Science of Information. In *Lecture Notes in Computer Science*; Nielsen, F., Barbaresco F., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8085, pp. 604–611.
78. Cartan, E. *Le rôle de la Théorie des Groupes de Lie Dans L'évolution de la Géométrie Moderne*; International Mathematical Union: Berlin, Germany, 1936. (In French)
79. Souriau, J.M. Thermodynamique et géométrie. In *Differential Geometrical Methods in Mathematical Physics II*; Bleuler, K., Reetz, A., Petry, H.R., Eds.; Springer: Berlin/Heidelberg, Germany, 1978; pp. 369–397.
80. Souriau, J.M. Géométrie de l'espace de phases. *Comm. Math. Phys.* **1966**, *374*, 1–30 (In French)
81. Souriau J.M. On Geometric Mechanics. *Discret. Cont. Dyn. Syst. J.* **2007**, *19*, 595–607.
82. Souriau, J.M. *Structure des systèmes dynamiques*; Editions Jacques Gabay: Paris, France, 1970. (In French)
83. Souriau, J.M. Structure of Dynamical Systems, volume 149 of Progress in Mathematics. In *A Symplectic View of Physics*; Birkhäuser Boston Inc.: Boston, MA, USA, 1997.
84. Souriau, J.M. Thermodynamique Relativiste des Fluides; Centre de Physique Théorique: Marseille, France, 1977. (In French)
85. Souriau, J.M.; Iglesias, P. *Heat Cold and Geometry. Differential Geometry and Mathematical Physics, Mathematical Physics Studies Volume*; Springer: Amsterdam, The Netherlands, 1983; pp. 37–68.
86. Souriau, J.M. Dynamic Systems Structure (chap.16 convexité, chap. 17 Mesures, chap. 18 Etats statistiques, Chap. 19 Thermodynamique), available in Souriau archive (document sent by C. Vallée), unpublished technical notes, 1980.
87. Souriau, J.M. Géométrie *Symplectique et Physique Mathématique*; Éditions du C.N.R.S.: Paris, France, 1975. (In French)
88. Ruelle, D. *Statistical Mechanics. Rigorous Results*; Imperial College Press: London, UK, 1999.
89. Ruelle, D. *Hasard et Chaos*; Editions Odile Jacob: Paris, France, 1991.
90. Ruelle, D. *Thermodynamic Formalism: The Mathematical Structure of Equilibrium Statistical Mechanics*, 2nd ed.; Cambridge Mathematical Library; Cambridge University Press: Cambridge, UK, 2004.
91. Vallée, C. Lois de Comportement des Milieux Continus Dissipatifs Compatibles Avec la Physique Relativiste. Ph.D Thesis, Poitiers University, Poitiers, France, 1987. (In French)
92. Vallée, C. Relativistic thermodynamics of continua. *Int. J. Eng. Sci*. **1981**, *19*, 589–601.
93. De Saxcé, G.; Vallée, C. Bargmann group, momentum tensor and Galilean invariance of Clausius-Duhem inequality. *Int. J. Eng. Sci*. **2012**, *50*, 216–232.
94. Vallée, C.; Hjiaj, M.; Fortuné, D.; de Saxcé, G. Canonical and Anti-Canonical Transformations Preserving Convexity of Potentials. *J. Elast*. **2011**, *103*, 247–267.
95. Iglésias, P. Equilibre statistiques et géométrie symplectique en relativité générale. A*nnales de l'institut Henri Poincaré (A) Physique théorique* **1982**, *36*, 257–270. (In French)
96. Iglésias, P. Essai de thermodynamique rationnelle des milieux continus. A*nnales de l'institut Henri Poincaré (A) Physique théorique*. **1981**, *34*, 1–24. (In French)

97. Dubois, F. Conservation Laws Invariants for Galileo Group; Cemracs Preliminary Results. Available online: http://hal.archives-ouvertes.fr/docs/00/55/53/13/PDF/dubois-cemracs99.janv2011.pdf (accessed on 20 June 2014).
98. Roblin, T. Ergodicité et équidistribution en courbure négative. *Mémoire de la Société Mathématique de France* **2003**, *95*, 96. (In French)
99. Paulin, F.; Pollicott, M.; Schapira, B. Equilibrium states in negative curvature. **2013**, arXiv:1211.6242.
100. Coudène, Y. Gibbs measures on negatively curved manifolds. *J. Dynam. Control Syst.* **2003**, *9*, 89–101.
101. Coudène, Y.; Schapira, B. Generic measures for geodesic flows on nonpositively curved manifolds. *Dyn. Syst.* **2014**, submitted.
102. Haydn, N.T.A.; Ruelle, D. Equivalence of Gibbs and equilibrium states for homeomorphisms satisfying expansiveness and specification. *Comm. Math. Phys.* **1992**, *148*, 155–167.
103. Mohsen, O. Le bas du spectre d'une variété hyperbolique est un point selle. *Ann. Sci. École Norm. Sup.* **2007**, *40*, 191–207. (In French)
104. Donaldson, S.K. Scalar Curvature and Stability of Toric Variety. *J. Differ. Geom*. **2002**, *62*, 289–349.
105. Abreu, M. Kähler geometry of toric varieties and extremal metrics. *Int. J. Math.* **1998**, *9*, 641–651.
106. Guan, D. On modified Mabuchi functional and Mabuchi moduli space of kahler metrics on toric bundles. *Math. Res. Lett*. **1999**, *6*, 547–555.
107. Guillemin, V. Kaehler structures on toric varieties. *J. Differ. Geom*. **1994**, *40*, 285–309.
108. Guillemin, V. *Moment Maps and Combinatorial Invariants of Hamiltonian Tn-Spaces*; Birkhau-ser: Basel, Switzerland, 1994.
109. Atiyah, M.; Bott, R. The moment map and equivariant cohomology. *Topology* **1984**, *23*, 1–28.
110. Bruguières, A. Propriétés de convexité de l'application moment. *Séminaire Bourbaki* **1985**, *28*, 63–87. (In French)
111. Condevaux, M.; Dazord, P.; Molino, P. *Géométrie du moment. Dans Travaux du séminaire Sud Rhodanien de Géométrie*; University Lyon: Lyon, France, 1988; pp. 131–160.
112. Delzant, T. Hamiltoniens périodiques et images convexes de l'application moment. *Bull. Soc. Math. Fr.* **1988**, *116*, 315–339. (In French)
113. Guillemin, V.; Sternberg, S. Convexity properties of the moment mapping. *Inv. Math*. **1982**, *67*, 491–513.
114. Kirwan, F. Convexity properties of the moment mapping. *Inv.Math*. **1984**, *77*, 547–552.
115. Kapranov, M. Thermodynamics and the moment map. **2011**, arXiv:1108.3472v1.
116. Biquard, O. *Métriques Kählériennes Extrémales sur les Surfaces Toriques*; Société Mathématique de France: Paris, France, **2011**; Volume 1018. (In French)
117. Pavlov, V.P.; Sergeev, V.M. Thermodynamics from the Differential Geometry Standpoint. *Theor. Math. Phys*. **2008**, *157*, 1484–1490.
118. Kozlov, V.V. Heat Equilibrium by Gibbs and Poincaré; *Dokl. Ross. Akad. Nauk* **2002**, *382*, 602–606. (In French)
119. Berezin, F.A. *Lectures on Statistical Physics*; World Scientific: Singapore, Singapore, 2007.
120. Poincaré, H. Réflexions sur la théorie cinétique des gaz. *J. Phys. Theor. Appl.* **1906**, *5*, 369–403.

121. Carathéodory, C. Untersuchungen über die Grundlagen der Thermodynamik. *Math. Ann*. **1909**, *67*, 355–386. (In German)
122. Souriau, J.M. Mécanique Classique et Géométrie Symplectique; Report ref. CPT-84/PE-1695 CNRS Centre de Physique Théorique: Marseille, France, 1984. (In French)
123. Viterbo, C. Generating Functions, Symplectic Geometry and Applications. *Proc. Intern. Congr. Math., Zurich* **1994**, *1*, 537–547.
124. Viterbo, C. Symplectic topology as the geometry of generating functions. *Math. Ann*. **1992**, *292*, 685–710.
125. Hörmander, L. Fourier integral operators I. *Acta Math.* **1971**, *127*, 79–183.
126. Théret, D. A complete proof of Viterbo's uniqueness theorem on generating functions. *Topol. Appl.* **1999**, *96*, 249–266.
127. Chentsov, N.N. Statistical Decision Rules and Optimal Inferences (Transactions of Mathematics Monograph); American Mathematical Society: Providence, RI, USA, 1982; Volume 53.
128. Villani, C. (Ir)rréversibilité et entropie/(Ir)reversibility and entropy. In *Séminaire Poincaré Le temps*; École Polytechnique: Paris, France, 2010. (In French)
129. Ollivier, Y. Aspects de l'entropie en mathématiques et en physique. Available online: http://www.yann-ollivier.org/entropie/entropie.pdf (accessed on 1 January 2014). (In French)
130. Avantaggiati, A.; Loreti, P. On Fourier Transform, Parseval Equality, and the Inversion Formula in Idempotent Analysis. In Proceedings of the 2013 European Control Conference (ECC), Zürich, Switzerland, 17–19 July 2013.
131. Litvinov, G.L. The Maslov Dequantization, Idempotent and Tropical Mathematics: A brief Introduction. *J. Math. Sci.* **2007**, *140*, 426–444.
132. 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.). *Bull. Soc. Math. Fr.* **1962**, *90*, 39–156. (In French)
133. Leray, J. Le calcul différentiel et intégral sur une variété analytique complexe Problème de Cauchy, III. *Bull. Soc. Math. Fr*. **1959**, *82*, 6–180. (In French)
134. Maslov, V.P. *Operational Methods*; MIR: Moscow, Russia, 1976.
135. Del Moral, P.; Doisy, M. Maslov idempotent probability calculus, I, II. *Theory Probab. Appl.* **2000**, *44*, 319–332.
136. Del Moral, P.; Doisy, M. On the applications of Maslov optimization theory. *Math. Notes* **2001**, *69*, 232–244.
137. Laplace P.S. Mémoire sur la probabilité des causes sur les évènements. In *Mémoires de Mathématique et de Physique*; De l'Imprimerie Royale: Paris, France, 1774. (In French)
138. Legendre, A.M. *Mémoire Sur L'intégration de Quelques Equations aux Différences Partielles*; Mémoires de l'Académie des Sciences: Paris, France, 1787; pp. 309–351. (In French)
139. Monge, G. *Sur le Calcul Intégral des Equations aux Différences Partielles*; Mémoires de l'Académie des Sciences: Paris, France, 1784; pp. 118–192. (In French)
140. Darboux, G. *Leçons sur la Théorie Générale des Surfaces et les Applications Géométriques du Calcul Infinitésimal: Premiere Partie (Généralités, Coordonnées Curvilignes, Surface Minima)*; Gauthier-Villars: Paris, France, 1887. (In French)

141. Hadamard, J. Les surfaces à courbures opposées et leurs lignes géodésiques. *Journal de Mathématiques Pures et Appliquées* **1898**, *4*, 27–74. (In French)
142. Vesentini, E. *Geometry of Homogeneous Bounded Domains*; Springer-Verlag: Berlin/Heidelberg, Germany, 2011.
143. Bourdon, M. Structure conforme au bord et flot géodésique d'un CAT(−1)-espace. *L'Enseignement Mathématique* **1995**, *41*, 63–102. (In French)
144. Deza, E.; Deza, M.M. Dictionary of Distances; Elsevier: Amsterdam, The Netherlands, 2006.
145. Cartan, E. *Leçons sur les Invariants Intégraux*; Hermann: Paris, France, 1922. (In French)
146. Libermann, P.; Marle, C.M. *Symplectic Geometry and Analytical Mechanics*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1987.