Lie Group Statistics and Lie Group Machine Learning Based on Souriau Lie Groups Thermodynamics & Koszul-Souriau-Fisher Metric: New Entropy Definition as Generalized Casimir Invariant Function in Coadjoint Representation

Frédéric Barbaresco

Key Technology Domain PCC (Processing, Control & Cognition) Representative, Thales Land & Air Systems, Voie Pierre-Gilles de Gennes, F91470 Limours, France; frederic.barbaresco@thalesgroup.com

Abstract: In 1969, Jean-Marie Souriau introduced a "Lie Groups Thermodynamics" in Statistical Mechanics in the framework of Geometric Mechanics. This Souriau's model considers the statistical mechanics of dynamic systems in their "space of evolution" associated to a homogeneous symplectic manifold by a Lagrange 2-form, and defines in case of non null cohomology (non equivariance of the coadjoint action on the moment map with appearance of an additional cocyle) a Gibbs density (of maximum entropy) that is covariant under the action of dynamic groups of physics (e.g., Galileo's group in classical physics). Souriau Lie Group Thermodynamics was also addressed 30 years after Souriau by R.F. Streater in the framework of Quantum Physics by Information Geometry for some Lie algebras, but only in the case of null cohomology. Souriau method could then be applied on Lie groups to define a covariant maximum entropy density by Kirillov representation theory. We will illustrate this method for homogeneous Siegel domains and more especially for Poincaré unit disk by considering SU(1,1) group coadjoint orbit and by using its Souriau's moment map. For this case, the coadjoint action on moment map is equivariant. For non-null cohomology, we give the case of Lie group SE(2). Finally, we will propose a new geometric definition of Entropy that could be built as a generalized Casimir invariant function in coadjoint representation, and Massieu characteristic function, dual of Entropy by Legendre transform, as a generalized Casimir invariant function in adjoint representation, where Souriau cocycle is a measure of the lack of equivariance of the moment mapping.

Keywords: Lie groups thermodynamics; Lie group machine learning; Kirillov representation theory; coadjoint orbits; moment map; covariant Gibbs density; maximum entropy density; Souriau-Fisher metric; generalized Casimir invariant function

---

La thèse de Kirillov, parue en 1962, a suscité immédiatement beaucoup d'intérêt ... En outre, quantité de notions naturelles concernant les représentations s'interprètent géométriquement en terme d'orbites coadjointes: restriction à un sous-groupe, induction unitaire, produit tensoriel, mesure de Plancherel, la topologie de l'ensemble représentations unitaires irréductibles ... Kirillov s'est vite convaincu, et il a convaincu la communauté mathématique que cette « méthode des orbites » devait être applicable à des groupes bien plus généraux que les groupes nilpotents. Il n'a pas hésité à aborder le cas des groupes de Lie connexes quelconques. Evidemment, des difficultés considérables ont surgi immédiatement. Néanmoins, Kirillov a indiqué une voie d'accès, qui ensuite a été largement utilisée. - **Jacques Dixmier, Brèves remarques sur l'œuvre de A.A. Kirillov**

On comprend ainsi comment Lagrange a pu développer les lois de la Mécanique des systèmes formés de solides sans s'occuper des variations de la température de ces corps et Fourier traiter des variations de la température de ces mêmes corps solides sans s'occuper de leur mouvement; comment on peut étudier le mouvement de la Terre, assimilée à un solide rigide, sans se préoccuper de la température de cet astre et étudier le refroidissement du globe terrestre sans se préoccuper de son mouvement. Une telle indépendance entre les problèmes qui ressortissent à la Mécanique et les problèmes qui ressortissent à la Théorie de la chaleur n'existe plus lorsque les systèmes auxquels on a affaire ne sont plus des systèmes classiques; si, par exemple, au lieu de regarder la Terre comme un solide rigide, d'état invariable, on tient compte des changements de volume, de forme, d'état physique et chimique qui accompagnent son refroidissement, on ne peut plus séparer le problème du mouvement de la Terre et le problème du refroidissement terrestre. . . . On sait que cette forme de relations supplémentaires avait été introduite par Newton et les géomètres du XVIIIème siècle dans la théorie du son. Ces considérations montrent que les questions qui ressortissent à la Thermodynamique ont dû solliciter l'attention des physiciens dès qu'on a voulu aborder l'étude des systèmes autres que des systèmes classiques; et, en fait, c'est la théorie de la propagation du son dans l'air qui a provoqué Laplace à créer la Thermodynamique.
- **P. Duhem, L'intégrale des forces vives en thermodynamique, JMPA 4:5-19, 1898 [1–4]**

Sous cette aspiration, la physique qui était d'abord une science des "agents" doit devenir une science des "milieux". C'est en s'adressant à des milieux nouveaux que l'on peut espérer pousser la diversification et l'analyse des phénomènes jusqu'à en provoquer la géométrisation fine et complexe, vraiment intrinsèque . . . Sans doute, la réalité ne nous a pas encore livré tous ses modèles, mais nous savons déjà qu'elle ne peut en posséder un plus grand nombre que celui qui lui est assigné par la théorie mathématique des groupes - **Gaston Bachelard, Etude sur l'Evolution d'un problème de Physique –La propagation thermique dans les solides, 1928**

1. Introduction

The previous French quotes by the Mathematician Jacques Dixmier, the Physicist Pierre Duhem, and the Philosopher Gaston Bachelard are important to introduce the epistemological context of models that will be developed in the paper. Jacques Diximer refers to Alexander Kirillov seminal idea of coadjoint orbits method to consider Lie group representation model. Pierre Duhem makes comments to the origin of the gap between the theory of heat and the theory of Mechanics. Finally, Gaston Bachelard make prediction that new Thermodynamics foundations will be given by groups. We will try in this paper, to prove that these ideas could be reconciled by the Souriau model of Lie groups Thermodynamics through the mathematical structure of Lie algebra cohomology.

After a the state of the art and trends in Machine Learning based on Information Geometry, we will present, in this introduction, the main objective of this paper to jointly apply models from geometric statistical mechanics and tools from Information geometry to solve "Gauss density" definition problem for statistics on Lie groups and homogeneous manifolds. We will also present use-cases motivation for Lie group machine learning illustrating for Doppler statistics analysis with SU(1,1) statistics, and for kinematics data analysis with SE(2) statistics.

1.1. State of the Art and Trends in Machine Learning Based on Information Geometry

The classical simple gradient descent used in Deep Learning has two drawbacks: the use of the same non-adaptive learning rate for all parameter components, and a non-invariance with respect to parameter re-encoding inducing different learning rates. As the parameter space of multilayer networks forms a Riemannian space equipped with Fisher information metric, instead of the usual gradient descent method, the natural gradient or Riemannian gradient method, which takes account of the geometric structure of the Riemannian space, is more effective for learning. The natural gradient preserves this invariance to be insensitive to the characteristic scale of each parameter direction. The Fisher metric defines a Riemannian metric as the Hessian of two dual potential functions (the Entropy and the log-partition function). Yann Ollivier and Gaétan Marceau-Caron provided in 2016 [5] the first

experimental results on non-synthetic data sets for the quasi-diagonal Riemannian Natural gradient descents for neural networks introduced previously by Yann Ollivier in [6] (MNIST, SVHN, and FACE data sets). The quasi-diagonal Riemannian algorithms consistently beat simple stochastic gradient gradient descents by a varying margin. The computational overhead with respect to simple backpropagation is around a factor 2, and reach their final performance quickly, thus requiring fewer training epochs and a smaller total computation time. The main goal of natural gradient is to obtain invariance properties, such as, for a neural network, insensitivity of the training algorithm to whether a logistic or tanh activation function is used, or insensitivity to simple changes of variables in the parameters, such as scaling some parameters. In 2017, same authors have introduced the resulting natural Langevin dynamics [7] combining the advantages of natural gradient descent and Fisher-preconditioned Langevin dynamics for large neural networks, validated on MNIST with Fisher matrix preconditioning. With all invariance properties of natural gradient, this Langevin Dynamics avoids overfitting as a regularization method, and replaces classical methods based on a controlled amount of noise to stochastic gradient descents, that ensures convergence to the Bayesian posterior on model parameters. The theoretically optimal covariance of the noise is the inverse Fisher metric, and Y. Ollivier and G. Marceau-Caron have shown how to implement this in practice with neural networks using efficient Fisher metric approximations. In 2017, Yann Ollivier has also introduced TANGO algorithm (True Asymptotic Natural Gradient Optimization) [8], which converges to a true natural gradient descent in the limit of small learning rates, without explicit Fisher matrix estimation, and where in large dimension, small learning rates will be required to approximate the natural gradient well. Y. Ollivier has also shown that it is possible to get arbitrarily close to exact natural gradient descent with a lightweight algorithm. About natural gradient for Deep Learning, we can refer to [9,10]. This year, Shun-ichi Amari [11] has given an elementary geometrical proof that any target function is realized in a sufficiently small neighborhood of any randomly connected deep network, provided the width (the number of neurons in a layer) is sufficiently large.

In this paper, we will introduce how to extend these approaches for data as elements of Lie groups or data lying on a homogeneous manifold where a Lie group acts transitively. This extension is considered in the framework and interconnexion of Souriau "Lie groups Thermodynamics", Information Geometry and Kirillov representation theory [12] to define probability densities for Lie groups, as Souriau covariant Gibbs densities (density of Maximum of Entropy). We will develop this case for the matrix Lie group SU(1,1) (case with null cohomology) through the computation of Souriau's moment map, and Kirillov's orbit method. We will also develop the method for SE(2) Lie group (case with non-null cohomology) where a Souriau cocycle should be taken into account due to the defect of equivariance of the coadjoint action on the moment map.

Supervised learning approaches are based on neural networks whose parameters are estimated by natural gradient algorithms. Non-supervised algorithm are based on clustering by using technics called "k-means" or "Mean-shift" using distance between elements of the dataset. In both cases, if we want to extend these approaches for Lie groups dataset, we have to extend the notion of Gaussian densities and distance between elements. We propose to use Geometric Statistical Model coming from Geometric Statistical Mechanics to introduce "Gauss density" of Lie group elements. Jointly, we can associate a natural distance between these Lie group elements on the Symplectic manifold by means of KKS 2-form, introducing a natural Riemannian metric associated to Fisher Metric from Information Geometry. The objective of this paper is to explain how to use Geometric Statistical Mechanics tools in this context.

1.2. Objectives of this Paper

The purpose of this article is multiple. The work of Professor Jean-Marie Souriau is well known in the field of "Geometric Mechanics" of which he is one of the founders with his book "structure of dynamic systems" published in 1969, and in which he introduced the foundations of Symplectic Geometry. Inside this book, chapter IV dealing with the extension of Geometric Mechanics to Statistical

Mechanics, has been little read or misunderstood by this community. We have discovered that this model was part of and generalized another discipline, which is called Information Geometry. We have demonstrated in other previous articles that one could generalize Fisher metric (invariant metric used in Information Geometry) for Lie groups, with this model. It is therefore a question of rehabilitating the work of Jean-Marie Souriau in a broader framework, which concerns statistics and machine learning extended to objects considered as elements of a Lie group or a homogeneous manifold.

The second goal is to solve with these new tools problems that were still unsolved in statistics and machine learning. These unresolved problems concern the definition and calculation of the expression of probability densities, playing the role of Gaussian density, for elements of a Lie group or elements of homogeneous manifolds. In this article, we completely solve the problem for 2 Lie groups very useful in machine learning but also in physics, the Lie groups SU(1,1) and SE(2). The calculation is not a simple application of the Souriau model, because it is necessary to establish the "moment map" associated with these groups and define a Laplace transform on their coadjoint orbits of these groups (action of the group on the dual space of Lie algebra). In a second step, we must use Information Geometry to write these covariant Gibbs densities in the correct parametrization which parametrizes the generalized Gaussian law from statistical moments on the homogeneous symplectic manifold associated with coadjoint orbits. In the case SU(1,1), which corresponds to a case of null cohomology (equivariance of the coadjoint operator on the moment application), as the homogeneous symplectic manifold to the coadjoint orbit is the Poincaré unit disk, we solve jointly, an open problem to define mathematically the notion of Gaussian density in this disk in hyperbolic geometry. With the property that this density is by construction invariant under the action of the group SU(1,1), which is the condition sine qua none to preserve the symmetries and the invariance of the associated Fisher metric. We show that this model achieves a breakthrough in machine learning, because we have a Gibbs density and a Fisher metric invariant by change of parametrization and invariant under the action of symmetries. Gibbs density allows us to extend the classical supervised statistical machine learning algorithms, and Fisher metric allows us to adress unsupervised learning problem as k-means problems in metric space. The model opens the way to machine learning for Lie groups with multiple applications in robotics, sensor signal processing, image processing.

In the last part of the article, based on this model, we give a new "geometric" definition of Entropy by showing that Entropy is an invariant Casimir function in coadjoint representation. The Casimir functions have been widely studied within the framework of Poisson structures and manifolds [13–16]. This characterization of Entropy is new, because previously Entropy was defined axiomatically. Using this Casimir function property, we show that it is possible to use full geometric approaches to construct the Entropy function only from the structure coefficients of the Lie group associated with the symmetries involved. We show that we can also introduce an Euler-Poincaré equation and its stochastic variant to study other open problems in statistics and thermodynamics. The application of this Casimir characterization, which is demonstrated in this article, are developed in another twin article published in the same special issue with François Gay-Balmaz [17].

1.3. Motivation of Lie Group Machine Learning with Use-Cases

Machine learning is a field of study of artificial intelligence, which is based on statistical approaches to give computers the ability to "learn" from data, that is, to classify data from observations in a supervised or non-supervised way. Machine learning generally has two phases. The first consists in estimating a model from data, called observations. This so-called "training" phase is generally carried out before the practical use of the model. The second phase corresponds to the start of production: the model being determined, new data can then be classified. According to the information available during the learning phase, learning is qualified in different ways. If the data is labeled (that is, the task response is known for that data), it is supervised learning. We speak of classification if the labels are discrete, or of regression if they are continuous. In the most general case, without a label, we seek to determine the underlying structure of the data (which can be a probability density) and it is then

unsupervised learning. Machine learning can be applied to different types of data, such as graphs, trees, curves, or more simply feature vectors, which can be continuous or discrete. We propose to extend the approach, when datasets are element of matrix lie groups.

Learning algorithms can be categorized according to their learning mode. For supervised learning, the classes are predetermined and the examples known, and then the system learns to classify according to a classification model. An expert must label examples beforehand. The process takes place in two phases. For unsupervised learning, when the system or operator has only examples, but no label, and the number of classes and their nature have not been predetermined, we speak of unsupervised learning or clustering. No expert is required. The algorithm must discover for itself the more or less hidden structure of the data, by data partitioning and data clustering. The system must cluster the data according to their available attributes, to classify them into homogeneous groups of examples. Similarity is generally calculated according to a distance function between pairs of examples.

We will illustrate two problems of Machine Learning on Lie groups coming from Radar Industry. Target recognition on Radar micro-Doppler data could be modeled by a problem of classification of dataset considered as elements of SU(1,1) Lie group (see Figure 1). Radar complex time series of micro-Doppler observation of data are classically processed on sliding time window to estimate their associated covariance matrices that are characterized by a Toeplitz Hermitian Positive-definitiveness structure. Using a well-known Verblunsky/Trench Theorem, we can parametrize all Toeplitz Hermitian Positive Definite Covariance matrices of stationary Radar Time series in a product space with a real positive axis (for signal power) and a Poincaré polydisk (for Doppler Spectrum shape). If we consider the Poincaré Unit Disk as an homogeneous space where SU(1,1) Lie group acts transitively. Each data in Poincaré unit disk of this polydisk could be then coded by SU(1,1) matrix Lie group element. We have transformed the problem into a statistical learning challenge processing data of SU(1,1) matrix Lie group. Another exemple considers flying object recognition on their kinematics coded in SE(2) or SE(3) Lie Groups. 3D (or 2D) trajectories could be coded by SE(3) (or SE(2)) Lie group time series provided through Invariant Extended Kalman Filter (IEKF) Radar Tracker, that locally estimates displacement of Frenet-Seret frame. Object kinematics will be then coded by time series of SE(3) (or SE(2)) matrix Lie groups characterizing local rotation/translation of Frenet frame along the drone 3D (or 2D) trajectory. Statistics of this SE(3) (or SE(2)) Lie group elements will characterize flight mechanics of different kinds of object (birds, drones, . . .).

SU(1,1) or SE(2) are also fundamental tools in Image Processing (Sub-Riemannian Geometry of vision with SE(2)), in robotics (rigid bodies statistical analysis with SE(2)), in Natural Langage Processing (methods of graph-embedding in Poincaré disk with SU(1,1)), ... For instance, SU(1,1) Lie group which acts on Poincaré unit Disk is highly studied to embed isometrically a graph in an hyperbolic space. It is used by GAFAM (Google, Facebook, . . .) for Natural Language Processing by reducing graph analysis to a Machine Learning problem in Hyperbolic Poincaré Unit Disk. Hyperbolic Neural network [18] have been developed in this framework. SU(1,1) Lie group is also fundamental in Quantum physics to describe Coherent states of an electron in a magnetic field for instance [19] and Coherent states in Quantum Optics [20] (some statistical photon-counting aspects of SU(1,1) coherent states are emphasized). SE(2) Lie group is especially fundamental for Geometry of Vision considering sub-Riemannian approaches of the Citti-Petitot-Sarti Model [21] but also also in neuroimagery [22].

1.3.1. SU(1,1) Lie Group Machine Learning for Doppler Data Statistics Analysis

Lie group structure appears naturally on Doppler data, if we consider time series of locally stationary signal and their associated covariance matrix. Covariance matrix is Toeplitz Hermitian Positive Definite. Based on Theorem due to Verblunsky [23,24] and Trench [25], we can parametrize Hermitian Positive Definite Matrix in product space involving the Poincaré unit Polydisk:

$$\varphi : THDP(n) \rightarrow R_+^* \times D^{n-1}$$
$$R_n \mapsto (P_0, \mu_1, \ldots, \mu_{n-1}) \tag{1}$$

where D is the Poincaré Unit Disk:

$$D = \{z = x + iy \in C / |z| < 1\} \tag{2}$$

The Poincaré unit disk is an homogeneous bounded domain where the Lie group SU(1,1) act transitively. This Matrix Group is given by:

$$SU(1,1) = \left\{ \begin{bmatrix} a & b \\ b^* & a^* \end{bmatrix} / |a|^2 - |b|^2 = 1,\ a, b \in C \right\} \tag{3}$$

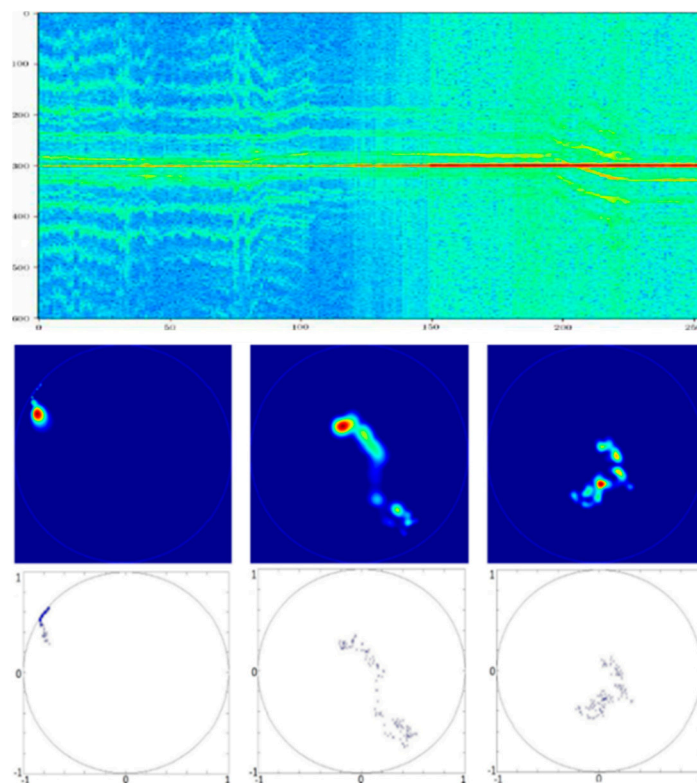where SU(1,1) acts on the Poincaré Unit Disk by:

$$g \in SU(1,1) \Rightarrow g.z = \frac{az + b}{b^*z + a} \tag{4}$$

with Cartan Decomposition of SU(1,1)

$$\begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} = |a| \begin{pmatrix} 1 & z \\ z^* & 1 \end{pmatrix} \begin{pmatrix} a/|a| & 0 \\ 0 & a^*/|a| \end{pmatrix} \tag{5}$$

$$\text{with } z = b(a^*)^{-1}, |a| = \left(1 - |z|^2\right)^{-1/2}$$

Figure 1. Illustration of Statistics of Doppler Spectrum as densities of points in Poincaré polydisc where SU(1,1) Lie group act transitively. Top is a time-doppler frequency spectrum of a drone. Bottom is a Doppler spectrum fluctuation coded as densities in Poincaré Polydisk.

We can observe that $z = b(a^*)^{-1}$ could be considered as action of $g \in SU(1,1)$ on the centre on the unit disk $z = g.0 = b(a^*)^{-1}$. The principal idea is that we can code any point $z = b(a^*)^{-1}$ in the unit disk by an element of the Lie group SU(1,1). Main advantage is that the point position is no

longer coded by coordinates but intrinsically by transformation from the orogin 0 to this point. Finally, a covariance matrix of a stationary signal could be coded by $(n-1)$ Matrix SU(1,1) Lie group elements:

$$\text{THPD} \rightarrow R_+^* \times D^{n-1} \rightarrow R_+^* \times SU(1,1)^{n-1}$$

$$R_n \mapsto (P_0, \mu_1, \ldots, \mu_{n-1}) \mapsto \left(P_0, \begin{bmatrix} a_1 & b_1 \\ b_1^* & a_1^* \end{bmatrix}, \ldots, \begin{bmatrix} a_{n-1} & b_{n-1} \\ b_{n-1}^* & a_{n-1}^* \end{bmatrix} \right) \tag{6}$$

1.3.2. SE(2) and SE(3) Lie Groups Machine Learning for Kinematics Data Statistics Analysis

When we consider a 3D trajectory of a mobile target, we can describe this curve by a time evolution of the local Frenet-Serret frame (local frame with tangent vector, normal vector and binormal vector) as illustrated in Figure 2. This frame evolution is described by the Frenet-Serret formula that gives the kinematic properties of the target moving along the continuous, differentiable curve in 3D Euclidean space $\mathbb{R}^3$. More specifically, the formulas describe the derivatives of the so-called tangent, normal, and binormal unit vectors in terms of each other.

$$\frac{d}{dt} \begin{pmatrix} \vec{t} \\ \vec{n} \\ \vec{b} \end{pmatrix} = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \gamma \\ 0 & -\gamma & 0 \end{bmatrix} \begin{pmatrix} \vec{t} \\ \vec{n} \\ \vec{b} \end{pmatrix} \text{ with } \begin{cases} \kappa : \text{curvature} \\ \gamma : \text{torsion} \end{cases} \tag{7}$$
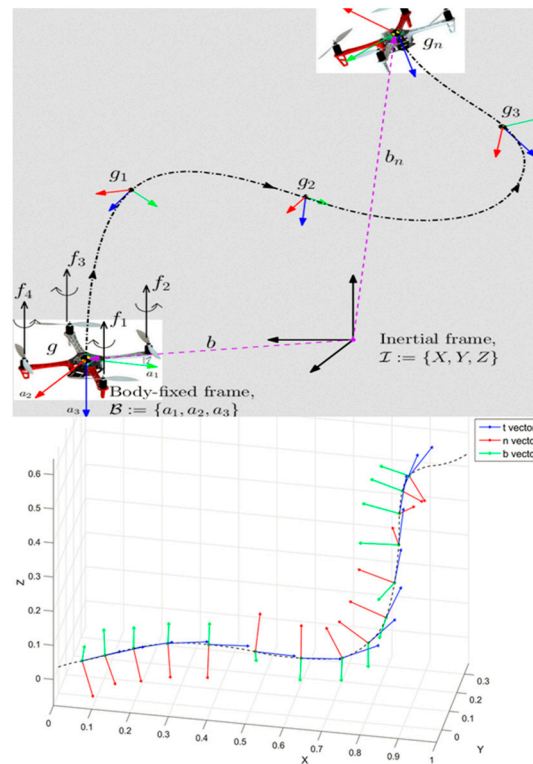
Figure 2. Illustration of times series of local Frenet-Serret Frame where SE(3) Lie group act between two successive frames.

We will consider motions determined by exponentials of paths in the Lie algebra. Such a motion is determined by a unit speed space-curve $\tau(t)$. Now in a Frenet-Serret motion a point in the moving body moves along the curve and the coordinate frame in the moving body remains aligned with the

tangent $\vec{t}$, normal $\vec{n}$, and binormal $\vec{b}$, of the curve. Using the 4-dimensional representation of the Lie group SE(3), the motion can be specified as:

$$G(t) = \begin{pmatrix} R(t) & \tau(t) \\ 0 & 1 \end{pmatrix} \in SE(3) \tag{8}$$

where $\tau(t)$ is the curve and the rotation matrix $R(t)$ has the unit vectors $\vec{t}$, $\vec{n}$, and $\vec{b}$ as columns:

$$R(t) = \begin{pmatrix} \vec{t} & \vec{n} & \vec{b} \end{pmatrix} \in SO(3) \tag{9}$$

If we introduce the Darboux vector $\vec{\omega} = \gamma \vec{t} + \kappa \vec{b}$ that we can rewritte from Frenet-Serret Formulas:

$$\frac{d\vec{t}}{dt} = \vec{\omega} \times \vec{t} \, , \frac{d\vec{n}}{dt} = \vec{\omega} \times \vec{n} \, , \frac{d\vec{b}}{dt} = \vec{\omega} \times \vec{b} \tag{10}$$

Then, we can write with Ω is the 3×3 anti-symmetric matrix corresponding to $\vec{\omega}$:

$$\frac{dR}{dt} = \Omega R \tag{11}$$

We note that $\frac{d\tau(t)}{dt} = \vec{t}$ and $\frac{d\vec{\omega}}{dt} = \frac{d\gamma}{dt}\vec{t} + \frac{d\kappa}{dt}\vec{b}$.

The instantaneous twist of the motion $G(t)$ is given by:

$$S_d = \frac{dG(t)}{dt}G^{-1}(t) = \begin{pmatrix} \Omega & v \\ 0 & 0 \end{pmatrix} \tag{12}$$

This is the Lie algebra element corresponding to the tangent vector to the curve $G(t)$. It is well known that elements of the Lie algebra $se(3)$ can be described as lines with a pitch. The fixed axode of a motion $G(t) \in SE(3)$ is given by the axis of S_d as t varies. The instantaneous twist in the moving reference frame is given by $S_b = G^{-1}(t)S_dG(t)$, that is, by the adjoint action on the twist in the fixed frame. The instantaneous twist S_b can also be found from the relation:

$$S_b = G^{-1}(t)\frac{dG(t)}{dt} \tag{13}$$

$$S_b = G^{-1}\frac{dG}{dt} = \begin{pmatrix} R^T & -R^T\tau \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \Omega R & \vec{t} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} R^T\Omega R & R\vec{t} \\ 0 & 0 \end{pmatrix} \tag{14}$$

We can observe that we could describe a 3D trajectory by a time series of *SE(3)* Lie group elements:

$$SE(3) = \left\{ \begin{bmatrix} R & \tau \\ 0 & 1 \end{bmatrix} / R \in SO(3), \tau \in R^3 \right\} \tag{15}$$

with

$$SO(3) = \left\{ R / R^T R = RR^T = I, \det{}^2 R = 1 \right\} \tag{16}$$

Then, the trajectory will be given by the following time series of *SE(3)* elements:

$$\left\{ \begin{bmatrix} R_1 & \tau_1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} R_2 & \tau_2 \\ 0 & 1 \end{bmatrix}, \cdots, \begin{bmatrix} R_n & \tau_n \\ 0 & 1 \end{bmatrix} \right\} \in SE(3)^n \tag{17}$$

2. New Results Introduced in the Paper

The paper is structured in two parts:

- *1st Part on "Gauss Density on Lie groups"*: This part is totally new in Machine learning with an extension of "Gauss densities" (defined as Maximum Entropy model) for Lie groups coupling both Souriau model (introduced in statistical physics domain), with Information Geometry in Geometric Machine Learning domain. We illustrate with two use-cases SU(1,1) and SE(2) that are the most useful Lie groups in Image Processing (Sub-Riemannian Geometry of vision with SE(2)), in robotics (rigid bodies statistical analysis with SE(2)), in Natural Langage Processing (SU(1,1) with methods of graph-embedding in Poincaré disk), ... Some tentatives have been developed to define noise on Lie groups by adding additional Gaussian components on elements of the Lie algebra [26–29], but these models are not mathematically correct because they do not preserve the symmetries and the moment map associated to these symmetries by the Noether Theorem.

- *2nd part on "Entropy definition extension as Casimir Function"*: This part gives a new geometric definition of Entropy as invariant Casimir function in coadjoint representation, explaining the invariance of entropy under the affine coadjoint action on moment map in the dual space of Lie algebra. This definition was not in the paper of Souriau. With this new definition, we can compute Entropy only by structure constraints given by the Lie group. It opens the door to new generalization of Maximum Entropy method and first of all computation of "Gaussian densities" for any Lie group. Applications of this new property is not developed in this paper but in a twin paper in the same special issue [17]. We refert to M. Gromov papers to consider more geometric structures of Entropy [30,31].

The main new results of this paper are the introduction of "Gauss density" for Lie groups or data on homogeneous space where a Lie groups acts transitively, and the full computation for SU(1,1) Lie group. This group acts transitively on the Poincaré unit disk, and so we have also solved an open problem related to Gauss density on this homogeneous space. For this purpose, the main approach has considered an extended definition of classical "Gauss density", as introduced by Jaynes, in term of density of Maximum Entropy. In this way, the initial problem was transfert to a new one related to the good definition of Entropy for Lie groups. To address this problem, first, we have recalled the classical Euclidean case, where the Entropy $S(\eta)$ could be defined as the Legendre transform of minus the log-partition function $\Phi(\theta) = -\log \int_R e^{-\langle \theta, y \rangle} dy$ (defined by Laplace transform) by the following equation $S(\eta) = \langle \theta, \eta \rangle - \Phi(\theta)$ with $\eta_i = \frac{\partial \Phi(\theta)}{\partial \theta_i}$ and $\theta_i = \frac{\partial S(\eta)}{\partial \eta_i}$. The next step was to explain how to extend the log-partition function for Lie groups. We have then considered the Laplace transform in the framework of Lie group representation theory as introduced by Alexander Kirillov and Geometric Statical Mechanics as modeled by Jean-Marie Souriau. We have preserved the same Legendre structure, and have defined the Entropy $S(Q)$, parametrized on the dual space of the Lie algebra $Q \in \mathfrak{g}^*$ (called geometric heat), as Legendre transform of minus of the log-partition function $\Phi(\beta) = -\log \int_M e^{-\langle J(\xi), \beta \rangle} d\lambda_\omega$, parametrized on the Lie algebra by $\beta \in \mathfrak{g}$ (called geometric Planck Temperature), from a Laplace transform defined on the homogeneous symplectic manifold (associated to the Lie group by the Kirrilov-Kostant-Souriau 2-form called KKS 2-form in the litterature). By introducing the moment map $J : M \to \mathfrak{g}^*$, fundamental tool of representation theory introduced by Souriau, we were able to define the log-partition function on the coadjoint orbit of the Lie group, $\Phi(\beta) = -\log \int_{\mathfrak{g}^*} e^{-\langle J(\xi), \beta \rangle} d\lambda_\omega$. The entropy is then given by the Legendre transform $S(Q) = \langle Q, \beta \rangle - \Phi(\beta)$ with $Q = \frac{\partial \Phi(\beta)}{\partial \beta} \in \mathfrak{g}^*$ and $\beta = \frac{\partial S(Q)}{\partial Q} \in \mathfrak{g}$. We have then defined the Gauss density for Lie groups as the density that maximizes this Entropy $S(Q)$ under the constraint of its associated first moment $Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \int_M J(\xi) p_{Gibbs}(\xi) d\lambda_\omega$. The Gauss density is then established by

analogy with thermodynamics as the Gibbs density $p_{Gibbs}(\xi) = e^{\Phi(\beta) - \langle J(\xi), \beta \rangle} = \frac{e^{-\langle J(\xi), \beta \rangle}}{\int_M e^{-\langle J(\xi), \beta \rangle} d\lambda_\omega}$. But this

is not enough, because this density is not given in the good parametrization. We have proposed to express the Gibbs density with respect to the 1st statistical moment Q (statistical mean of moment map) by inverting the relation $Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \Theta(\beta)$. The Gibbs density $p_{Gibbs,Q}(\xi) = e^{\Phi(\beta) - \langle J(\xi), \Theta^{-1}(Q) \rangle}$ with $\beta = \Theta^{-1}(Q)$ will provide the extended definition of Gauss density in final good parametrization.

For the time being, no "Gaussian density" was defined on Poincaré unit disk with the mandatory property to be covariant under the action of $SU(1,1)$ Lie group that acts transitively on this homogeneous bounded domain. We have applied the previous model via computation of moment map and developed the full computation of this extended Gauss density for $SU(1,1)$ Lie group, $SU(1,1) = \left\{ \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} / a, b \in C, |a|^2 - |b|^2 = 1 \right\}$ and then deduced as consequence the gauss density for the Poincaré unit disk considered as the homogeneous symplectic manifold associated to the coadjoint orbit of the $SU(1,1)$ Lie group via KKS 2 form. Considering the Lie algebra $su(1,1) = \left\{ \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix} / r \in R, \eta \in C \right\}$ and the dual space of the Lie algebra $su(1,1)^* = \left\{ \begin{pmatrix} z & x + iy \\ -x + iy & -z \end{pmatrix} / x, y, z \in R \right\}$, we have computed the moment map $J : D \to su^*(1,1)$ defined by $J(z).u_i = J_i(z, z^*)$, that maps D the Poincaré unit disk into a coadjoint orbit in $su^*(1,1)$, $J(z) = J_1(z, z^*)u_1^* + J_2(z, z^*)u_2^* + J_3(z, z^*)u_3^* = \rho \begin{pmatrix} \frac{1+|z|^2}{1-|z|^2} & -2\frac{z^*}{1-|z|^2} \\ 2\frac{z}{1-|z|^2} & -\frac{1+|z|^2}{1-|z|^2} \end{pmatrix} \in g^*$

The moment map J is a diffeomorphism of D onto one sheet of the two-sheeted hyperboloid in $su^*(1,1)$, determined by the following equation $J_1^2 - J_2^2 - J_3^2 = \rho^2$, $J_1 \geq \rho$ with $J_1 u_1^* + J_2 u_2^* + J_3 u_3^* \in su^*(1,1)$. But the full $SU(1,1)$ Lie group is not related to any equilibrium Gibbs state (the open subset of the Lie algebra, associated to this Gibbs state is empty). We have then considered one-parameter subgroups of the Lie group $SU(1,1)$ such that the open subset $\Lambda_\beta = \left\{ \beta \in g / \int_D e^{-\langle J(z), \beta \rangle} d\lambda(z) < +\infty \right\}$ is not empty. In the neighborhood of the identity element, the elements of $g \in SU(1,1)$ can be written as the exponential of an element β of its Lie algebra. If we make the remark that we have the following relation $\beta^2 = \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix} \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix} = \left(|\eta|^2 - r^2 \right)I$, we can developed the exponential map by a Taylor expansion of the exponential function, which is given by the following relation $g = \exp(\varepsilon\beta) = \sum_{k=0}^{\infty} \frac{(\varepsilon\beta)^k}{k!} = \begin{pmatrix} \cosh(\varepsilon R) + ir\frac{\sinh(\varepsilon R)}{R} & \eta\frac{\sinh(\varepsilon R)}{R} \\ \eta^*\frac{\sinh(\varepsilon R)}{R} & \cosh(\varepsilon R) - ir\frac{\sinh(\varepsilon R)}{R} \end{pmatrix}$ with $R^2 = |\eta|^2 - r^2$.

We can observe that one condition is that $|\eta|^2 - r^2 > 0$ then the subset to consider is given by the subset $\Lambda_\beta = \left\{ \beta = \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix}, r \in R, \eta \in C / |\eta|^2 - r^2 > 0 \right\}$ such that $\int_D e^{-\langle J(z), \beta \rangle} d\lambda(z) < +\infty$. Finally, we have computed the covariant Gibbs density in the unit disk given by $\beta \in \Lambda_\beta$ and by the moment map of the Lie group $SU(1,1)$, that could be expressed in the following equation:

$$p_{Gibbs}(z) = \frac{e^{-\langle J(z), \beta \rangle}}{\int_D e^{-\langle J(z), \beta \rangle} d\lambda(z)} = = e^{-\left\langle \rho \begin{pmatrix} \frac{1+|z|^2}{(1-|z|^2)} & \frac{-2z^*}{(1-|z|^2)} \\ \frac{2z}{(1-|z|^2)} & -\frac{1+|z|^2}{(1-|z|^2)} \end{pmatrix}, \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix} \right\rangle} \Big/ \int_D e^{-\langle J(z), \beta \rangle} d\lambda(z)$$ with $d\lambda(z) = 2i\rho\frac{dz \wedge dz^*}{(1-|z|^2)^2}$. To write

the final Gibbs density with respect to its statistical moment, we rewrite the density with $Q = E[J(z)]$, by $\beta = \Theta^{-1}(Q) \in g$ where $Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \Theta(\beta) \in g^*$ and $Q = E[J(z)] = E\left[\rho \begin{pmatrix} \frac{1+|w|^2}{(1-|w|^2)} & \frac{-2w^*}{(1-|w|^2)} \\ \frac{2w}{(1-|w|^2)} & -\frac{1+|w|^2}{(1-|w|^2)} \end{pmatrix} \right]$.

To extend this approach for covariant Gibbs density on Siegel Unit Disk $SD = \left\{ Z \in M_{pq}(C) / I_p - ZZ^+ > 0 \right\}$, that is a classical matrix extension of Poincaré unit Disk, we have proposed to consider $G = SU(p,q)$ unitary group and the

homogeneous space $G/K = SU(p,q)/S(U(p),U(q))$ with $K = S(U(p) \times U(q)) = \left\{ \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} / A \in U(p), D \in U(q), \det(A)\det(D) = 1 \right\}$ and the moment map given by $J(Z) = $

$$i\lambda \begin{pmatrix} (I_p - ZZ^+)^{-1}(-pZZ^+ - qI_p) & (p+q)Z(I_q - Z^+Z)^{-1} \\ -(p+q)(I_q - Z^+Z)^{-1}Z^+ & (pI_q + qZ^+Z)(I_q - Z^+Z)^{-1} \end{pmatrix}.$$

After $SU(1,1)$ Lie group (case with null cohomology), we have considered the same model for $SE(2)$ Lie group with non-null cohomology that needs the use of symplectic one-cocycle to manage the defect of cohomology. We have considered the special Euclidean group $SE(2) = \left\{ \begin{bmatrix} R_\varphi & \tau \\ 0 & 1 \end{bmatrix} / R_\varphi \in SO(2), \tau \in R^2 \right\}$ with $SO(2) = \left\{ R_\varphi = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} / \varphi \in R \right\}$, and the Lie algebra $se(2)$ of $SE(2)$ $(\xi, u) \in se(2) = R \times R^2 \Rightarrow \begin{bmatrix} -\xi\Im & u \\ 0 & 0 \end{bmatrix} \in se(2)$ with $\Im = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, to define the moment map $J_{(\xi,u)}(x) : R^2 \to se^*(2)$ that is given by the expression $J_{(\xi,u)}(x) = J(x).(\xi, u)$ with $J(x) = -2\left(\frac{1}{2}\|x\|^2, -\Im x\right)$, $x \in R^2$. Then, the Gibbs density is deduced for generalized temperature $\beta \in \Omega = \left\{ (b,B) \in se(2) / b < 0, B \in R^2 \right\}$ by $p_{Gibbs}(x) = \dfrac{e^{-\langle J(x), \beta \rangle}}{\int\limits_{R^2} e^{-\langle J(x), \beta \rangle} d\lambda(x)} = \dfrac{e^{\frac{1}{2}b\|x\|^2 - B.\Im x}}{\int\limits_{R^2} e^{\frac{1}{2}b\|x\|^2 - B.\Im x} d\lambda(x)}$, with the log-partition function given by the following expression $\Phi(\beta) = \log \int\limits_{R^2} e^{\frac{1}{2}b\|x\|^2 - B.\Im x} d\lambda(x) = \log\left(-\frac{2\pi}{b} e^{-\frac{1}{2b}\|B\|^2}\right)$ with $Q = \frac{\partial\Phi(\beta)}{\partial\beta} = \left(\frac{1}{b} - \frac{\|B\|^2}{2b^2}, \frac{1}{b}B\right) = \Theta(\beta)$ and where $Q \in \Omega^* = \left\{ (m,M) \in se^*(2) / m + \frac{\|M\|^2}{2} < 0 \right\}$. To obtain the good parametrization related to statical moments, we have inverted the relation $\beta = \Theta^{-1}(Q) = \left(\left(m + \frac{1}{2}\|M\|^2\right)^{-1}, \left(m + \frac{1}{2}\|M\|^2\right)^{-1}M \right)$, to provide the covariant Gibbs density parametrized by $(m,M) = E(J(x)) = E\left[-2\left(\frac{1}{2}\|x\|^2, -\Im x\right)\right] = \left[-E(\|x\|^2), 2\Im E(x)\right]$. The final Gauss density for $SE(2)$ is then

$$p_{Gibbs}(x) = \dfrac{e^{\frac{\frac{1}{2}\|x\|^2 - M.\Im x}{(m + \frac{1}{2}\|M\|^2)}}}{\int\limits_{R^2} e^{\frac{\frac{1}{2}\|x\|^2 - M.\Im x}{(m + \frac{1}{2}\|M\|^2)}} d\lambda(x)}.$$

We conclude the paper by a deeper study of Souriau model structure. We observe that Souriau Entropy $S(Q)$ defined on coadjoint orbit of the group has a property of invariance $S\left(Ad_g^\#(Q)\right) = S(Q)$ with respect to Souriau affine definition of coadjoint action $Ad_g^\#(Q) = Ad_g^*(Q) + \theta(g)$ where $\theta(g)$ is called the Souriau cocyle. In the framework of Souriau Lie groups Thermodynamics, we can then characterize the Entropy as a generalized Casimir invariant function in coadjoint representation, and Massieu characteristic function (or log-partition function), dual of Entropy by Legendre transform, as a generalized Casimir function in adjoint representation. When M is a Poisson manifold, a function on M is a Casimir function if and only if this function is constant on each symplectic leaf (the non-empty open subsets of the symplectic leaves are the smallest embedded manifolds of M which are Poisson submanifolds) [15]. Classically, the Entropy is defined axiomatically as Shannon or von Neumann Entropies without any geometric structures constraints. In this paper, the Entropy is also presented as solution of the Casimir equation $\left(ad^*_{\frac{\partial S}{\partial Q}} Q\right)_j + \Theta\left(\frac{\partial S}{\partial Q}\right)_j = C_{ij}^k ad^*_{\left(\frac{\partial S}{\partial Q}\right)_i} Q_k + \Theta_j = 0$ with $\widetilde{\Theta}(X,Y) = \langle \Theta(X), Y \rangle = J_{[X,Y]} - \{J_X, J_Y\} = -\langle d\theta(X), Y \rangle$, $X, Y \in g$, where $\Theta(X) = T_e\theta(X(e))$ appears in case of non-null cohomology (non-equivariance of coadjoint operator on the moment map), with $\theta(g)$ the Souriau Symplectic cocycle. The dual space of the Lie algebra foliates into coadjoint orbits that are also the level sets on the entropy. The KKS (Kostant-Kirillov Souriau) 2-form, and the Souriau-Koszul-Fisher metric transform each orbit into a homogeneous Symplectic manifold. The information manifold foliates into level sets of the entropy that could be interpreted in the framework of Thermodynamics by the fact that motion remaining on this complex surfaces is non-dissipative, whereas motion transversal to these surfaces is dissipative, where the dynamic is given by $\frac{dQ}{dt} = \{Q, H\}_{\widetilde{\Theta}} = ad^*_{\frac{\partial H}{\partial Q}} Q + \Theta\left(\frac{\partial H}{\partial Q}\right)$ with

stable equilibrium when $H = S \Rightarrow \frac{dQ}{dt} = \{Q, S\}_{\widetilde{\Theta}} = ad^*_{\frac{\partial S}{\partial Q}} Q + \Theta\left(\frac{\partial S}{\partial Q}\right) = 0$. We have finally also observed that $dS = \widetilde{\Theta}_\beta\left(\frac{\partial H}{\partial Q}, \beta\right) dt$ where $\widetilde{\Theta}_\beta\left(\frac{\partial H}{\partial Q}, \beta\right) = \widetilde{\Theta}\left(\frac{\partial H}{\partial Q}, \beta\right) + \left\langle Q, \left[\frac{\partial H}{\partial Q}, \beta\right]\right\rangle$, showing that Entropy production is linked with Souriau tensor related to Fisher metric.

The Casimir equations that we have introduced in non-zero cohomology case are consequences of the constancy of the entropy on adjoint orbits of the Lie algebra and of the equivariance of the map between the set of generalized temperatures and the dual space of the Lie algebra, as introduced by Jean-Marie in his 1974 paper. We explained this fact in the paper by starting elaboration of Casimir equations from the Souriau equation. Casimir equations are then presented in this context, as a fully equivalent form written in a new way, especially in the framework of Souriau Lie groups Thermodynamics. Souriau has not observed that the Entropy is an invariant Casimir function in coadjoint representation, but we can assume that he was fully aware of this invariant structure.

From Souriau equation $\langle Q, [\beta, Z]\rangle + \widetilde{\Theta}(\beta, Z) = 0$ published in 1974, we have rewritten as direct consequence this equation on a Casimir form $ad^*_{\frac{\partial S}{\partial Q}} Q + \Theta\left(\frac{\partial S}{\partial Q}\right) = 0$. This equation preserves the geometric structures included in Souriau equation but allow us to consider the Entropy from the point of view of Casimir invariant function. The concept of Entropy and the concept of Casimir function were, for the time being, two disjoint concepts that have been developed independently in the past. There is a large literature on Casimir function, especially the russian one that have characterized properties of Casimir function. We refer to Igor V. Shirokov who has proposed a method for constructing invariants of the coadjoint representation of Lie groups with an arbitrary dimension and structure based on local symplectic coordinates on the coadjoint orbits. With Oleg L. Kurnyavko, Igor V. Shirokov has also proposed a general method for constructing invariant Casimir functions. The second reference is about A.T. Fomenko and V.V. Trofimov who have also deeply studied Casimir functions (but in case of null cohomology) and have developed the following equation that we can write for Entropy in null cohomology case $S\left(Ad^*_{e^{t\xi}} Q\right) = S(Q) + \sum_{n=1}^{\infty} \frac{(-\phi(\xi))^n S}{n!}(Q).t^n$ with $\phi : g \rightarrow Vec(\Gamma)$ a representation of Lie algebras defined on basis $(e_1, e_2, \ldots, e_n)$ in g. We refer to a twin paper [17] developing consequences of this new definition of Entropy as an invariant Casimir function. In this twin paper, we study the associated Euler-Poincaré equation $\frac{dQ}{dt} = ad^*_{\frac{\partial H}{\partial Q}} Q + \Theta\left(\frac{\partial H}{\partial Q}\right)$ and the stochastic extension based on a new Stratonovich differential equation for the stochastic process given by the following relation by mean of Souriau's symplectic cocycle $dQ + \left[ad^*_{\frac{\partial H}{\partial Q}} Q + \Theta\left(\frac{\partial H}{\partial Q}\right)\right] dt + \sum_{i=1}^{N} \left[ad^*_{\frac{\partial H_i}{\partial Q}} Q + \Theta\left(\frac{\partial H_i}{\partial Q}\right)\right] \circ dW_i(t) = 0$. These kind of stochastic equations have been also studied by Alexis Arnaudon and Daryl Holm but only in the restricted case of null-cohomology [32].

We give references from classical textbooks (as Souriau book and papers) to preprints because different approaches have been developed in parallel to address Lie groups statistics, as soon as mid of last century, but without bridges between these disciplines which have developed specific tools to address this problem. We have limited these references to main and important documents, which are characterized as seminal and as tutorial of their domains. We have preserved references in French, because some works as Souriau Lie groups Thermodynamics model have not been yet largely spread towards the different communities.

3. Learning Inference Lie Groups Thermodynamics and Covariant Gibbs Density

We identify the Riemanian metric introduced by Souriau based on cohomology, in the framework of "Lie groups thermodynamics" as an extension of classical Fisher metric introduced in information geometry. We have observed that Souriau metric preserves Fisher metric structure as the Hessian of the minus logarithm of a partition function, where the partition function is defined as a generalized Laplace transform on a sharp convex cone. Souriau's definition of Fisher metric extends the classical one in case of Lie groups or homogeneous manifolds. Souriau has developed this "Lie groups thermodynamics" theory in the framework of homogeneous symplectic manifolds in geometric statistical mechanics for

dynamical systems, but as observed by Souriau, these model equations are no longer linked to the symplectic manifold but equations only depend on the Lie group and the associated cocycle [33,34]. This analogy with Fisher metric opens potential applications in machine learning, where the Fisher metric is used in the framework of information geometry, to define the "natural gradient" tool for improving ordinary stochastic gradient descent sensitivity to rescaling or changes of variable in parameter space. In machine learning revised by natural gradient of information geometry, the ordinary gradient is designed to integrate the Fisher matrix. Amari has theoretically proved the asymptotic optimality of the natural gradient compared to classical gradient. With the Souriau approach, the Fisher metric could be extended, by Souriau-Fisher metric, to design natural gradients for data on homogeneous manifolds. Information geometry has been derived from invariant geometrical structure involved in statistical inference. The Fisher metric defines a Riemannian metric as the Hessian of two dual potential functions, linked to dually coupled affine connections in a manifold of probability distributions. With the Souriau model, this structure is extended preserving the Legendre transform between two dual potential function parametrized in Lie algebra of the group acting transentively on the homogeneous manifold.

3.1. Inference by Natutal Gradient and Legendre Structure

Classically, to optimize the parameter θ of a probabilistic model, based on a sequence of observations y_t, is an online gradient descent:

$$\theta_t \leftarrow \theta_{t-1} - \eta_t \frac{\partial l_t(y_t)^T}{\partial \theta} \tag{18}$$

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

$$I(\theta) = \left[g_{ij}\right] \text{ with } g_{ij} = \left[-E_{y\sim p(y/\theta)}\left[\frac{\partial^2 \log p(y/\theta)}{\partial \theta_i \partial \theta_j}\right]\right]_{ij} \tag{19}$$

with natural gradient:

$$\theta_t \leftarrow \theta_{t-1} - \eta_t I(\theta)^{-1} \frac{\partial l_t(y_t)^T}{\partial \theta} \tag{20}$$

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

$$g_{ij} = -\left[\frac{\partial^2 \Phi}{\partial \theta_i \partial \theta_j}\right]_{ij} \text{ with } \Phi(\theta) = -\log \int_R e^{-\langle \theta, y \rangle} dy \tag{21}$$

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

$$S(\eta) = \langle \theta, \eta \rangle - \Phi(\theta) \text{ with } \eta_i = \frac{\partial \Phi(\theta)}{\partial \theta_i} \text{ and } \theta_i = \frac{\partial S(\eta)}{\partial \eta_i} \tag{22}$$

We can observe that $\Phi(\theta) = -\log \int_R e^{-\langle \theta, y \rangle} dy = -\log \psi(\theta)$ is linked with the cumulant generating function.

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

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

$$\psi_\Omega(\theta) = \int_{\Omega^*} e^{-\langle \theta, y \rangle} dy \text{ with Koszul-Vinberg Characteristic function} \tag{23}$$

Jean-Louis Koszul has introduced the following forms
1st Koszul form:

$$\alpha = d\Phi_\Omega(\theta) = -d \log \psi_\Omega(\theta) \tag{24}$$

2nd Koszul form:

$$\gamma = D\alpha = Dd \log \psi_\Omega(\theta) \tag{25}$$

with the following property of positive definitiveness:

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

$$\text{with } F(\xi) = e^{-\frac{1}{2}\langle x, \xi \rangle} \text{ and } G(\xi) = e^{-\frac{1}{2}\langle x, \xi \rangle} \langle u, \xi \rangle \tag{26}$$

Koszul has defined the following Diffeomorphism:

$$\eta = \alpha = -d \log \psi_\Omega(\theta) = \int_{\Omega^*} \xi p_\theta(\xi) d\xi \text{ with } p_\theta(\xi) = \frac{e^{-\langle \xi, \theta \rangle}}{\int_{\Omega^*} e^{-\langle \xi, \theta \rangle} d\xi} \tag{27}$$

with preservation of Legendre transform:

$$S_\Omega(\eta) = \langle \theta, \eta \rangle - \Phi_\Omega(\theta) \text{ with } \eta = d\Phi_\Omega(\theta) \text{ and } \theta = dS_\Omega(\eta) \tag{28}$$

3.2. Souriau Lie Groups Thermodynamique and Souriau-Koszul-Fisher Metric

This relations have been extended by Jean-Marie Souriau in geometric statistical mechanics, where he developed a "Lie groups thermodynamics" of dynamical systems where the (maximum entropy) Gibbs density is covariant with respect to the action of the Lie group. In the Souriau model, previous structures of information geometry are preserved:

$$I(\beta) = -\frac{\partial^2 \Phi}{\partial \beta^2} \text{ with } \Phi(\beta) = -\log \int_M e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega \text{ and } U : M \to g^* \tag{29}$$

$$S(Q) = \langle Q, \beta \rangle - \Phi(\beta) \text{ with } Q = \frac{\partial \Phi(\beta)}{\partial \beta} \in g^* \text{ and } \beta = \frac{\partial S(Q)}{\partial Q} \in g \tag{30}$$

In the Souriau Lie groups thermodynamics model, β is a "geometric" (Planck) temperature, element of Lie algebra g of the group, and Q is a "geometric" heat, element of the dual space of the Lie algebra g^* of the group. Souriau has proposed a Riemannian metric that we have identified as a generalization of the Fisher metric:

$$I(\beta) = \left[g_\beta \right] \text{ with } g_\beta([\beta, Z_1], [\beta, Z_2]) = \widetilde{\Theta}_\beta(Z_1, [\beta, Z_2]) \tag{31}$$

$$\text{with } \widetilde{\Theta}_\beta(Z_1, Z_2) = \widetilde{\Theta}(Z_1, Z_2) + \left\langle Q, ad_{Z_1}(Z_2) \right\rangle \text{ where } ad_{Z_1}(Z_2) = [Z_1, Z_2] \tag{32}$$

Souriau has proved that all co-adjoint orbit of a Lie group given by $O_F = \left\{ Ad_g^* F = g^{-1} Fg, g \in G \right\}$ subset of g^*, $F \in g^*$ carries a natural homogeneous symplectic structure by

a closed *G*-invariant 2-form. If we define $K = Ad_g^* = \left(Ad_{g^{-1}}\right)^*$ and $K_*(X) = -(ad_X)^*$ with $\left\langle Ad_g^* F, Y \right\rangle = \left\langle F, Ad_{g^{-1}} Y \right\rangle, \forall g \in G, Y \in \mathbf{g}, F \in \mathbf{g}^*$ where if $X \in \mathbf{g}, Ad_g(X) = gXg^{-1} \in \mathbf{g}$, the G-invariant 2-form is given by the following expression $\sigma_\Omega(ad_X F, ad_Y F) = B_F(X, Y) = \langle F, [X, Y] \rangle, X, Y \in \mathbf{g}$. Souriau Foundamental Theorem is that « Every symplectic manifold on which a Lie group acts transitively by a Hamiltonian action is a covering space of a coadjoint orbit ». We can observe that for Souriau model, Fisher metric is an extension of this 2-form in non-equivariant case $g_\beta([\beta, Z_1], [\beta, Z_2]) = \widetilde{\Theta}(Z_1, [\beta, Z_2]) + \langle Q, [Z_1, [\beta, Z_2]] \rangle$ ».

The Souriau additional term $\widetilde{\Theta}(Z_1, [\beta, Z_2])$ is generated by non-equivariance through Symplectic cocycle. The tensor $\widetilde{\Theta}$ used to define this extended Fisher metric is defined by the moment map $J(x)$, application from *M* (homogeneous symplectic manifold) to the dual space of the Lie algebra $\mathbf{g}^*$, given by:

$$\widetilde{\Theta}(X, Y) = J_{[X,Y]} - \{J_X, J_Y\} \tag{33}$$

$$\text{with } J(x) : M \to \mathbf{g}^* \text{ such that } J_X(x) = \langle J(x), X \rangle, X \in \mathbf{g} \tag{34}$$

This tensor $\widetilde{\Theta}$ is also defined in tangent space of the cocycle $\theta(g) \in \mathbf{g}^*$ (this cocycle appears due to the non-equivariance of the coadjoint operator Ad_g^*, action of the group on the dual space of the lie algebra; the action of the group on the dual space of the Lie algebra is modified with a cocycle so that the momentu map becomes equivariant relative to this new affine action):

$$Q\left(Ad_g(\beta)\right) = Ad_g^*(Q) + \theta(g) \tag{35}$$

$\theta(g) \in \mathbf{g}^*$ is called nonequivariance one-cocycle, and it is a measure of the lack of equivariance of the moment map.

$$\widetilde{\Theta}(X, Y) : \mathbf{g} \times \mathbf{g} \to \mathfrak{R} \qquad \text{with } \Theta(X) = T_e\theta(X(e))$$
$$X, Y \mapsto \langle \Theta(X), Y \rangle \tag{36}$$

The cocycle should verify:

$$\begin{aligned}
\theta(st) &= J((st).x) - Ad_{st}^* J(x) \\
\theta(st) &= [J(s.(t.x)) - Ad_s^* J(t.x)] + \left[Ad_s^* J(t.x) - Ad_s^* Ad_t^* J(x)\right] \\
\theta(st) &= \theta(s) + Ad_s^* \left[J(t.x) - Ad_t^* J(x)\right] \\
\theta(st) &= \theta(s) + Ad_s^* \theta(t)
\end{aligned} \tag{37}$$

We can also compute tangent of one-cocycle θ at neutral element, to compute 2-cocycle Θ:

$$\begin{aligned}
\zeta \in \mathbf{g}, \theta_\zeta(s) &= \langle \theta(s), \zeta \rangle = \langle J(s.x), \zeta \rangle - \langle Ad_s^* J(x), \zeta \rangle = \langle J(s.x), \zeta \rangle - \left\langle J(x), Ad_{s^{-1}} \zeta \right\rangle \\
T_e\theta_\zeta(\xi) &= \left\langle T_x J.\xi_p(x), \zeta \right\rangle + \left\langle J(x), ad_\xi \zeta \right\rangle \text{ with } \xi_p = X_{\langle J, \xi \rangle} \\
T_e\theta_\zeta(\xi) &= X_{\langle J(x), \xi \rangle}[\langle J(x), \zeta \rangle] + \langle J(x), [\xi, \zeta] \rangle \\
T_e\theta_\zeta(\xi) &= -\{\langle J, \xi \rangle, \langle J, \zeta \rangle\} + \langle J(x), [\xi, \zeta] \rangle = \Theta(\xi)
\end{aligned} \tag{38}$$

We can also write: $T_x J\left(\xi_p(x)\right) = -ad_\xi^* J(x) + \Theta(\xi, .)$
By differentiating the equation on affine action, we have:

$$dJ(Xx) = ad_X J(x) + d\theta(X) , x \in M, X \in \mathbf{g} \tag{39}$$

$$\begin{aligned}
\langle dJ(Xx), Y \rangle &= \langle ad_X J(x), Y \rangle + \langle d\theta(X), Y \rangle, x \in M, X, Y \in \mathbf{g} \\
\langle dJ(Xx), Y \rangle &= \langle J(x), [X, Y] \rangle + \langle d\theta(X), Y \rangle = \{\langle J, X \rangle, \langle J, Y \rangle\}(x) \\
\langle J(x), [X, Y] \rangle &- \{\langle J, X \rangle, \langle J, Y \rangle\}(x) = -\langle d\theta(X), Y \rangle
\end{aligned} \tag{40}$$

It can be then deduced that the tensor could be also written:

$$\widetilde{\Theta}(X, Y) = J_{[X,Y]} - \{J_X, J_Y\} = -\langle d\theta(X), Y \rangle \,, \, X, Y \in \mathrm{g} \tag{41}$$

with the cocycle property:

$$\widetilde{\Theta}([X, Y], Z) + \widetilde{\Theta}([X, Y], Z) + \widetilde{\Theta}([X, Y], Z) = 0 \,, \, X, Y, Z \in \mathrm{g} \tag{42}$$

By noting the action of the group on the dual space of the Lie algebra:

$$G \times \mathrm{g}^* \to \mathrm{g}^*, (s, \xi) \mapsto s\xi = Ad_s^*\xi + \theta(s) \tag{43}$$

Associativity is also derived:

$$\begin{aligned}
(s_1 s_2)\xi &= Ad_{s_1 s_2}^*\xi + \theta(s_1 s_2) = Ad_{s_1}^* Ad_{s_2}^*\xi + \theta(s_1) + Ad_{s_1}^*\theta(s_2) \\
(s_1 s_2)\xi &= Ad_{s_1}^*\left(Ad_{s_2}^*\xi + \theta(s_2)\right) + \theta(s_1) = s_1(s_2\xi) \,, \, \forall s_1, s_2 \in G, \xi \in \mathrm{g}^*
\end{aligned} \tag{44}$$

This study of the moment map J equivariance, and the existence of an affine action of G on g^*, whose linear part is the coadjoint action, for which the moment J is equivariant, is at the cornerstone of Souriau theory of geometric mechanics and Lie groups thermodynamics.

3.3. Souriau Entropy and Souriau-Fisher-Koszul Metric Invariance under the Action of the Group and Covariant Souriau Gibbs Density

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

$$\beta \in \mathrm{g} \to Ad_g(\beta) \Rightarrow \begin{cases} S\left[Q\left(Ad_g(\beta)\right)\right] = S(Q) \\ I\left[Ad_g(\beta)\right] = I(\beta) \end{cases} \tag{45}$$

In the framework of Lie group action on a symplectic manifold, equivariance of moment map could be studied to prove that there is a unique action $a(.,.)$ of the Lie group G on the dual g^* of its Lie algebra for which the moment map J is equivariant, that means for each $x \in M$:

$$J\left(\Phi_g(x)\right) = a(g, J(x)) = Ad_g^*(J(x)) + \theta(g) \tag{46}$$

When coadjoint action is not equivariant, the symmetry is broken, and new "cohomological" relations should be verified in Lie algebra of the group. A natural equilibrium state will thus be characterized by an element of the Lie algebra of the Lie group, determining the equilibrium temperature β. The entropy $s(Q)$, parametrized by Q the geometric heat (mean of energy U, element of the dual space of the Lie algebra) is defined by the Legendre transform of the Massieu potential $\Phi(\beta)$ parametrized by β ($\Phi(\beta)$ is the minus logarithm of the partition function $\psi_\Omega(\beta)$).

A Gibbs state, in the usual sense, is a statistical state at which the entropy is stationary with respect to all infinitesimal variations of the statistical state for which the mean value of the energy remains constant. In the sense of Souriau, a generalized Gibbs state is a statistical state at which the entropy is stationary with respect to all infinitesimal variations of the statistical state for which the mean value of the moment map remains constant. This generalization is very natural, since the energy can be considered as the moment map of the Hamiltonian action of the one-dimensional Lie group of time translations. Furthermore, each generalized Gibbs state is associated to an element of the Lie algebra of the group, called by Souriau a generalized temperature, and that the set of possible generalized temperature is not, in general the whole Lie algeba, but an open convex subset of the Lie algebra,

which may be empty, for which some integrals encountered in the expression of the generalized Gibbs state are normally convergent. So, for some Lie groups, generalized Gibbs states do not exist, and there is no Souriau Lie groups thermodynamics.

Souriau has then defined a Gibbs density that is covariant under the action of the group:

$$p_{Gibbs}(\xi) = e^{\Phi(\beta) - \langle U(\xi), \beta \rangle} = \frac{e^{-\langle U(\xi), \beta \rangle}}{\int_M e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega} \text{ , with } \Phi(\beta) = -\log \int_M e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega$$

$$Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \frac{\int_M U(\xi) e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega}{\int_M e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega} = \int_M U(\xi) p(\xi) d\lambda_\omega \tag{47}$$

We can express the Gibbs density with respect to Q by inverting the relation $Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \Theta(\beta)$. Then $p_{Gibbs,Q}(\xi) = e^{\Phi(\beta) - \langle U(\xi), \Theta^{-1}(Q) \rangle}$ with $\beta = \Theta^{-1}(Q)$. All Souriau equations of Lie groups Thermodynamics are illustrated in Figures 3 and 4.
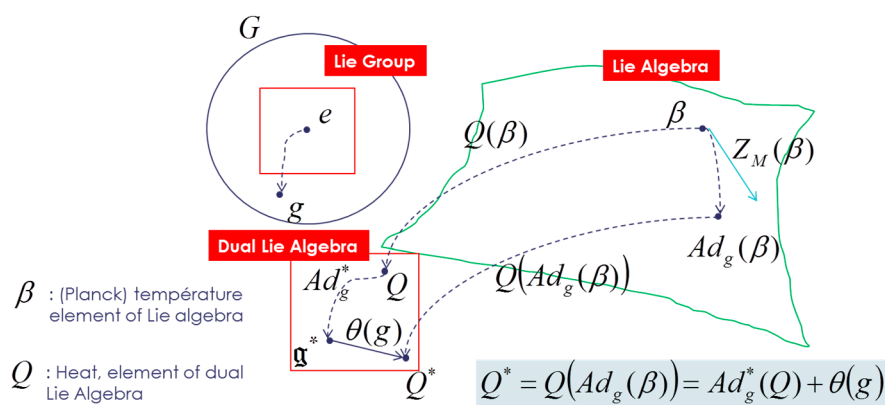
Figure 3. Fondamental Equation of Souriau Lie groups Thermodynamics. Q is the geometric heat in dual Lie algebra, β is the geometric temperature in Lie algebra.
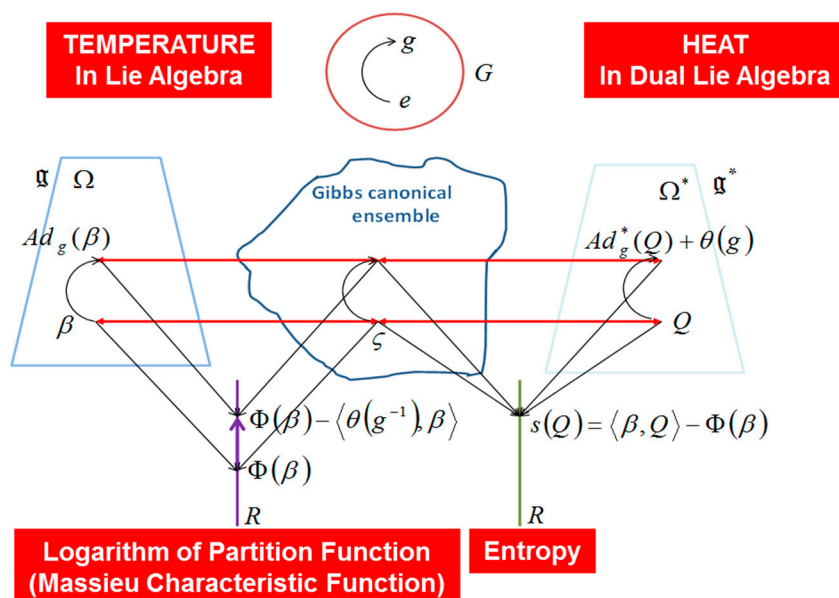
Figure 4. Souriau Model of Lie Groups Thermodynamics.

Souriau completed his "geometric heat theory" by introducing a 2-form in the Lie algebra, that is a Riemannian metric tensor in the values of adjoint orbit of β, $[\beta, Z]$ with Z an element of the Lie algebra. This metric is given for (β, Q):

$$g_\beta([\beta, Z_1], [\beta, Z_2]) = \langle \Theta(Z_1), [\beta, Z_2] \rangle + \langle Q, [Z_1, [\beta, Z_2]] \rangle \tag{48}$$

where Θ is a cocycle of the Lie algebra, defined by $\Theta = T_e \theta$ with θ a cocycle of the Lie group defined by $\theta(M) = Q(Ad_M(\beta)) - Ad_M^* Q$.

We observe that Souriau Riemannian metric, introduced with symplectic cocycle, is a generalization of the Fisher metric, that we call the Souriau-Fisher metric, that preserves the property to be defined as a Hessian of the partition function logarithm $g_\beta = -\frac{\partial^2 \Phi}{\partial \beta^2} = \frac{\partial^2 \log \psi_\Omega}{\partial \beta^2}$ as in classical information geometry. We will establish the equality of two terms, between Souriau definition based on Lie group cocycle Θ and parameterized by "geometric heat" Q (element of the dual space of the Lie algebra) and "geometric temperature" β (element of Lie algebra) and hessian of characteristic function $\Phi(\beta) = -\log \psi_\Omega(\beta)$ with respect to the variable β (as illustrated in Figure 5):
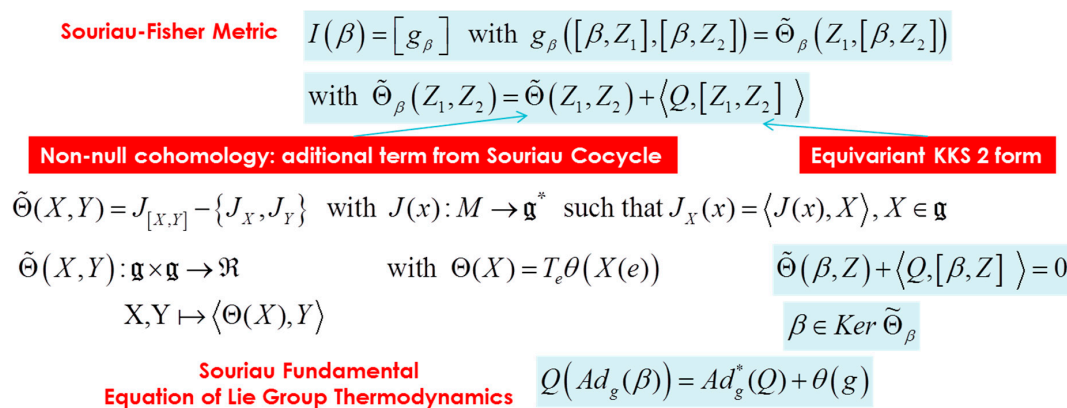
$$g_\beta([\beta, Z_1], [\beta, Z_2]) = \langle \Theta(Z_1), [\beta, Z_2] \rangle + \langle Q, [Z_1, [\beta, Z_2]] \rangle = \frac{\partial^2 \log \psi_\Omega}{\partial \beta^2} \tag{49}$$

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

$$\frac{\partial Q}{\partial \beta}(-[Z_1, \beta], .) = \widetilde{\Theta}(Z_1, [\beta, .]) + \langle Q, Ad_{.Z_1}([\beta, .]) \rangle = \widetilde{\Theta}_\beta(Z_1, [\beta, .]) \tag{50}$$

$$-\frac{\partial Q}{\partial \beta}([Z_1, \beta], Z_2.) = \widetilde{\Theta}(Z_1, [\beta, Z_2]) + \langle Q, Ad_{.Z_1}([\beta, Z_2]) \rangle = \widetilde{\Theta}_\beta(Z_1, [\beta, Z_2]) \tag{51}$$

$$\Rightarrow -\frac{\partial Q}{\partial \beta} = g_\beta([\beta, Z_1], [\beta, Z_2]) \tag{52}$$

Souriau-Fisher Metric　$I(\beta) = [g_\beta]$ with $g_\beta([\beta, Z_1], [\beta, Z_2]) = \tilde{\Theta}_\beta(Z_1, [\beta, Z_2])$

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

Non-null cohomology: aditional term from Souriau Cocycle　　**Equivariant KKS 2 form**

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

$\tilde{\Theta}(X, Y): \mathfrak{g} \times \mathfrak{g} \to \Re$　　with $\Theta(X) = T_e \theta(X(e))$　　$\tilde{\Theta}(\beta, Z) + \langle Q, [\beta, Z] \rangle = 0$

$X, Y \mapsto \langle \Theta(X), Y \rangle$　　　　$\beta \in Ker\ \tilde{\Theta}_\beta$

Souriau Fundamental Equation of Lie Group Thermodynamics　$Q(Ad_g(\beta)) = Ad_g^*(Q) + \theta(g)$

Figure 5. Souriau-Fisher metric as extension of KKS 2-form in case of non-null Cohomogy.

As the entropy is defined by the Legendre transform of the characteristic function, a dual metric of the Fisher metric is also given by the hessian of "geometric entropy" $S(Q)$ with respect to the dual variable given by Q: $\frac{\partial^2 S(Q)}{\partial Q^2}$.

For the maximum entropy density (Gibbs density), the following three terms coincide: $\frac{\partial^2 \log \psi_\Omega}{\partial \beta^2}$ that describes the convexity of the log-likelihood function, $I(\beta) = -E\left[\frac{\partial^2 \log p_\beta(\xi)}{\partial \beta^2}\right]$ the Fisher metric that describes the covariance of the log-likelihood gradient, whereas $I(\beta) = E\left[(\xi - Q)(\xi - Q)^T\right] = Var(\xi)$

that describes the covariance of the observables. We can also observe that the Fisher metric $I(\beta) = -\frac{\partial Q}{\partial \beta}$ is exactly the Souriau metric defined through symplectic cocycle:

$$I(\beta) = \widetilde{\Theta}_\beta(Z_1, [\beta, Z_2]) = g_\beta([\beta, Z_1], [\beta, Z_2]) \tag{53}$$

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

3.4. Covariant Souriau Gibbs Density and Information Manifold Foliation

R.F. Streater has studied in 1999, Information Geometry for some Lie algebra where for certain unitary representation of a Lie algebra, he has defined the statistical manifold of states as convex cone for which the partition function is finite, making reference to Bogoliubov-Kubo-Mori metric. But Streater has only developed the case with null cohomology for so (3) and sl (2,R) Lie alebras. Nevertheless, as observed by R.F. Streater in his paper "**Information Geometry for some Lie algebras**" [35], referring to Kirillov work and Roger Balian paper, "*We can expect further natural structures to arise in this case. Indeed, it is known (*) that **the dual to the Lie algebra**, which parametrizes the state-space in this case, **foliates into coadjoint orbits; there are also the level sets on the entropy; Kirillov form, and the BKM (Bogoliubov-Kubo-Mori) metric, together make each orbit into kähler space**, along the lines proposed by Kostant. **Motion along these holomorphic directions is nondissipative. The transversal to the orbits is a real half-line, which represents the dissipative direction** ... We study the case of sl (2,R) in the discrete series of representations. We show **the information manifold foliates into level sets of the entropy**, each being isometric to H, the Poincaré upper half-plane ... The states of constant entropy are the hyperboloids and β is the dissipative coordinate ... For an integrable system described by a Lie algebra in a traceable representation, we find that **the information manifold foliates into complex spaces; the level sets of entropy can be given a complex structure by the method of Kostant**. Motion remaining on the complex surfaces is nondissipative, whereas motion transversal to these surfaces is dissipative. In information geometry, the state is parametrized by the canonical coordinates. Which function of them is measured by a thermometer? In our models, it is reasonable to designate $1/\beta$ to be the temperature; it is a dissipative coordinate, and it increases with time, showing that the system is thermalizing*".

4. Mathematical Definition of Souriau Moment Map

Previously, we have introduced the concept of Souriau's moment map. In this chapter, we will introduce a mathematical definition of this tool, as defined in Souriau's book [36] with modern notations [37–41]. Other details on moment map are also given in Jean-Louis Koszul's Book [42].

4.1. Operations on Vector Fields

Consider a map $F : X \subset R^M \to Y \subset R^N$, $y = F(x)$, the derivative of F at $x \in X$, $DF : X \to R^{N \times M}$ is given by:

$$\begin{pmatrix} \delta y^1 \\ \vdots \\ \delta y^N \end{pmatrix} = \begin{pmatrix} \frac{\partial y^1}{\partial x^1} & \cdots & \frac{\partial y^1}{\partial x^M} \\ \vdots & \ddots & \vdots \\ \frac{\partial y^N}{\partial x^1} & \cdots & \frac{\partial y^N}{\partial x^M} \end{pmatrix} \begin{pmatrix} \delta x^1 \\ \vdots \\ \delta x^M \end{pmatrix} = DF(x)(\delta x) = \underset{t \to 0}{Lim} \frac{F(x + t\delta x) - F(x)}{t} \tag{54}$$

Second derivative is given by the linear map $D^2 F : X \to R^{N \times M \times M}$:

$$\delta \left[\frac{\partial y}{\partial x} \right] = \frac{\partial^2 y}{\partial x^2}(\delta x) = D^2 F(x)(\delta x) \tag{55}$$

Consider a vector Field V on $X \subset R^M$ defined by: $V : X \subset R^M \to R^M$, operations on vector fields are given by adjoint action and Lie bracket:

$$Ad_F V(y) = \frac{d}{dt}\left[F \circ e^{tV} \circ F^{-1}\right](y)\Big|_{t=0} = DF(x)(V(x)) \text{ with } x = F^{-1}(y) \tag{56}$$

$$[U,V](x) = \frac{d}{ds}Ad_{e^{sU}}V(x)\Big|_{s=0} = DU(x)(V(x)) - DV(x)(U(x)) \tag{57}$$

0-form is a scalar, 1-form are row $\omega = (\omega_1 \cdots \omega_M)$ in dual space. 2-forms can be regarded as antisymmetric matrices (ω_{ij}) with $\omega(u,v) = u^t \begin{pmatrix} \omega_{11} & \cdots & \omega_{1M} \\ \vdots & \ddots & \vdots \\ \omega_{M1} & \cdots & \omega_{MM} \end{pmatrix} v$. m-forms are all scalar multiples of the standard volume form vol, defined by $Vol(v_1, \cdots, v_m) = \det(\text{matrix with columns } v_1, \ldots, v_m)$.

4.2. Derivative Rules by Sophus Lie, Elie Cartan and Henri Cartan

With the following classical definitions:

- **Pull back**: $F^*\omega$ is a p-form on X

$$F^*\omega\left(v_1, \cdots, v_p\right) = \omega_{F(x)}\left(DF(x)(v_1), \cdots, DF(x)(v_p)\right) \tag{58}$$

- **Interior product**: $i_V\omega$ is the (p−1)form on M obtained by inserting $V(x)$ as the first argument of ω

$$i_V\omega\left(v_2, \cdots v_p\right) = \omega\left(V(x), v_2, \cdots, v_p\right) \tag{59}$$

- **Exterior product:** $\theta \wedge \omega$ is the (p + 1)-form on X where ω is a p-form and θ is a 1-form on M (where the hat indicates a term to be omitted):

$$\theta \wedge \omega\left(v_0, \cdots, v_p\right) = \sum_{i=0}^{p} (-1)^i \theta(v_i)\omega\left(v_0, \cdots, \hat{v}_i, \cdots, v_p\right) \tag{60}$$

- **Lie derivative**: $L_V\omega$ is a p-form on M, and $L_V\omega = 0$ if the flow of V consists of symmetries of ω:

$$L_V\omega\left(v_1, \cdots, v_p\right) = \frac{d}{dt}e^{tV*}\omega\left(v_1, \cdots, v_p\right)\Big|_{t=0} \tag{61}$$

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

- **Exterior derivative:**

$$d\omega\left(v_0, \cdots, v_p\right) = \sum_{i=0}^{p} (-1)^i \frac{\partial \omega}{\partial x}(v_i)(v_0, \cdots, \hat{v}_i, \cdots, v_p) \tag{62}$$

$$p = 0, [d\omega]_i = \partial_i\omega \; ; p = 1, [d\omega]_{ij} = \partial_i\omega_j - \partial_j\omega_i \; ; p = 2, [d\omega]_{ijk} = \partial_i\omega_{jk} + \partial_j\omega_{ki} + \partial_k\omega_{ij} \tag{63}$$

From these definitions, the properties of the exterior and Lie Derivative were established by Sophus Lie, Elie Cartan, and Henri Cartan:

$$L_V\omega = di_V\omega + i_V d\omega \tag{64}$$

(Elie Cartan equation)

$$i_{[U,V]}\omega = i_V L_U\omega - L_U i_V\omega \tag{65}$$

(Henri Cartan equation)

$$L_{[U,V]}\omega = L_V L_U \omega - L_U L_V \omega \tag{66}$$

(Sophus Lie equation)

4.3. Souriau Moment Map

Considering Manifolds and Lie groups, We define the tangent bundle TX of X as the disjoint union of the $T_x X$, or the set of all pairs $\begin{pmatrix} \delta x \\ x \end{pmatrix}$ with $x \in X$ and $\delta x \in T_x X$. If $F : X \to Y$ is a smooth map between manifolds, its tangent map is the map:

$$F_* \begin{pmatrix} \delta x \\ x \end{pmatrix} = \begin{pmatrix} DF(x)(\delta x) \\ F(x) \end{pmatrix} \tag{67}$$

A Lie group is a group G with a manifold structure such that the product $(g, h) \mapsto gh$ and the inversion $g \mapsto g^{-1}$ are smooth maps from $G \times G$ (resp. G) to G. Its Lie algebra is the tangent space $g = T_e G$ at the identity element. A smooth action of G on a manifold X is a group morphism:

$$\begin{aligned} \Phi : G \times X &\to Diff(X) \\ (g, x) &\mapsto g.x \end{aligned} \tag{68}$$

The orbit of $x \in X$ is $G(x) = \{g.x : g \in G\}$.
The tangent space to an orbit at x:
$T_x G(x) = \{Z(x) : Z \in g\} = g/g_x$ with $Z(x) = \frac{d}{dt} e^{tZ}(x)\big|_{t=0}$ and where

$$g_x = \{Z \in g : Z(x) = 0\} \tag{69}$$

Let (M, σ) be a connected symplectic manifold. A vector field η on M is called symplectic if its flow preserves the 2-form: $L_\eta \sigma = 0$. If we use Elie Cartan's formula, we can deduce that $L_\eta \sigma = d i_\eta \sigma + i_\eta d\sigma = 0$ but as $d\sigma = 0$ then $d i_\eta \sigma = 0$. We observe that the 1-form $i_\eta \sigma$ is closed. When this 1-form is exact, there is a smooth function $x \mapsto H$ on M with:

$$i_\eta \sigma = -dH \tag{70}$$

This vector field η is called Hamiltonian and could be defined as symplectic gradient $\eta = \nabla_{Symp} H$.

Let a Lie group G that acts on M and that also preserve σ. A moment map exists if these infinitesimal generators are actually hamiltonian, so that a map $J : M \to g^*$ exists with:
$i_{Z_X} \sigma = -dH_Z$ where

$$H_Z = \langle J(x), Z \rangle \tag{71}$$

The Poisson bracket of two functions H, H' is defined by:

$$\{H, H'\} = \sigma(\eta, \eta') = \sigma\left(\nabla_{Symp} H', \nabla_{Symp} H\right) \text{ with } i_\eta \sigma = -dH \text{ and} \tag{72}$$
$$i_{\eta'} \sigma = -dH'$$

If G is connected, then the moment map is G-equivariant if and only if it satisfies $\{H_Z, H_{Z'}\} = H_{[Z,Z']}$.

Souriau has proved thet every coadjoint orbit of a Lie group is a homogeneous symplectic manifold when endowed with the KKS 2-form $\sigma(Z(x), Z'(x)) = \langle x, [Z', Z] \rangle$, and conversely, every homogeneous symplectic manifold of a connected Lie group G is, up to a possible covering, a coadjoint orbit of some central extension of G. σ is G-invariant.

5. Poincaré Unit Disk, SU(1,1) Lie Group and Souriau Moment Map

We will introduce Souriau moment map for SU(1,1)/K group that acts transitively on Poincaré Unit Disk, based on moment map. More details on computation of moment map for SU(1,1)/K Lie group is given in Appendix A of this document.

5.1. Poincaré Unit Disk and SU(1,1) Lie Group

The group of complex unimodular pseudo-unitary matrices $SU(1,1)$, is the set of elements u such that [43–52]:

$$uMu^+ = M \text{ with } M = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix} \tag{73}$$

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

$$G = SU(1,1) = \left\{ \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} / |a|^2 - |b|^2 = 1, \ a, b \in C \right\} \tag{74}$$

Its Cartan decomposition is given by:

$$\begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} = |a| \begin{pmatrix} 1 & z \\ z^* & 1 \end{pmatrix} \begin{pmatrix} a/|a| & 0 \\ 0 & a^*/|a| \end{pmatrix} \text{ with } z = b(a^*)^{-1}, |a| = \left(1 - |z|^2\right)^{-1/2} \tag{75}$$

$$\begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} \begin{pmatrix} 1 & z \\ z^* & 1 \end{pmatrix} = |a'| \begin{pmatrix} 1 & z' \\ z'^* & 1 \end{pmatrix} \begin{pmatrix} a'/|a'| & 0 \\ 0 & a'^*/|a'| \end{pmatrix} \text{ with } \begin{cases} a' = bz^* + a \\ z' = \frac{az+b}{b^*z+a^*} \end{cases} \tag{76}$$

$SU(1,1)$ is associated to group of holomorphic automorphisms of the Poincaré unit disk $D = \{z = x + iy \in C / |z| < 1\}$ in the complex plane, by considering its action on the disk as $g(z) = (az + b)/(b^*z + a^*)$. The following measure on Unit disk:

$$d\mu_0(z, z^*) = \frac{1}{2\pi i} \frac{dz \wedge dz^*}{\left(1 - |z|^2\right)^2} \tag{77}$$

is invariant under the action of $SU(1,1)$ captured by the fractional holomorphic transformation:

$$\frac{dz' \wedge dz'^*}{\left(1 - |z'|^2\right)^2} = \frac{dz \wedge dz^*}{\left(1 - |z|^2\right)^2} \tag{78}$$

The complex unit disk admits a Kähler structure determined by potential function:

$$\Phi(z', z^*) = -\log(1 - z'z^*) \tag{79}$$

The invariant 2-form is:

$$\Omega = \frac{1}{i} \frac{\partial^2 \Phi(z, z^*)}{\partial z \partial z^*} dz \wedge dz^* = \frac{1}{i} \frac{dz \wedge dz^*}{\left(1 - |z|^2\right)^2} \tag{80}$$

which is closed $d\Omega = 0$. This group $SU(1,1)$ is isomorphic to the group $SL(2,R)$ as a real Lie group, and the Lie algebra $\mathsf{g} = su(1,1)$ is given by:

$$g = \left\{ \begin{pmatrix} -ir & \eta \\ \eta^* & ir \end{pmatrix} / r \in R, \eta \in C \right\} \tag{81}$$

with the bases $(u_1, u_2, u_3) \in \mathsf{g}$: $u_1 = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, u_2 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, u_3 = \frac{1}{2} \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$

with the commutation relation:

$$[u_3, u_2] = u_1, [u_3, u_1] = -u_2, [u_2, u_1] = -u_3 \tag{82}$$

Dual base on the dual space of the Lie algebra is named $\left(u_1^*, u_2^*, u_3^*\right) \in g^*$. The dual vector space $g^* = su^*(1,1)$ can be identified with the subspace of $sl(2,C)$ of the form:

$$g^* = \left\{ \begin{pmatrix} z & x+iy \\ -x+iy & -z \end{pmatrix} = x \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} + z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} / x, y, z \in R \right\} \tag{83}$$

Coadjoint action of $g \in G$ on dual space of the Lie algebra $\xi \in g^*$ is written $g.\xi$.

5.2. Coadjoint Orbit of SU(1,1) and Souriau Moment Map

We will use results of C. Cishahayo and S. de Bièvre [53] and B. Cahen [54,55] for computation of moment map of $SU(1,1)$. Let $r \in R^{*+}$, orbit $O\left(ru_3^*\right)$ of ru_3^* for the coadjoint action of $g \in G$ could be identified with the upper half sheet $x_3 > 0$ of $\left\{\xi = x_1 u_1^* + x_2 u_2^* + x_3 u_3^* / -x_1^2 - x_2^2 + x_3^2 = r^2\right\}$, the two-sheet hyperboloid. The stabilizer of ru_3^* for the coadjoint action of G is torus $K = \left\{ \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}, \theta \in R \right\}$. K induces rotations of the unit disk, and leaves 0 invariant. The stabilizer for the origin 0 of unit disk is maximal compact subgroup K of *SU(1,1)*. We can observe [54] that $O(ru_3^*) = G/K$. On the other hand $O\left(ru_3^*\right) = G/K$ is diffeomorphic to the unit disk $D = \{z \in C/|z| < 1\}$, then by composition, the Souriau moment map is given by:

$$
\begin{aligned}
J &: D \to O\left(ru_3^*\right) \\
z &\mapsto J(z) = r\left(\frac{z+z^*}{(1-|z|^2)}u_1^* + \frac{z-z^*}{i(1-|z|^2)}u_2^* + \frac{1+|z|^2}{(1-|z|^2)}u_3^*\right)
\end{aligned}
\tag{84}
$$

J is linked to the natural action of G on D (by fractional linear transforms) but also the coadjoint action of G on $O\left(ru_3^*\right) = G/K$. J^{-1} could be interpreted as the stereographic projection from the two-sphere S^2 onto $C \cup \infty$ [56]. In case $r = \frac{n}{2}$ where $n \in N^+, n \geq 2$ then the coadjoint orbit is given by $O_n = O(\zeta_n)$ with $\xi_n = \frac{n}{2}u_3^* \in g^*$, with stabilizer of ξ_n for coadjoint action the torus $K = \left\{ \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}, \theta \in R \right\}$ with Lie algebra Ru_3. $O_n = O(\zeta_n)$ is associated with a holomorphic discrete series representation π_n of G by the KKS (Kirillov-Kostant-Souriau) method of orbits.

$$
\begin{aligned}
J &: D \to O_n \\
z &\mapsto J(z) = \frac{n}{2}\left(\frac{z+z^*}{(1-|z|^2)}u_1^* + \frac{z-z^*}{i(1-|z|^2)}u_2^* + \frac{1+|z|^2}{(1-|z|^2)}u_3^*\right)
\end{aligned}
\tag{85}
$$

Group G act on D by homography $g.z = \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix}.z = \frac{az+b}{b^*z+a^*}$. **This action corresponds with coadjoint action of G on O_n.** The Kirillov-Kostant-Souriau 2-form of O_n is given by:

$$\Omega_n(\zeta)(X(\zeta), Y(\zeta)) = \langle \zeta, [X, Y] \rangle, \; X, Y \in g \text{ and } \zeta \in O_n \tag{86}$$

and is associated in the frame by J with:

$$\omega_n = \frac{in}{\left(1-|z|^2\right)^2} dz \wedge dz^* \tag{87}$$

with the corresponding Poisson Bracket:

$$\{f, g\} = i\left(1 - |z|^2\right)^2\left(\frac{\partial f}{\partial z}\frac{\partial g}{\partial z^*} - \frac{\partial f}{\partial z^*}\frac{\partial g}{\partial z}\right) \tag{88}$$

It has been also observed that there are 3 basic observables generating the $SU(1,1)$ symmetry on classical level:

$$\begin{cases} D \to R \\ z \mapsto k_3(z) = \frac{1+|z|^2}{1-|z|^2} \end{cases}, \begin{cases} D \to R \\ z \mapsto k_1(z) = \frac{1}{i}\frac{z-z^*}{1-|z|^2} \end{cases}, \begin{cases} D \to R \\ z \mapsto k_2(z) = \frac{z+z^*}{1-|z|^2} \end{cases} \tag{89}$$

with the Poisson commutation rule:

$$\{k_3, k_1\} = k_2, \{k_3, k_2\} = -k_1, \{k_1, k_2\} = -k_3 \tag{90}$$

(k_1, k_2, k_3) vector points to the upper sheet of the two-sheeted hyperboloid in R^3 given by $k_3^2 - k_1^2 - k_2^2 = 1$, whose the stereographic projection onto the open unit disk is:

$$\begin{cases} (k_1, k_2, k_3) \in \mathrm{H}^+ \to D \\ z = \frac{k_2 + ik_1}{1 + k_3} = \sqrt{\frac{k_3 - 1}{k_3 + 1}}e^{i\mathrm{arg}z} \end{cases} \tag{91}$$

Under the action of $g \in G = SU(1,1) = \left\{\begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix}/|a|^2 - |b|^2 = 1, \; a, b \in C\right\}$:

$$\begin{pmatrix} k_- & k_3 \\ k_3 & k_+ \end{pmatrix} = \begin{pmatrix} k_2 + ik_1 & k_3 \\ k_3 & k_2 - ik_1 \end{pmatrix} = \frac{1}{1-|z|^2}\begin{pmatrix} 2z & 1 + |z|^2 \\ 1 + |z|^2 & 2z^* \end{pmatrix} \text{ is transform in:}$$

$$\begin{pmatrix} k'_- & k'_3 \\ k'_3 & k'_+ \end{pmatrix} = \begin{pmatrix} k_-\left(g^{-1}.z\right) & k_3\left(g^{-1}.z\right) \\ k_3\left(g^{-1}.z\right) & k_+\left(g^{-1}.z\right) \end{pmatrix} = g^{-1}\begin{pmatrix} k_- & k_3 \\ k_3 & k_+ \end{pmatrix}\left(g^{-1}\right)^t \tag{92}$$

This transform can be viewed as the co-adjoint action of $SU(1,1)$ ***on the coadjoint orbit identified with*** $k_3^2 - k_1^2 - k_2^2 = 1$. We can also observe that the quotient $SU(1,1)/K$ is isomorphic to the upper sheet of the hyperboloid described by $k_3^2 - k_1^2 - k_2^2 = 1$, by the following parametrization (τ, φ), given by $\vec{n} = (\cosh\tau, \sinh\tau\cos\varphi, \sinh\tau\sin\varphi)$, and its stereographic projection onto the inside of the unit disk, parametrized by $\varsigma = \tanh\frac{\tau}{2}e^{-i\varphi}$.

6. Covariant Gibbs Density by Souriau Thermodynamics for Poincaré Unit Disk

6.1. Fourier Transform, Laplace Transform and Lie Group Representation Theory

In Souriau Lie Group Thermododynamic, we have to consider Laplace Transform defined on coadjoint orbits to define Massieu Potential Function and Gibbs density. This problem has been solved in the domain of Kirillov Representation Theory. Representation theory studies abstract algebraic structures by representing their elements as linear transformations of vector spaces, and algebraic objects (Lie groups, Lie algebras) by describing its elements by matrices and the algebraic operations in terms of matrix addition and matrix multiplication, reducing problems of abstract algebra to problems in linear algebra. Representation theory generalizes Fourier analysis via harmonic analysis. The modern development of Fourier analysis during XXth century has explored the generalization of Fourier and Fourier-Plancherel formula for non-commutative harmonic analysis, applied to locally compact non-Abelian groups. This has been solved by geometric approaches based on "orbits methods" (Fourier-Plancherel formula for G is given by coadjoint representation of G in dual vector space of its Lie algebra) with many contributors (Dixmier, Kirillov, Bernat, Arnold, Berezin, Kostant, Souriau, Duflo, Guichardet, Torasso, Vergne, Paradan, etc.) [57–68].

For classical commutative harmonic analysis, we consider the following groups:

$$G = \mathrm{T}^n = R^n / Z^n \text{ for Fourier series, } G = R^n \text{ for Fourier Transform}$$

G group character (linked to e^{ikx}) : $\chi : G \to U$ with $U = \{z \in C / |z| = 1\}$

$$\hat{G} = \{\chi / \chi_1 . \chi_2(g) = \chi_1(g)\chi_2(g)\} \text{ and Fourier transform is given by :} \tag{93}$$

$$\varphi : G \to C \qquad\qquad \hat{\varphi} : \hat{G} \to C$$
$$g \mapsto \varphi(g) = \int_{\hat{G}} \hat{\varphi}(\chi)\chi(g)^{-1} d\chi \qquad \chi \mapsto \hat{\varphi}(\chi) = \int_{G} \varphi(g)\chi(g) dg$$

For non-commutative harmonic analysis, Group unitary irreducible representation is $U : G \to U(H)$ with H Hilbert space and character by $\chi_U(g) = tr U_g$. Fourier transform for non-commutative group is $U_\varphi = \int_G \varphi(g) U_g dg$ with character $\chi_U(g) = tr U_\varphi$. If we describe group element with exponential map $U_\psi = \int_g \psi(X) U_{\exp(X)} dX$, we have:

$$tr U_\psi = \dim \tau . \mu_{G.f}\left(\overset{\wedge}{\psi . j^{-1}}\right) \quad \textbf{with} \begin{cases} \mu_{G.f}\colon \text{ Liouville meas. on } O = G.f, f \in g^* \\ \mu_{G.f}\left(\overset{\wedge}{\psi . j^{-1}}\right)\colon \text{ Integral of } \overset{\wedge}{\psi . j^{-1}} \text{wrt } \mu_{G.f} \end{cases} \tag{94}$$

$$\overset{\wedge}{\psi . j^{-1}} : g \to g^*, \text{ Four. Transf.}$$

where

$$j(X) = (\det s(ad_X))^{1/2} \text{ with } s(x) = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!}\left(\frac{x}{2}\right)^{2n} = sh\left(\frac{x}{2}\right) / \left(\frac{x}{2}\right) \tag{95}$$

Kirillov Character formula is:

$$\chi_U(\exp(X)) = tr U_{\exp(X)} = j(X)^{-1} \int_O e^{i\langle f, X\rangle} d\mu_O(f) \tag{96}$$

$$\int_O e^{i\langle f, X\rangle} d\mu_O(f) = j(X) tr U_{\exp(X)} \text{ with } j(X) = \left(\det\left(\frac{e^{ad_X/2} - e^{-ad_X/2}}{ad_X/2}\right)\right)^{1/2} \tag{97}$$

We will use Kirillov representation theory and his character formula to compute Souriau covariant Gibbs density in the unit Poincaré disk. For any Lie group G, a coadjoint orbit $O \subset g^*$ has a canonical symplectic form ω_0 given by KKS 2-form. As seen, if G is finite dimensional, the corresponding volume element defines a G-invariant measure supported on O, which can be interpreted as a tempered distribution. The Fourier transform (where d is the half of the dimension of the orbit O):

$$\Im(x) = \int_{O \subset g^*} e^{-i\langle \lambda, x\rangle} \frac{1}{d!} d\omega_{O^d} \text{ with } \lambda \in g^* \text{ and } x \in g \tag{98}$$

is Ad G-invariant. When $O \subset g^*$ is an integral coadjoint orbit, Kirillov formula, given previously, expresses Fourier transform $\Im(x)$ by Kirillov character χ_O:

$$\Im(x) = j(x)\chi_O(e^x) \text{ where } j(x) = \det^{1/2}\left(\frac{\sinh(ad(x/2))}{ad(x/2)}\right) \tag{99}$$

χ_O is, as defined previously, the "*Kirillov character*" of a unitary representation associated to the orbit.

6.2. Souriau Covariant Gibbs Density in Poincaré Unit Disk for SU(1,1) Lie Group

In the following, we will give the full development to compute the Souriau covariant Gibbs density. As the Gibbs density is not defined for all geometric temperature, as observed by Souriau,

we have used his approach by considering a one-parameter subgroup of the Lie group generated by exponential map from a one element of Lie algebra given by geometric temperature. The subset of Lie algebra where the Gibbs density is deduced from the contraints related to this one-parameter subgroup generation.

Considering the Lie group $SU(1,1) = \left\{ \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} / a, b \in C, |a|^2 - |b|^2 = 1 \right\}$ and its Lie algebra given by elements $su(1,1) = \left\{ \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix} / r \in R, \eta \in C \right\}$. A basis for this Lie algebra $su(1,1)$ is $(u_1, u_2, u_3) \in$ g with $u_1 = \frac{i}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, u_2 = -\frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $u_3 = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ with $[u_1, u_3] = -u_2, [u_1, u_2] = u_3, [u_2, u_3] = -u_1$.

The compact subgroup is generated by u_1, while u_2 and u_3 generate a hyperbolic subgroup. The dual space of the Lie algebra is given by $su(1,1)^* = \left\{ \begin{pmatrix} z & x + iy \\ -x + iy & -z \end{pmatrix} / x, y, z \in R \right\}$ with the basis $(u_1^*, u_2^*, u_3^*) \in g^*$ with $u_1^* = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, u_2^* = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$ and $u_3^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Let consider $D = \{z \in C / |z| < 1\}$ be the open unit disk of Poincaré. For each $\rho > 0$, the pair (D, ω_ρ) is a symplectic homogeneous manifold with $\omega_\rho = 2i\rho \frac{dz \wedge dz^*}{(1 - |z|^2)^2}$, where ω_ρ is invariant under the action:

$$SU(1,1) \times D \to D$$
$$(g, z) \mapsto g.z = \frac{az + b}{b^* z + a^*} .$$

This action is transitive and is globally and strongly Hamiltonian. Its generators are the hamiltonian vector fields associated to the functions:

$$J_1(z, z^*) = \rho \frac{1 + |z|^2}{1 - |z|^2}, \; J_2(z, z^*) = \frac{\rho}{i} \frac{z - z^*}{1 - |z|^2}, \; J_3(z, z^*) = -\rho \frac{z + z^*}{1 - |z|^2} \tag{100}$$

The associated moment map $J : D \to su^*(1,1)$ defined by $J(z).u_i = J_i(z, z^*)$, maps D into a coadjoint orbit in $su^*(1,1)$. Then, we can write the moment map as a matrix element of $su^*(1,1)$:

$$J(z) = J_1(z, z^*)u_1^* + J_2(z, z^*)u_2^* + J_3(z, z^*)u_3^* = \begin{pmatrix} \rho \frac{1 + |z|^2}{1 - |z|^2} & \rho \frac{z - z^*}{1 - |z|^2} - \rho \frac{z + z^*}{1 - |z|^2} \\ \rho \frac{z - z^*}{1 - |z|^2} + \rho \frac{z + z^*}{1 - |z|^2} & -\rho \frac{1 + |z|^2}{1 - |z|^2} \end{pmatrix}$$

$$J(z) = J_1(z, z^*)u_1^* + J_2(z, z^*)u_2^* + J_3(z, z^*)u_3^* = \rho \begin{pmatrix} \frac{1 + |z|^2}{1 - |z|^2} & -2\frac{z^*}{1 - |z|^2} \\ 2\frac{z}{1 - |z|^2} & -\frac{1 + |z|^2}{1 - |z|^2} \end{pmatrix} \in g^* \tag{101}$$

The moment map J is a diffeomorphism of D onto one sheet of the two-sheeted hyperboloid in $su^*(1,1)$, determined by $J_1^2 - J_2^2 - J_3^2 = \rho^2$, $J_1 \geq \rho$ with $J_1 u_1^* + J_2 u_2^* + J_3 u_3^* \in su^*(1,1)$. We note O_ρ^+ the coadjoint orbit $Ad^*_{SU(1,1)}$ of $SU(1,1)$, given by the upper sheet of the two-sheeted hyperboloid given by previous equation. The orbit method of Kostant-Kirillov-Souriau associates to each of these coadjoint orbits a representation of the discrete series of $SU(1,1)$, provided that ρ is a half integer greater or equal than 1 ($\rho = \frac{k}{2}, k \in N$ and $\rho \geq 1$). When explicitly executing the Kostant-Kirillov construction, the representation Hilbert spaces H_ρ are realized as closed reproducing kernel subspaces of $L^2(D, \omega_\rho)$. The Kostant-Kirillov-Souriau orbit method shows that to each coadjoint orbit of a connected Lie group is associated a unitary irreducible representation of G acting in a Hilbert space H.

Souriau has oberved that action of the full Galilean group on the space of motions of an isolated mechanical system is not related to any equilibrium Gibbs state (the open subset of the Lie algebra, associated to this Gibbs state is empty). The main Souriau idea was to define the Gibbs states for one-parameter subgroups of the Galilean group. We will use the same approach, in this case We will consider action of the Lie group $SU(1,1)$ on the symplectic manifold (M, ω) (Poincaré unit disk)

and its momentum map J are such that the open subset $\Lambda_\beta = \left\{ \beta \in g / \int_D e^{-\langle J(z), \beta \rangle} d\lambda(z) < +\infty \right\}$ is not empty. This condition is not always satisfied when (M, ω) is a cotangent bundle, but of course it is satisfied when it is a compact manifold. The idea of Souriau is to consider a one parameter subgroup of $SU(1,1)$. To parametrize elements of $SU(1,1)$ is through its Lie algebra. In the neighborhood of the identity element, the elements of $g \in SU(1,1)$ can be written as the exponential of an element β of its Lie algebra:

$$g = \exp(\varepsilon\beta) \text{ with } \beta \in g \tag{102}$$

The condition $g^+ M g = M$ for $M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ can be expanded for $\varepsilon << 1$ and is equivalent to

$\beta^+ M + M\beta = 0$ which then implies $\beta = \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix}$, $r \in R, \eta \in C$. We can observe that r and $\eta = \eta_R + i\eta_I$

contain 3 degrees of freedom, as required. Also because $\det g = 1$, we get $Tr(\beta) = 0$. We can then exponentiate β with exponential map to get:

$$g = \exp(\varepsilon\beta) = \sum_{k=0}^{\infty} \frac{(\varepsilon\beta)^k}{k!} = \begin{pmatrix} a_\varepsilon(\beta) & b_\varepsilon(\beta) \\ b_\varepsilon^*(\beta) & a_\varepsilon^*(\beta) \end{pmatrix} \tag{103}$$

If we make the remark that $\beta^2 = \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix} \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix} = \left(|\eta|^2 - r^2 \right) I$, we can developed the exponential map:

$$g = \exp(\varepsilon\beta) = \begin{pmatrix} \cosh(\varepsilon R) + ir\frac{\sinh(\varepsilon R)}{R} & \eta\frac{\sinh(\varepsilon R)}{R} \\ \eta^*\frac{\sinh(\varepsilon R)}{R} & \cosh(\varepsilon R) - ir\frac{\sinh(\varepsilon R)}{R} \end{pmatrix} \text{ with } R^2 = |\eta|^2 - r^2 \tag{104}$$

We can observe that one condition is that $|\eta|^2 - r^2 > 0$ then the subset to consider is $\Lambda_\beta = \left\{ \beta = \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix}, r \in R, \eta \in C / |\eta|^2 - r^2 > 0 \right\}$ such that $\int_D e^{-\langle J(z), \beta \rangle} d\lambda(z) < +\infty$. The generalized Gibbs states of the full $SU(1,1)$ group do not exist. However, generalized Gibbs states for the one-parameter subgroups $\exp(\alpha\beta)$, $\beta \in \Lambda_\beta$, of the $SU(1,1)$ group do exist. The generalized Gibbs state associated to β remains invariant under the restriction of the action to the one-parameter subgroup of $SU(1,1)$ generated by $\exp(\varepsilon\beta)$.

To futher, we will develop the Souriau Gibbs density from the Souriau moment map $J(z)$ and the Souriau temperature $\beta \in \Lambda_\beta$. If we note $b = \frac{1}{1-|z|^2} \begin{bmatrix} 1 \\ -z \end{bmatrix}$, we can write the moment map:

$$J(z) = \rho \begin{pmatrix} \frac{1+|z|^2}{1-|z|^2} & -2\frac{z^*}{1-|z|^2} \\ 2\frac{z}{1-|z|^2} & -\frac{1+|z|^2}{1-|z|^2} \end{pmatrix} = \rho\left(2Mbb^+ - Tr\left(Mbb^+ \right)I \right) \text{ with } M = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{105}$$

We can the write the covariant Gibbs density in the unit disk given by moment map of the Lie group $SU(1,1)$ and geometric temperature in its Lie algebra $\beta \in \Lambda_\beta$:

$$p_{Gibbs}(z) = \frac{e^{-\langle J(z), \beta \rangle}}{\int_D e^{-\langle J(z), \beta \rangle} d\lambda(z)} \text{ where } d\lambda(z) = 2i\rho \frac{dz \wedge dz^*}{\left(1 - |z|^2 \right)^2} \tag{106}$$

$$p_{Gibbs}(z) = \frac{e^{-\langle \rho(2\Im bb^+ - Tr(\Im bb^+)I), \beta \rangle}}{\int_D e^{-\langle J(z), \beta \rangle} d\lambda(z)} = \frac{e^{-\left\langle \rho\left(\begin{smallmatrix} \frac{1+|z|^2}{(1-|z|^2)} & \frac{-2z^*}{(1-|z|^2)} \\ \frac{2z}{(1-|z|^2)} & -\frac{1+|z|^2}{(1-|z|^2)} \end{smallmatrix} \right), \left(\begin{smallmatrix} ir & \eta \\ \eta^* & -ir \end{smallmatrix} \right) \right\rangle}}{\int_D e^{-\langle J(z), \beta \rangle} d\lambda(z)} \tag{107}$$

To write the Gibbs density with respect to its statistical moments, we have to express the density with respect to $Q = E[J(z)]$. Then, we have to invert the relation between Q and β, to replace this last variable $\beta = \begin{pmatrix} ir & \eta \\ \eta^* & -ir \end{pmatrix} \in \Lambda_\beta$ by $\beta = \Theta^{-1}(Q) \in g$ where $Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \Theta(\beta) \in g^*$ with $\Phi(\beta) = -\log \int_D e^{-\langle J(z), \beta \rangle} d\lambda(z)$, deduce from Legendre tranform. The mean moment map is given by:

$$Q = E[J(z)] = E\left[\rho \left(\begin{matrix} \frac{1+|w|^2}{(1-|w|^2)} & \frac{-2w^*}{(1-|w|^2)} \\ \frac{2w}{(1-|w|^2)} & -\frac{1+|w|^2}{(1-|w|^2)} \end{matrix} \right) \right] \text{ where } w \in D \tag{108}$$

This mean moment map can be obtained by Karcher mean computation on the one-sheet hyperboloid corresponding to the coadjoint orbit. For the dual pairing, we can observed that $J(z) = J_1(z, z^*)u_1^* + J_2(z, z^*)u_2^* + J_3(z, z^*)u_3^* \in g^*$ with $J_1(z, z^*) = \rho\frac{1+|z|^2}{1-|z|^2}, J_2(z, z^*) = \frac{\rho}{i}\frac{z-z^*}{1-|z|^2}, J_3(z, z^*) = -\rho\frac{z+z^*}{1-|z|^2}$ and $\beta = \beta_1 u_1 + \beta_2 u_2 + \beta_3 u3 \in g^*$ with $\beta = 2(r, -\eta_R, -\eta_I), \eta = \eta_R + i\eta_I$.

The integral of normalization in Gibbs density could be computed through Kirillov character formula by $\chi_m\left(\exp\left(\begin{pmatrix} x & . \\ . & -x \end{pmatrix} \right) \right) = j(x)^{-1} \int_{O_{m-1}^+} e^{-i \left\langle \left(\begin{smallmatrix} x & . \\ . & -x \end{smallmatrix} \right), \left(\begin{smallmatrix} ir & \eta \\ \eta^* & -ir \end{smallmatrix} \right) \right\rangle} \omega_{O_{m-1}^+}$ where

$$j(x) = \det^{1/2}\left[\sinh\left(ad\left(\begin{matrix} x/2 & \\ & -x/2 \end{matrix} \right) \right) / ad\left(\begin{matrix} x/2 & \\ & -x/2 \end{matrix} \right) \right] = \frac{\sinh(x)}{x}$$

with following relation $\frac{e^{mx}}{1-e^{2x}} j(x) = \int_D e^{(m-1)x\frac{1+|w|^2}{1-|w|^2}} \frac{1}{(1-|w|^2)^2} dw \wedge dw^*$.

Recently, Enrico De Micheli [69] has introduced a Laplace-type transform (the so-called Spherical Laplace Transform) with a connection to the Non-Euclidean Fourier Transform in the sense of Helgason, and the principal series of the unitary representation of *SU(1,1)*.

6.3. Extension to SU (p,q) Unitary Group for Siegel Unit Disk

Mode details are given in Appendix B, on parameterization of SU(1,1) and extension to SU (p,q). To address computation of covariant Gibbs density for Siegel Unit Disk, we will consider in this section $SU(p, q)$ Unitary Group:

$$G = SU(p, q) \text{ and } K = S(U(p) \times U(q)) = \left\{ \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} / A \in U(p), D \in U(q), \det(A)\det(D) = 1 \right\} \tag{109}$$

We can use the following decomposition for $g \in G^C$:

$$g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in G^C, g = \begin{pmatrix} I_p & BD^{-1} \\ 0 & I_q \end{pmatrix} \begin{pmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I_p & 0 \\ D^{-1}C & I_q \end{pmatrix} \tag{110}$$

and consider the action of $g \in G^C$ on Siegel Unit Disk $SD = \left\{ Z \in M_{pq}(C) / I_p - ZZ^+ > 0 \right\}$ given by:

$$g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in G^C, g = \begin{pmatrix} I_p & BD^{-1} \\ 0 & I_q \end{pmatrix} \begin{pmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I_p & 0 \\ D^{-1}C & I_q \end{pmatrix} \tag{111}$$

Benjamin Cahen has study this case and introduced the moment map by identifing G-equivariantly g^* with g by means of the Killing form β on g^C:

$$g^* \, G - \text{equivariant with gby Killing form } \beta(X,Y) = 2(p+q)Tr(XY)$$

The set of all elements of g fixed by K is h:

$$h = \{\text{element of } G \text{ fixed by } K\}, \xi_0 \in h, \xi_0 = i\lambda \begin{pmatrix} -qI_p & 0 \\ 0 & pI_q \end{pmatrix} \tag{112}$$
$$\Rightarrow \langle \xi_0, [Z, Z^+] \rangle = -2i\lambda(p+q)^2 Tr(ZZ^+), \forall Z \in D$$

Then, we the equivatiant moment map is given by:

$$\forall X \in g^C, \ Z \in D, \psi(Z) = Ad^*(\exp(-Z^+)\zeta(\exp Z^+ \exp Z))\xi_0$$
$$\forall g \in G, Z \in D \text{ then } \psi(g.Z) = Ad_g^* \psi(Z) \tag{113}$$
$$\psi \text{ is a diffeomorphism from } SD \text{ onto orbit } O(\xi_0)$$

with:

$$\psi(Z) = i\lambda \begin{pmatrix} \left(I_p - ZZ^+\right)^{-1}\left(-pZZ^+ - qI_p\right) & (p+q)Z\left(I_q - Z^+Z\right)^{-1} \\ -(p+q)\left(I_q - Z^+Z\right)^{-1}Z^+ & \left(pI_q + qZ^+Z\right)\left(I_q - Z^+Z\right)^{-1} \end{pmatrix} \tag{114}$$

$$\zeta\left(\exp Z^+ \exp Z\right) = \begin{pmatrix} I_p & Z\left(I_q - Z^+Z\right)^{-1} \\ 0 & I_q \end{pmatrix} \tag{115}$$

7. Lie Groups Thermodynamics for SE(2) Lie Group

After $SU(1,1)$ Lie group with null cohomology and then without Souriau one-cocycle, we will consider Souriau model for $SE(2)$ Lie group with non-null cohomology and then with introduction of Souriau one-cocycle [70].

We will consider first $SO(2)$ Lie group:

$$SO(2) = \left\{ R_\varphi = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} / \varphi \in R \right\} \tag{116}$$

A vector at the identity to $SO(2)$ is given by:

$$\frac{dR_{t\eta}}{dt}\bigg|_{t=0} = -\eta \mathfrak{I} \text{ with } \mathfrak{I} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \mathfrak{I}^T = \mathfrak{I}^{-1} = -\mathfrak{I} \tag{117}$$

We consider the special Euclidean group $SE(2) = SO(2) \times R^2$.

$$SE(2) = \left\{ \begin{bmatrix} R_\varphi & \tau \\ 0 & 1 \end{bmatrix} / R_\varphi \in SO(2), \tau \in R^2 \right\} \tag{118}$$

the group operation is given by:

$$\begin{bmatrix} R_{\varphi_1} & \tau_1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} R_{\varphi_2} & \tau_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_{\varphi_1}R_{\varphi_2} & R_{\varphi_1}\tau_2 + \tau_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_{\varphi_1+\varphi_2} & R_{\varphi_1}\tau_2 + \tau_1 \\ 0 & 1 \end{bmatrix} \tag{119}$$
$$\Rightarrow (R_1,\tau_1).\big(R_{\varphi_2},\tau_2\big) = \big(R_{\varphi_1+\varphi_2}, R_{\varphi_1}\tau_2 + \tau_1\big)$$

$$\begin{bmatrix} R_{\varphi_1} & \tau_1 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} R_{-\varphi_1} & -R_{-\varphi_1}\tau_1 \\ 0 & 1 \end{bmatrix} \Rightarrow \big(R_{\varphi_1},\tau_1\big)^{-1} = \big(R_{-\varphi_1}, -R_{-\varphi_1}\tau_1\big) \tag{120}$$

The Lie algebra $se(2)$ of $SE(2)$ has underlying vector space R^3 and Lie bracket:

$$(\xi,u) \in se(2) = R \times R^2 \Rightarrow \begin{bmatrix} -\xi\Im & u \\ 0 & 0 \end{bmatrix} \in se(2) \tag{121}$$

Lie bracket is given by:

$$[(\xi,u),(\eta,v)] = (0, \xi\Im v + \eta\Im u) \tag{122}$$

Adjoint action of $SE(2)$ is given by:

$$Ad_{(R_\varphi,\tau)}(\xi,u) = \begin{bmatrix} R_\varphi & \tau \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -\xi\Im & u \\ 0 & 0 \end{bmatrix}\begin{bmatrix} R_{-\varphi} & -R_{-\varphi}\tau \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -\xi\Im & \xi\Im\tau + R_\varphi u \\ 0 & 0 \end{bmatrix} \tag{123}$$
$$Ad_{(R_\varphi,\tau)}(\xi,u) = \big(\xi, R_\varphi u + \xi\Im\tau\big)$$

Coadjoint action of $SE(2)$ is given by:

$$Ad^*_{(R_\varphi,\tau)}(m,\rho) = \big(m + \Im R_\varphi\rho.\tau, R_\varphi\rho\big) \tag{124}$$

The moment map $J : R^2 \to se^*(2)$ of $SE(2)$ is defined by:

$$J_{(\xi,u)}(x) = J(x).(\xi,u) \tag{125}$$

with the right action of $SE(2)$ on R^2:

$$x.(R\varphi,\tau) = R_{-\varphi}(x - \tau) \tag{126}$$

the infinitesimal generator of $(\xi,u) \in se(2)$ has the expression:

$$(\xi,u)_{R^2}(x) = \left.\frac{d[x.(R_{t\xi}, tu)]}{dt}\right|_{t=0} = \left.\frac{d[R_{-t\xi}(x - tu)]}{dt}\right|_{t=0} = \xi\Im x - u \tag{127}$$

Let $J_{(\xi,u)}(x) : R^2 \to se^*(2)$ be the moment map of this action relative to the symplectic form, we can compute it from its definition:

$$\begin{aligned}
&dJ_{(\xi,u)}(x).y = -2\omega((\xi,u)_{R^2}, y) \\
&\text{with } \omega((\xi,u)_{R^2}, y) = \omega(\xi\Im x - u, y) = (\xi\Im x - u).\Im y = (\xi x + \Im u).y \\
&\Rightarrow dJ_{(\xi,u)}(x).y = -2(\xi x + \Im u).y \\
&\Rightarrow J_{(\xi,u)}(x) = -2\big(\tfrac{1}{2}\xi\|x\|^2 + \Im u.x\big) = -2\big(\tfrac{1}{2}\|x\|^2, -\Im x\big).(\xi,u) \\
&J_{(\xi,u)}(x) = J(x).(\xi,u) \Rightarrow J(x) = -2\big(\tfrac{1}{2}\|x\|^2, -\Im x\big), x \in R^2
\end{aligned} \tag{128}$$

We then compute the one-cocycle of $SE(2)$ from the moment map:

$$\theta\big((R_{\varphi,\tau})\big) = J\big(0.(R_\varphi,\tau)\big) - Ad^*_{(R_{\varphi,\tau})} J(0) = J\big(-R_{-\varphi}\tau\big)$$
$$\theta\big((R_{\varphi,\tau})\big) = -2\big(\tfrac{1}{2}\|\tau\|^2, \Im R_{-\varphi}\tau\big) = -2\big(\tfrac{1}{2}\|\tau\|^2, R_{-\varphi-\frac{\pi}{2}}\tau\big)$$

(129)

Coadjoint orbit of $SE(2)$ are generated by:

$$O_{(m,\rho)} = \Big\{ A^*_{(R_{\varphi,\tau})}(m,\rho) + \theta\big((R_\varphi,\tau)\big) / \big(R_\varphi,\tau\big) \in SE(2) \Big\}$$
$$O_{(m,\rho)} = \Big\{ \big(x - R_{-\frac{\pi}{2}}\rho.\tau - \|\tau\|^2, R_{-\varphi}\rho - 2R_{-\varphi-\frac{\pi}{2}}\tau \big) / \big(R_\varphi,\tau\big) \in SE(2) \Big\}$$

(130)

The Souriau Symplectic form in this case of non-null cohomology is given by:

$$\omega_{(m,\rho)(m',\rho')}\Big(ad^*_{(\xi,u)}(m',\rho') - (0,2\Im u), ad^*_{(\eta,v)}(m',\rho') - (0,2\Im v) \Big) = \rho'.(-\xi\Im v + \eta\Im u) + 2u.\Im v$$
$$\text{with } (m',\rho') = \big(x - R_{-\frac{\pi}{2}}\rho.\tau - \|\tau\|^2, R_{-\varphi}\rho - 2R_{-\varphi-\frac{\pi}{2}}\tau \big) \in O_{(m,\rho)} \subset R^3$$

(131)

With the expression of moment map, we can compute Souriau covariant Gibbs density of Maximum Entropy.

Considering the symplectic form $\omega(\zeta, v) = \zeta.\Im v$ with $\Im = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ on R^2, we have seen that the action of $SE(2)$ is symplectic and admits the momentum map, $J(x) = -\big(\tfrac{1}{2}\|x\|^2, -\Im x \big)$, $x \in R^2$.

Souriau Gibbs density is defined for generalized temperature $\beta \in \Omega = \big\{ (b,B) \in se(2) / b < 0, B \in R^2 \big\}$ and given by:

$$p_{Gibbs}(x) = \frac{e^{-\langle J(x),\beta\rangle}}{\displaystyle\int_{R^2} e^{-\langle J(x),\beta\rangle} d\lambda(x)} = \frac{e^{\frac{1}{2}b\|x\|^2 - B.\Im x}}{\displaystyle\int_{R^2} e^{\frac{1}{2}b\|x\|^2 - B.\Im x} d\lambda(x)}$$

(132)

The Massieu Potential could be computed:

$$\Phi(\beta) = \log \int_{R^2} e^{\frac{1}{2}b\|x\|^2 - B.\Im x} d\lambda(x) = \log\Big(-\frac{2\pi}{b} e^{-\frac{1}{2b}\|B\|^2} \Big)$$

(133)

By derivation of Massieu potential, we can deduce expression of Heat:

$$Q \in \Omega^* = \Big\{ (m,M) \in se^*(2) / m + \tfrac{\|M\|^2}{2} < 0 \Big\}$$
$$Q = \frac{\partial \Phi(\beta)}{\partial \beta} = \Big(\tfrac{1}{b} - \tfrac{\|B\|^2}{2b^2}, \tfrac{1}{b}B \Big) = \Theta(\beta)$$

(134)

We can the inverse this relation to express generalized temperature with respect to the heat:

$$\beta = \Theta^{-1}(Q) = \Big(\big(m + \tfrac{1}{2}\|M\|^2\big)^{-1}, \big(m + \tfrac{1}{2}\|M\|^2\big)^{-1} M \Big)$$

(135)

We can the express the Gibbs density with respect to the Heat Q which is the mean of moment map:

$$p_{Gibbs}(x) = \frac{e^{\frac{\frac{1}{2}\|x\|^2 - M.\Im x}{(m + \frac{1}{2}\|M\|^2)}}}{\displaystyle\int_{R^2} e^{\frac{\frac{1}{2}\|x\|^2 - M.\Im x}{(m + \frac{1}{2}\|M\|^2)}} d\lambda(x)} \quad \text{with } (m,M) = E(J(x)) = E\Big[-2\big(\tfrac{1}{2}\|x\|^2, -\Im x\big) \Big] = \Big[-E\big(\|x\|^2\big), 2\Im E(x) \Big]$$

(136)

So we can rewrite the Gibbs density:

$$p_{Gibbs}(x) = \frac{e^{\frac{\frac{1}{2}\|x\|^2 + 2E(x).Ix}{(-E(\|x\|^2) + 2\|E(x)\|^2)}}}{\int\limits_{R^2} e^{\frac{\frac{1}{2}\|x\|^2 + 2E(x).Ix}{(-E(\|x\|^2) + 2\|E(x)\|^2)}} d\lambda(x)} \tag{137}$$

We can also provide a Fisher metric in dual Lie algebra as hessian of the Entropy:

$$S(Q) = \langle Q, \beta \rangle - \Phi(\beta) = 1 + \log(2\pi) + \log\left(-m - \frac{\|M\|^2}{2}\right) \tag{138}$$

$$I_{Fisher}(Q) = \left(m + \frac{1}{2}\|M\|^2\right)^{-1} \begin{bmatrix} I & M^T \\ M^T & \frac{1}{2}\|M\|^2 - m \end{bmatrix} \tag{139}$$

and as $(m, M) = E(J(x)) = E\left[-2\left(\frac{1}{2}\|x\|^2, -\Im x\right)\right] = \left[-E(\|x\|^2), 2\Im E(x)\right]$, Fisher metric in dual space of Lie Algebra parameterization could be written:

$$I_{Fisher}(Q) = \left(2\|E(x)\|^2 - E(\|x\|^2)\right)^{-1} \begin{bmatrix} I & (2\Im E(x))^T \\ 2\Im E(x) & 2\|E(x)\|^2 + E(\|x\|^2) \end{bmatrix} \tag{140}$$

8. New Entropy Definition as Generalized Casimir Invariant Functions for Coadjoint and Adjoint Representation

In his paper written in 1974, Jean-Marie Souriau has observed that if we consider the heat expression $Q = \frac{d\Phi}{d\beta}$, that we can write $\delta\Phi - \langle Q, \delta\beta \rangle = 0$. For each $\delta\beta$ tangent to the orbit, and so generated by an element Z of the Lie algebra, if we consider the relation $\Phi\left(Ad_g(\beta)\right) = \Phi(\beta) - \left\langle \theta\left(g^{-1}\right), \beta \right\rangle$ and we differentiate it at $g = e$ using the property that $\widetilde{\Theta}(X, Y) = -\langle d\theta(X), Y \rangle$, $X, Y \in$ g, we obtain $\langle Q, [\beta, Z] \rangle + \widetilde{\Theta}(\beta, Z) = 0$. Souriau has stopped by this last equation, the characterization of Group action on $Q = \frac{\partial\Phi}{\partial\beta}$. Souriau has also observed that $S\left[Q\left(Ad_g(\beta)\right)\right] = S\left[Ad_g^*(Q) + \theta(g)\right] = S(Q)$. We propose to characterize more explicitly this invariance, by characterizing Entropy as an invariant Casimir function in coadjoint representation.

From last Souriau equation, if we use the identities $\beta = \frac{\partial S}{\partial Q}$, $ad_\beta Z = [\beta, Z]$ and $\widetilde{\Theta}(\beta, Z) = \langle \Theta(\beta), Z \rangle$, then we can deduce that $\left\langle ad^*_{\frac{\partial S}{\partial Q}} Q + \Theta\left(\frac{\partial S}{\partial Q}\right), Z \right\rangle = 0, \forall Z$. So, Entropy $S(Q)$ should verify $ad^*_{\frac{\partial S}{\partial Q}} Q + \Theta\left(\frac{\partial S}{\partial Q}\right) = 0$, characterizes an invariant Casimir function in case of non-null cohomology, that we propose to write with Poisson brackets, $\{S, H\}_{\widetilde{\Theta}}(Q) = 0$ where $\{S, H\}_{\widetilde{\Theta}}(Q) = \left\langle Q, \left[\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q}\right] \right\rangle + \widetilde{\Theta}\left(\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q}\right) = 0$, $\forall H : g^* \to R, Q \in g^*$.

In a Poisson manifold, Casimir functions $S \in C^\infty(g^*)$, in case of null cohomology, are functions whose Poisson brackets will all functions vanish, $\{S, H\}(Q) = 0$, $\forall S \in C^\infty(g^*), Q \in g^*$. In the dual of the Lie algebra of a connected Lie group G, the Casimir functions are the Ad^*-invariant functions, because if $S, H \in C^\infty(g^*)$ and $Q \in g^*$, then $\{S, H\}(Q) = \left\langle Q, \left[\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q}\right] \right\rangle = \left\langle Q, ad_{\frac{\partial S}{\partial Q}} \frac{\partial H}{\partial Q} \right\rangle = \left\langle ad^*_{\frac{\partial S}{\partial Q}} Q, \frac{\partial H}{\partial Q} \right\rangle$ vanishes for all $H \in C^\infty(g^*)$ if and only if $ad^*_{\frac{\partial S}{\partial Q}} Q = 0$. A function is S on g^* is Ad^*-invariant if $g.S = S$, $\forall g \in G$ where Lie group G acts on functions on g^* by $(g.S)(Q) = S\left(Ad_g^* Q\right), Q \in g^*, S \in C^\infty(g^*), g \in G$, and where infinitesimal characterizations of Ad^*-invariant functions on g^*, $\frac{d}{dt} S\left(Ad^*_{\exp(tx)} Q\right)\Big|_{t=0} = \left\langle ad_x^* Q, \frac{\partial S}{\partial Q} \right\rangle = -\left\langle ad^*_{\frac{\partial S}{\partial Q}} Q, x \right\rangle$. The symplectic leaves of a Poisson manifold are contained in the connected components of the level sets of the Casimir functions and Casimir function is constant on a symplectic leaf. Coadjoint orbits lie on level sets of the Casimir functions, which are conserved quantities. Casimir functions

Level sets are symplectic manifolds. Coadjoint motion of the moment map $Q(t) = Ad^*_{g(t)}Q(0)$ for a solution curve $g(t) \in C(G)$ take place on the intersections of levels sets of the Hamiltonian and the Casimir functions. Alexis Arnaudon has studied stochastic coadjoint processes whose solutions lie on coadjoint orbits.

　　We have observed that $\{S, H\}_{\widetilde{\Theta}}(Q) = \left\langle Q, \left[\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q}\right]\right\rangle + \widetilde{\Theta}\left(\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q}\right) = 0, \forall H : g^* \to R, Q \in g^*$, that shows that Souriau Entropy is a Casimir function in case with non-null cohomology when an additional cocycle should be taken into account. Indeed, infinitesimal variation is characterized by the following differentiation: $\frac{d}{dt}S\big(Q\big(Ad_{\exp(tx)}\beta\big)\big)\big|_{t=0} = \frac{d}{dt}S\big(Ad^*_{\exp(tx)}Q + \theta(\exp(tx))\big)\big|_{t=0} = -\left\langle ad^*_{\frac{\partial S}{\partial Q}}Q + \Theta\left(\frac{\partial S}{\partial Q}\right), x\right\rangle$. We recover extended Casimir equation in case of non-null cohomology verified by Entropy, $ad^*_{\frac{\partial S}{\partial Q}}Q + \Theta\left(\frac{\partial S}{\partial Q}\right) = 0$, and then the generalized Casimir condition $\{S, H\}_{\widetilde{\Theta}}(Q) = 0$. Hamiltonian motion on these affine coadjoint orbits is given by the solutions of the Lie-Poisson equations with cocycle.

　　The identification of Entropy as an Invariant Casimir Function in Coadjoint representation is also important in Information Theory, because classically Entropy is introduced axiomatically. With this new approach, we can build Entropy by constructing the Casimir Function associated to the Lie group and also in case of non-null cohomology. Igor V. Shirokov [71–75] has proposed a method for constructing invariants of the coadjoint representation of Lie groups with an arbitrary dimension and structure based on local symplectic coordinates on the coadjoint orbits. The idea of the method of constructing coadjoint invariants is to construct the canonical transition to the Darboux coordinates on the orbits of the dual Lie algebra g* of maximal dimension dual to the Lie algebra g of the Lie group G. These relations provide invariants of the coadjoint representation of the Lie group G.

　　This geometric framework unifies several earlier works on the subject, including Souriau's symplectic model of statistical mechanics, and approaches developed in Information Geometry and Quantum Information Geometry. This approach helps to identify the common geometric structures appearing in various domains from statistical mechanics to statistical learning. The emphasis is put on the role of the affine equivariance with respect to Lie group actions, as extension of the Fisher metric in presence of equivariance and the associated Lie-Poisson equations with cocycle (affine Lie-Poisson equations). The entropy of the Souriau model as a Casimir function can be used to apply a geometric model for energy preserving entropy production on Lie algebras. We can exploit the geometric framework of this new equation to build geometric numerical integrator schemes for some of the equations associated to Souriau's model and its polysymplectic extension. This new equation is important because it introduce new structure of differential equations in case of non-null cohomology and for an arbitrary Hamiltonian $H : g^* \to R$: $\frac{dQ}{dt} = ad^*_{\frac{\partial H}{\partial Q}}Q + \Theta\left(\frac{\partial H}{\partial Q}\right)$.

　　The equation $\frac{dQ}{dt} = ad^*_{\frac{\partial H}{\partial Q}}Q + \Theta\left(\frac{\partial H}{\partial Q}\right)$ is important because it allows extending stochastic perturbation of the Lie-Poisson equation with cocycle within the setting of stochastic Hamiltonian dynamics, which preserves the affine coadjoint orbits. We can extend model for stochastic geometric modeling in fluid dynamics via variational principles described in [32,76]. This extension results in the new Stratonovich differential equation for the stochastic process $dQ + \left[ad^*_{\frac{\partial H}{\partial Q}}Q + \Theta\left(\frac{\partial H}{\partial Q}\right)\right]dt + \sum_{i=1}^{N}\left[ad^*_{\frac{\partial H_i}{\partial Q}}Q + \Theta\left(\frac{\partial H_i}{\partial Q}\right)\right] \circ dW_i(t) = 0$.

　　This new equation is also very usefull for geometric symplectic Lie group integrator for Lie-Poisson equations with cocycle that preserves the affine coadjoint orbits for general Hamiltonian. This equation is also very relevant in the framework of dynamics with Casimir dissipation/production, to formulate a dynamical geometric model for dissipation/production of this Casimir. This allows to extend the general Lie algebraic approach developed in [77,78] for Casimir dissipation, to take into account of a

cocycle, and to a wider class of dissipation. Paper [17] will exploit this new Casimir equation in case of non-null cohomology.

This equation $\frac{dQ}{dt} = ad^*_{\frac{\partial H}{\partial Q}} Q + \Theta\left(\frac{\partial H}{\partial Q}\right)$ could be used also to make the link with 2nd principle of Thermodynamique, that will be deduced from positivity of Souriau-Fisher metric:

$$S(Q) = \langle Q, \beta \rangle - \Phi(\beta) \text{ with } \frac{dQ}{dt} = ad^*_{\frac{\partial H}{\partial Q}} Q + \Theta\left(\frac{\partial H}{\partial Q}\right)$$

$$\frac{dS}{dt} = \left\langle Q, \frac{d\beta}{dt} \right\rangle + \left\langle ad^*_{\frac{\partial H}{\partial Q}} Q + \Theta\left(\frac{\partial H}{\partial Q}\right), \beta \right\rangle - \frac{d\Phi}{dt} = \left\langle Q, \frac{d\beta}{dt} \right\rangle + \left\langle ad^*_{\frac{\partial H}{\partial Q}} Q, \beta \right\rangle + + \left\langle \Theta\left(\frac{\partial H}{\partial Q}\right), \beta \right\rangle - \frac{d\Phi}{dt}$$

$$\frac{dS}{dt} = \left\langle Q, \frac{d\beta}{dt} \right\rangle + \left\langle Q, \left[\frac{\partial H}{\partial Q}, \beta\right] \right\rangle + \widetilde{\Theta}\left(\frac{\partial H}{\partial Q}, \beta\right) - \frac{d\Phi}{dt} = \left\langle Q, \frac{d\beta}{dt} \right\rangle + \widetilde{\Theta}_\beta\left(\frac{\partial H}{\partial Q}, \beta\right) - \left\langle \frac{\partial \Phi}{\partial \beta}, \frac{d\beta}{dt} \right\rangle \qquad (141)$$

$$\frac{dS}{dt} = \left\langle Q, \frac{d\beta}{dt} \right\rangle + \widetilde{\Theta}_\beta\left(\frac{\partial H}{\partial Q}, \beta\right) - \left\langle \frac{\partial \Phi}{\partial \beta}, \frac{d\beta}{dt} \right\rangle \text{ with } \frac{\partial \Phi}{\partial \beta} = Q$$

$$\frac{dS}{dt} = \widetilde{\Theta}_\beta\left(\frac{\partial H}{\partial Q}, \beta\right) \geq 0, \forall H \text{ (link to positivity of Fisher metric)}$$

$$\text{if } H = S \underset{\frac{\partial S}{\partial Q} = Q}{\Rightarrow} \frac{dS}{dt} = \widetilde{\Theta}_\beta(\beta, \beta) = 0 \text{ because } \beta \in Ker\widetilde{\Theta}_\beta$$

Entropy production is then linked with Souriau-Fisher structure, $dS = \widetilde{\Theta}_\beta\left(\frac{\partial H}{\partial Q}, \beta\right)dt$ with $\widetilde{\Theta}_\beta\left(\frac{\partial H}{\partial Q}, \beta\right) = \Theta\left(\frac{\partial H}{\partial Q}, \beta\right) + \left\langle Q, \left[\frac{\partial H}{\partial Q}, \beta\right] \right\rangle$ Souriau tensor related to Fisher metric.

8.1. Casimir Invariant and Generalized Casimir Invariant

Hendrik Brugt Gerhard Casimir, a Dutch physicist, studied what is called Casimir operators and Casimir invariants (H. Casimir and Van der Waerden studied the SU(2) group, the group of isospin/angular momentum, as the model of the algebraic approach to the study of the unitary representations of semi-simple compact Lie groups). Kirillov has explained that Casimir operators are in one-to-one correspondence with polynomial invariants characterizing orbits of the coadjoint representation. Solutions are not necessarily polynomials and the nonpolynomial solutions are called ***generalized Casimir invariants***. For certain classes of Lie algebras, all invariants of the coadjoint representation are functions of polynomial ones. In physics, Hamiltonians and integrals of motion of classical integrable Hamiltonian systems are not polynomials in the momenta [71–75,79–92].

8.2. Souriau Entropy as Generalized Casimir Invariant in Coadjoint Representation

In Souriau Lie groups Thermodynamics, we will see that coadjoint orbits lie on level sets of the Entropy that could be considered as a Casimir invariant function:

$$\begin{aligned} S : g^* &\to R \\ Q &\mapsto S(Q) \end{aligned} \qquad (142)$$

We will consider first the case of null-cohomology, Entropy as Casimir invariant function is a conserved quantity, because Casimir function has null Lie Poisson brackets functions [93,94]:

$$\{S, H\}(Q) = \left\langle Q, \left[\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q}\right] \right\rangle = 0, \forall H : g^* \to R, Q \in g^*, \langle A, B \rangle = B(A, B) \text{ Cartan-Killing form}$$

$$\text{with } \partial S(Q) = \frac{d}{d\varepsilon} S(Q + \delta Q)\big|_{\varepsilon=0} = \left\langle \delta Q, \frac{\partial S}{\partial Q} \right\rangle \qquad (143)$$

We can observe that $\beta = \frac{\partial S}{\partial Q}$, then:

$$\left\langle Q, \left[\beta, \frac{\partial H}{\partial Q}\right] \right\rangle = \left\langle Q, ad_\beta \frac{\partial H}{\partial Q} \right\rangle = 0, \forall H : g^* \to R, Q \in g^*, ad_a b = [a, b] \qquad (144)$$

We can also write:

$$\left\langle Q, \left[\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q}\right]\right\rangle = \left\langle Q, ad_{\frac{\partial S}{\partial Q}} \frac{\partial H}{\partial Q}\right\rangle = \left\langle ad^*_{\frac{\partial S}{\partial Q}} Q, \frac{\partial H}{\partial Q}\right\rangle = 0 , \forall H : g^* \to R \tag{145}$$

It means that $ad^*_{\frac{\partial S}{\partial Q}} Q = ad^*_{\beta} Q = 0$, $\beta = \frac{\partial S}{\partial Q}$. We can remark that if we note $\left(ad^*_{\frac{\partial S}{\partial Q}} Q\right)_j = C^k_{ij} ad^*_{(\frac{\partial S}{\partial Q})^i} Q_k = 0$ with C^k_{ij} the structure tensor, we observe that this equation is in fact the Casimir condition for invariant function in coadjoint representation as we will see hereafter. The restriction of the Lie-Poisson bracket to an orbit generates a symplectic structure on the orbit, called the KKS (Kirillov-Kostant-Souriau) structure, or the canonical symplectic structure. Casimir function is characterized as a quantity which commutes with each linear functional on the Poisson manifold, and then it is conserved by dynamics of any Hamiltonian.

Given a Hamiltonian $H : g^* \to R$, the equation of motion for $Q \in g^*$ is:

$$\frac{dQ}{dt} = \{Q, H\} = ad^*_{\frac{\partial H}{\partial Q}} Q \text{ with } H = S \Rightarrow \frac{dQ}{dt} = \{Q, S\} = ad^*_{\frac{\partial S}{\partial Q}} Q = 0 \tag{146}$$

In case of non-null cohomology, the Lie Poisson brackets functions are given by:

$$\begin{aligned}
&\{S, H\}_{\widetilde{\Theta}}(Q) = \left\langle Q, \left[\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q}\right]\right\rangle + \widetilde{\Theta}\left(\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q}\right) = 0, \forall H : g^* \to R, Q \in g^* \\
&\text{with } \widetilde{\Theta}(X, Y) = J_{[X,Y]} - \{J_X, J_Y\} \text{ where } J_X(x) = \langle J(x), X\rangle \\
&\widetilde{\Theta}(X, Y) : g \times g \to \mathfrak{R} \qquad \text{with } \Theta(X) = T_e \theta(X(e)) \\
&\qquad\qquad X, Y \mapsto \langle \Theta(X), Y\rangle
\end{aligned} \tag{147}$$

That we can develop in the following:

$$\begin{aligned}
&\{S, H\}_{\widetilde{\Theta}}(Q) = \left\langle Q, \left[\frac{\partial S}{\partial Q}, \frac{\partial H}{\partial Q}\right]\right\rangle + \left\langle \Theta\left(\frac{\partial S}{\partial Q}\right), \frac{\partial H}{\partial Q}\right\rangle = 0 \\
&\{S, H\}_{\widetilde{\Theta}}(Q) = \left\langle Q, ad_{\frac{\partial S}{\partial Q}} \frac{\partial H}{\partial Q}\right\rangle + \left\langle \Theta\left(\frac{\partial S}{\partial Q}\right), \frac{\partial H}{\partial Q}\right\rangle = 0 \\
&\{S, H\}_{\widetilde{\Theta}}(Q) = \left\langle ad^*_{\frac{\partial S}{\partial Q}} Q, \frac{\partial H}{\partial Q}\right\rangle + \left\langle \Theta\left(\frac{\partial S}{\partial Q}\right), \frac{\partial H}{\partial Q}\right\rangle = 0 \\
&\forall H, \{S, H\}_{\widetilde{\Theta}}(Q) = \left\langle ad^*_{\frac{\partial S}{\partial Q}} Q + \Theta\left(\frac{\partial S}{\partial Q}\right) +, \frac{\partial H}{\partial Q}\right\rangle = 0 \Rightarrow ad^*_{\frac{\partial S}{\partial Q}} Q + \Theta\left(\frac{\partial S}{\partial Q}\right) = 0
\end{aligned} \tag{148}$$

We have found the generalized Casimir equation for Entropy in the non-null cohomology case:

$$\{S, H\}_{\widetilde{\Theta}}(Q) = 0 \tag{149}$$

That could be also written:

$$ad^*_{\frac{\partial S}{\partial Q}} Q + \Theta\left(\frac{\partial S}{\partial Q}\right) = 0 \tag{150}$$

This equation was observed by Souriau in his paper of 1974, where he has written that geometric temperature β is a kernel of $\widetilde{\Theta}_\beta$, that is written:

$$\beta \in Ker\widetilde{\Theta}_\beta \Rightarrow \langle Q, [\beta, Z]\rangle + \widetilde{\Theta}(\beta, Z) = 0 \tag{151}$$

That we can develop to recover the Casimir equation:

$$\Rightarrow \left\langle Q, ad_\beta Z \right\rangle + \widetilde{\Theta}(\beta, Z) = 0 \Rightarrow \left\langle ad_\beta^* Q, Z \right\rangle + \widetilde{\Theta}(\beta, Z) = 0$$

$$\beta = \frac{\partial S}{\partial Q} \Rightarrow \left\langle ad_{\frac{\partial S}{\partial Q}}^* Q, Z \right\rangle + \widetilde{\Theta}\left(\frac{\partial S}{\partial Q}, Z\right) = \left\langle ad_{\frac{\partial S}{\partial Q}}^* Q + \Theta\left(\frac{\partial S}{\partial Q}\right), Z \right\rangle = 0, \forall Z \tag{152}$$

$$\Rightarrow ad_{\frac{\partial S}{\partial Q}}^* Q + \Theta\left(\frac{\partial S}{\partial Q}\right) = 0$$

Then the generalized Casimir Equation in non-null cohomogy is given by:

$$\left(ad_{\frac{\partial S}{\partial Q}}^* Q\right)_j + \Theta\left(\frac{\partial S}{\partial Q}\right)_j = C_{ij}^k ad_{\left(\frac{\partial S}{\partial Q}\right)^i}^* Q_k + \Theta_j = 0 \tag{153}$$

Given a Hamiltonian $H : g^* \to R$, the equation of motion for $Q \in g^*$ is:

$$\frac{dQ}{dt} = ad_{\frac{\partial H}{\partial Q}}^* Q + \Theta\left(\frac{\partial H}{\partial Q}\right) \text{ with } H = S \Rightarrow \frac{dQ}{dt} = ad_{\frac{\partial S}{\partial Q}}^* Q + \Theta\left(\frac{\partial S}{\partial Q}\right) = 0 \tag{154}$$

Level sets of the Casimir Entropy function, on which the coadjoint orbits lie, are symplectic manifolds.

8.3. Souriau Entropy Invariance in Coadjoint Representation

If we note $An(g^*)$ the space of analytic function on the dual space of the Lie agebra g^*, a function $F^* \in An(g^*)$ is a Casimir invariant if for any $g \in G, X \in g^*$, we have $F^*\left(Ad_g^* X\right) = F^*(X)$. We have observed previously that Souriau's Entropy analytic function $S(Q)$ defined on dual space of the Lie algebra g^* by Legendre transform of Massieu Characteric analytic function $\Phi(\beta)$ (minus logarithm of Laplace transform) defined on Lie algebra g was an invariant function under the affine coadjoint action $S\left[Q\left(Ad_g(\beta)\right)\right] = S\left[Ad_g^*(Q) + \theta(g)\right] = S(Q)$. In case of null-cohomology, Souriau cocycle cancels $\theta(g) = 0$, and we recover Casimir invariant function in coadjoint representation $S\left[Ad_g^*(Q)\right] = S(Q)$.

We can then observe that Souriau Entropy is an extended Casimir invariant function in case of non-null cohomogy. This characteristic of Souriau Entropy could be a new characterization of Entropy. In Souriau Lie groups Thermodynamics, Entropy $S(Q)$ is a generalized Casimir invariant function for coadjoint representation in case of non-null cohomology, and Massieu Characteristic function by Legendre duality is a generalized Casimir function for adjoint representation.

We will explain how to prove that Souriau Entropy is invariant under the action of the group, starting from its definition:

$$S(Q) = \langle Q, \beta \rangle - \Phi(\beta) \text{ with } Q = \frac{\partial \Phi(\beta)}{\partial \beta} \in g^* \text{ and } \beta = \frac{\partial S(Q)}{\partial Q} \in g \tag{155}$$

with

$$\Phi(\beta) = -\log \int_M e^{-\langle U(\xi), \beta \rangle} d\lambda_\omega \text{ and } U : M \to g^* \tag{156}$$

Considering Souriau Entropy $S(Q)$ where the heat $Q = \frac{\partial \Phi(\beta)}{\partial \beta} \in g^*$ an element of the dual space of the Lie algebra is parameterized by $\beta \in g$ an element of the Lie algebra, the Lie group G acts through $g \in G$ by adjoint operator Ad_g, the entropy is given by $S\left[Q\left(Ad_g(\beta)\right)\right]$ with $Q\left(Ad_g(\beta)\right)$ given by fundamental Souriau equation:

$$Q\left(Ad_g(\beta)\right) = Ad_g^*(Q) + \theta(g) \tag{157}$$

The invariance of Souriau Entropy is deduced from the following developments:

$$\beta \in \mathfrak{g} \to Ad_g(\beta) \Rightarrow \Psi(Ad_g(\beta)) = \int_M e^{-\langle U, Ad_g(\beta) \rangle} d\lambda_\omega$$

$$\Psi(Ad_g(\beta)) = \int_M e^{-\langle Ad^*_{g^{-1}} U, \beta \rangle} d\lambda_\omega = \int_M e^{-\langle U(Ad_{g^{-1}}\beta) - \theta(g^{-1}), \beta \rangle} d\lambda_\omega$$

$$\Psi(Ad_g(\beta)) = e^{\langle \theta(g^{-1}), \beta \rangle} \Psi(\beta) \tag{158}$$

$$\theta(g^{-1}) = -Ad^*_{g^{-1}} \theta(g) \Rightarrow \Psi(Ad_g(\beta)) = e^{-\langle Ad^*_{g^{-1}} \theta(g), \beta \rangle} \Psi(\beta)$$

$$\Phi(\beta) = -\log \Psi(\beta) \Rightarrow \Phi(Ad_g(\beta)) = \Phi(\beta) - \langle \theta(g^{-1}), \beta \rangle = \Phi(\beta) + \langle Ad^*_{g^{-1}} \theta(g), \beta \rangle$$

Based on this expression of Massieu Characteristic function transform by action of the group, we can use Legendre transform to study how Souriau Entropy is changed:

$$S(Q) = \langle Q, \beta \rangle - \Phi(\beta) \Rightarrow S(Q(Ad_g\beta)) = \langle Q(Ad_g\beta), Ad_g\beta \rangle - \Phi(Ad_g\beta)$$

$$\begin{cases} Q(Ad_g(\beta)) = Ad^*_g(Q) + \theta(g) \\ \Phi(Ad_g(\beta)) = -\log \Psi(Ad_g(\beta)) = -\langle \theta(g^{-1}), \beta \rangle + \Phi(\beta) \end{cases}$$

$$\Rightarrow S(Q(Ad_g\beta)) = \langle Ad^*_g(Q) + \theta(g), Ad_g\beta \rangle + \langle \theta(g^{-1}), \beta \rangle - \Phi(\beta) \tag{159}$$

$$\Rightarrow S(Q(Ad_g\beta)) = \langle Ad^*_g(Q) + \theta(g), Ad_g\beta \rangle - \langle Ad^*_{g^{-1}} \theta(g), \beta \rangle - \Phi(\beta)$$

$$\Rightarrow S(Q(Ad_g\beta)) = \langle Ad^*_{g^{-1}} Ad^*_g(Q) + Ad^*_{g^{-1}} \theta(g), \beta \rangle - \langle Ad^*_{g^{-1}} \theta(g), \beta \rangle - \Phi(\beta)$$

$$Ad^*_{g^{-1}} Ad^*_g(Q) = Q \Rightarrow S(Q(Ad_g\beta)) = \langle Q, \beta \rangle - \Phi(\beta) = S(\beta)$$

We finally prove that Souriau Entropy is invariant in coadjoint representation $S(Ad^*_g(Q) + \theta(g)) = S(\beta)$ in general case of non-null cohomology, that we could write $S(Ad^\#_g(Q)) = S(\beta)$, if we note affine coadjoint action $Ad^\#_g(Q) = Ad^*_g(Q) + \theta(g)$. This is also true in case of null-cohomology when the Souriau cocycle cancels $\theta(g) = 0$, and we recover classical generalized Casimir invariant function definition on coadjoint representation for Entropy $S(Ad^*_g(Q)) = S(\beta)$ generalized Casimir invariant function definition on adjoint representation for Massieu Characteristic function $\Phi(Ad_g(\beta)) = \Phi(\beta)$.

8.4. Souriau Entropy Given by Casimir Invariant Functions Equations

Based on development given in the following we can state that:

As the Entropy S is a generalized Casimir invariant function in the coadjoint representation, $S(Ad^*_{e^{t\xi}} h) = S(h)$, then S should be solution of the following differential equation:

$$C^k_{ij} Q_k \frac{\partial S(Q)}{\partial Q_j} = 0 \, , \, i, j, k = \dim \mathfrak{g}, \text{ with } \begin{cases} C^k_{ij} Q_k = C_{ij}(Q) = B_{ij} \\ B_Q(x, y) = B_{ij} x^i y^j = \langle Q, [x, y] \rangle \end{cases} \tag{160}$$

where C^k_{ij} is the structure tensor of the Lie algebra $\mathfrak{g}$ in the basis $(e_1, e_2, \ldots, e_n)$, while X_k are the coordinates in $\mathfrak{g}^*$ in the basis $(e^1, e^2, \ldots, e^n)$ defined by $\langle e^j, e_i \rangle = \delta_{ij}$. The structure tensor s given by $[\phi(e_i), \phi(e_i)] = C^k_{ij} \phi(e_k)$ with $\phi(e_i) = C^k_{ij} X_k \frac{\partial}{\partial X_j}$, $i = 1, \ldots, n$.

8.5. Characterization of Generalized Casimir Invariant Functions in Coadjoint Representation

We will describe recent characterization of generalized Casimir invariant functions by Oleg L. Kurnyavko and Igor V. Shirokov [72,73,75] who have proposed Algebraic method for construction of Casimir invariants of Lie groups coadjoint representations (see Appendix C). Modern invariant theory based on geometric methods, which was credited classically as non-constructive, has some exception admitting a constructive solution related to the constructing invariants of Lie groups representations.

Let T be a connected Lie group, $T(G)$ a representation of the group G in the linear space V, T_g the operators associated to the representation of the group G on the linear space V, then the invariants are given by the following equation:

$$F\left(T_g x\right) = F(x) \, , x \in V, g \in G, T_g \in T(G), F(x) \in C^\infty(V) \tag{161}$$

With the properties that:

$$T_e = I \, , \, T_{g_a g_b} = T_{g_a} T_{g_b} \, , \, T_{g^{-1}} = \left(T_g\right)^{-1} \tag{162}$$

Solution is given by the following differential equation:

$$-\sum_{i,j}^{\dim V} t_{kj}^i x^j \frac{\partial F(x)}{\partial x^i} = 0 \text{ with } t_{kj}^i = \left.\frac{\partial \left(T_g\right)_j^i}{\partial g^k}\right|_{g=e} \text{ and } k = 1, \ldots, \dim G \tag{163}$$

t_{kj}^i are elements of the matrices of the Lie algebra representation basis of G.

That we can write $t_k = -t_{kj}^i x^j \frac{\partial}{\partial x^i}$ and $t_k F(x) = 0$.

If we consider the dual space V^*, the co-tangent representation is given by:

$$\langle T^*(g)X, T(g)x \rangle = \langle X, x \rangle \tag{164}$$

And co-represnetation invariants are given by:

$$t_k^* F^*(X) = 0 \text{ with } t_k^* = t_{kj}^i X_i \frac{\partial}{\partial X_j} \tag{165}$$

They have underlined the relationship between invariants of representations and conjugate representations, where the algebraic construction of Lie groups representations invariants are given by invariants of the conjugate representation with respect to the invariants of the original representation.

Shirokov Theorem 1. *Let $F(x)$ be a non-degenerate invariant of the representation $T(G)$, then conjugate representation invariant can be found by Legrendre tranform:*

$$F^*(X) = x^i X_i - F(x) = \langle x, X \rangle - F(x) \text{ with } X = \frac{\partial F(x)}{\partial x} \text{ such that } X_i = \frac{\partial F(x)}{\partial x^i} \tag{166}$$

and also the converse problem:

$$F(x) = x^i X_i - F^*(X) = \langle x, X \rangle - F^*(X) \text{ with } x = \frac{\partial F^*(X)}{\partial X} \text{ such that } x^i = \frac{\partial F^*(X)}{\partial X_i} \tag{167}$$

Shirokov has considered $F(x)$ the representation invariant $T(G)$, and $F^*(X)$ the representation invariant $T^*(G)$ conjugate to $T(G)$, with the conditions:

$$-t_{kj}^i x^j \frac{\partial F(x)}{\partial x^i} = 0 \text{ and } t_{lj}^i X_i \frac{\partial F^*(X)}{\partial X_j} = 0 \tag{168}$$

$$
\begin{aligned}
t_{lj}^i X_i \frac{\partial F^*(X)}{\partial X_j} &= t_{lj}^i X_i \frac{\partial}{\partial X_j}\left[x^k(X)X_k - F(x(X))\right] = t_{lj}^i X_i \frac{\partial x^k}{\partial X_j} X_k + t_{lj}^i X_i x^k \frac{\partial X_k}{\partial X_j} - t_{lj}^i X_i \frac{\partial F(x)}{\partial x^k} \frac{\partial x^k}{\partial X_j} \\
t_{lj}^i X_i \frac{\partial F^*(X)}{\partial X_j} &= t_{lj}^i X_i \frac{\partial x^k}{\partial X_j} \frac{\partial F(x)}{\partial x^k} + t_{lj}^i \frac{\partial F(x)}{\partial x^i} x^k \delta_k^j - t_{lj}^i \frac{\partial F(x)}{\partial x^i} \frac{\partial F(x)}{\partial x^k} \frac{\partial x^k}{\partial X_j} \\
t_{lj}^i X_i \frac{\partial F^*(X)}{\partial X_j} &= t_{lj}^i x_j \frac{\partial F(x)}{\partial x^i} = 0
\end{aligned}
\tag{169}
$$

Invariant Casimir Functions of the coadjoint representation has been studied for completely integrable Hamiltonian systems, as classical systems on the orbits of the coadjoint representation. Oleg L. Kurnyavko and Igor V. Shirokov have considered the relationship between invariants of representations of Lie groups and their conjugate dual representations.

Considering the coadjoint action given by:

$$\left\langle Ad_g^* X, x \right\rangle = \left\langle X, Ad_{g^{-1}} x \right\rangle, \, g \in G, X \in \mathfrak{g}^*, x \in \mathfrak{g} \tag{170}$$

Invariants of a coadjoint representation are called Casimir functions, with the property:

$$F^*\left(Ad_g^* X \right) = F^*(X) \tag{171}$$

the infinitesimal invariance is given by the equations:

$$C_{ij}(X) \frac{\partial F^*(X)}{\partial X_j} = 0 \text{ with } C_{ij}(X) = C_{ij}^k X_k, \, i, j, k = \dim\mathfrak{g} \tag{172}$$

The number of functionally independent invariants is given by the rank of the matrix $C_{ij}(X)$, called the index of the Lie algebra $\mathfrak{g}$: $\mathrm{ind}\mathfrak{g} = \dim\mathfrak{g}^* - \sup_{X \in \mathfrak{g}^*} rank C_{ij}(X)$.

From these adjoint and coadjoint representation, Shirokov has introduced the following theorem:

Shirokov Theorem 2. *Let* $F\left(Ad_g x \right) = F(x)$ *be a non-degenerate invariant of the adjoint representation* Ad_G, *then conjugate representation invariant, invariant of coadjoint representation* Ad_G^* *can be found by formula:*

$$F^*(X) = x^i X_i - F(x) = \langle x, X \rangle - F(x) \text{ with } X = \frac{\partial F(x)}{\partial x} \text{ such that } X_i = \frac{\partial F(x)}{\partial x^i} \tag{173}$$

and also the converse problem, let $F^*\left(Ad_g^* X \right) = F^*(X)$, *invariant of coadjoint representation* Ad_G *is given by:*

$$F(x) = x^i X_i - F^*(X) = \langle x, X \rangle - F^*(X) \text{ with } x = \frac{\partial F^*(X)}{\partial X} \text{ such that } x^i = \frac{\partial F^*(X)}{\partial X_i} \tag{174}$$

Nota:

$$C_{ij}^k X_k \frac{\partial F^*(X)}{\partial X_j} = 0 \, , \, i, j, k = \dim\mathfrak{g}, \text{ with } \begin{cases} C_{ij}^k X_k = C_{ij}(X) = B_{ij} \\ B_X(x, y) = B_{ij} x^i y^j = \langle X, [x, y] \rangle \end{cases} \tag{175}$$

8.6. Constructing Generalized Casimir Invariant Functions in Coadjoint Representation

I. V. Shirokov has proposed a method for constructing invariants of the coadjoint representation of Lie groups with an arbitrary dimension and structure based on local symplectic coordinates on the coadjoint orbits. Oleg L. Kurnyavko and Igor V. Shirokov have also proposed a general method for constructing Casimir invariants.

We will give some other developments of Casimir Invariant Functions by A.T. Fomenko and V.V. Trofimov, related to Orbits of the coadjoint representation and the associated canonical symplectic structure.

The coadjoint orbit O_h passing through the point $h \in \mathfrak{g}^*$ is given by

$$O_h = \left\{ Ad_g^* h \, / \, g \in G \right\} \text{ where } h \in \mathfrak{g}^* \tag{176}$$

$$T_h O_h = \left\{ ad_\rho^* h / \rho \in \mathfrak{g} \right\} \subset \mathfrak{g}^*, h \in \mathfrak{g}^*$$

$$v = \left. \frac{d\left[Ad_{\exp^{t\rho}}^* h \right]}{dt} \right|_{t=0} \in T_h O_h, \rho \in \mathfrak{g} \tag{177}$$

Let $(e_1, \ldots, e_n)$ basis of $\mathfrak{g}$, $\left(e^1, \ldots, e^n \right)$ basis of $\mathfrak{g}^*$, with $\left\langle e^i, e_j \right\rangle = \delta_{i,j}$

$$h = h_i e^i \Rightarrow v^i = \left. \frac{d\left[\left\langle h_i, Ad_{\exp^{t\rho}}^* h \right\rangle \right]}{dt} \right|_{t=0} = \left. \frac{d\left[\left\langle Ad_{\exp^{t\rho}}^* h, e_i \right\rangle \right]}{dt} \right|_{t=0}$$

$$v^i = \left. \frac{d\left[\left\langle h, Ad_{\exp^{-t\rho}} e_i \right\rangle \right]}{dt} \right|_{t=0} = \left\langle h, \left. \frac{d\left[Ad_{\exp^{-t\rho}} e_i \right]}{dt} \right|_{t=0} \right\rangle \tag{178}$$

$$v^i = \left\langle h, -[\rho, e_i] \right\rangle = -\left\langle ad_\rho^* h, e_i \right\rangle = \left\langle v, e_i \right\rangle \Rightarrow v = -ad_\rho^* h$$

Kirillov, Kostant and Souriau have introduced a KKS 2-form on co-adjoint co-orbits that then inherit a structure of homogeneous symplectic manifold:

$$\xi, \eta \in T_h O_h = \left\{ ad_\chi^* h / \rho \in \mathfrak{g} \right\} \subset \mathfrak{g}^*, h \in \mathfrak{g}^*$$

$$\omega_h(\xi, \eta) = \omega\left(ad_{\xi_1}^* h, ad_{\eta_1}^* h \right) = \left\langle h, [\xi_1, \eta_1] \right\rangle \text{ with } \xi = ad_{\xi_1}^* h \text{ and } \eta = ad_{\eta_1}^* h \tag{179}$$

This KKS 2-form ω is invariant with respect to the coadjoint action $\omega_g\left(Ad_f^* \xi, Ad_f^* \eta \right) = \omega_h(\xi, \eta)$:

$$\omega_g\left(Ad_f^* \xi, Ad_f^* \eta \right) = \omega_g\left(Ad_f^* ad_{\xi_1}^* h, Ad_f^* ad_{\eta_1}^* h \right)$$

$$\text{with } g = Ad_f^* h, \xi = ad_{\xi_1}^* h, \eta = ad_{\eta_1}^* h \text{ and } f \in G, g, h \in \mathfrak{g}^*$$

$$Ad_f^* ad_{\xi_1}^* h = ad_{Ad_f \xi_1}^* \left(Ad_f^* h \right) \text{ and } Ad_f^* ad_{\eta_1}^* h = ad_{Ad_f \eta_1}^* \left(Ad_f^* h \right)$$

$$\omega_g\left(Ad_f^* \xi, Ad_f^* \eta \right) = \omega_g\left(ad_{Ad_f \xi_1}^* \left(Ad_f^* h \right), ad_{Ad_f \eta_1}^* \left(Ad_f^* h \right) \right)$$

$$\omega_g\left(Ad_f^* \xi, Ad_f^* \eta \right) = \omega_g\left(ad_{Ad_f \xi_1}^* g, ad_{Ad_f \eta_1}^* g \right) = \left\langle g, \left[Ad_f \xi_1, Ad_f \eta_1 \right] \right\rangle \tag{180}$$

$$\omega_g\left(Ad_f^* \xi, Ad_f^* \eta \right) = \left\langle g, Ad_f [\xi_1, \eta_1] \right\rangle = \left\langle Ad_{f^{-1}}^* g, [\xi_1, \eta_1] \right\rangle \text{ with } h = Ad_{f^{-1}}^* g$$

$$\omega_g\left(Ad_f^* \xi, Ad_f^* \eta \right) = \left\langle h, [\xi_1, \eta_1] \right\rangle$$

$$\omega_g\left(Ad_f^* \xi, Ad_f^* \eta \right) = \omega_h(\xi, \eta)$$

The symplectic structure is given due to the property that $d\omega = 0$, that could be proved making link with Jacobi identity.

Let $grad_{skew} m$ such that $\omega(v, grad_{skew} m) = v(m) = \sum_i v^i \frac{\partial m}{\partial x^i}$ smooth vector field on M

$$\{m, n\} = \omega(grad_{skew} m, grad_{skew} n) = \sum_{i<j} \omega_{ij} (grad_{skew} m)^i (grad_{skew} n)^j \tag{181}$$

with $(grad_{skew} m)^i = \sum_j \omega^{ij} \frac{\partial m}{\partial x^j} \Rightarrow \{m, n\} = \sum_{i<j} \omega^{ij} \frac{\partial m}{\partial x^i} \frac{\partial n}{\partial x^j}$

Jacobi identity can be computed:

$$
\begin{aligned}
&\{m, \{n, p\}\} = -(grad_{Skew}m)\{n, p\} = -L_{grad_{Skew}m}\{n, p\} \text{ with } L_\zeta : \text{ Lie derivative} \\
&\text{If } \xi = grad_{Skew}m \\
&L_\xi\{n, p\} = L_\xi\left(\omega^{ij}\frac{\partial n}{\partial x^i}\frac{\partial p}{\partial x^j}\right) = L_\xi(\omega)^{ij}\frac{\partial n}{\partial x^i}\frac{\partial p}{\partial x^j} + \omega^{ij}\frac{\partial(\xi n)}{\partial x^i}\frac{\partial p}{\partial x^j} + \omega^{ij}\frac{\partial n}{\partial x^i}\frac{\partial(\xi p)}{\partial x^j} \\
&L_\xi\{n, p\} = L_\xi(\omega)^{ij}\frac{\partial n}{\partial x^i}\frac{\partial p}{\partial x^j} + \{\xi n, p\} + \{n, \xi p\} = L_\xi(\omega)^{ij}\frac{\partial n}{\partial x^i}\frac{\partial p}{\partial x^j} - \{\{m, n\}, p\} - \{n, \{m, p\}\} \\
&\Rightarrow \{m, \{n, p\}\} + \{\{m, n\}, p\} + \{n, \{m, p\}\} = L_\xi(\omega)^{ij}\frac{\partial n}{\partial x^i}\frac{\partial p}{\partial x^j}
\end{aligned}
\tag{182}
$$

Using Elie Cartan formula $L_\xi\omega = i(\xi)d\omega + di(\xi)\omega$. If ξ is a Hamiltonian vector field, $di(\xi)\omega = 0$ and then $L_\xi\omega = i(\xi)d\omega$. If $d\omega = 0$, then the Jacobi identity is satisfied $\{m, \{n, p\}\} + \{\{m, n\}, p\} + \{n, \{m, p\}\} = 0$ and conversely.

Let consider the Berezin Bracket:

$$
\begin{aligned}
&\{m, n\} = -C_{ij}^k x_k \frac{\partial m}{\partial x^i}\frac{\partial n}{\partial x^j} \text{ with } [e_i, e_j] = C_{ij}^k e_k \\
&\text{where } (e_1, e_2, \ldots, e_n) \text{ basis of Lie algebra g}, (e^1, e^2, \ldots, e^n) \text{basis of dual Lie algebra g}^* \\
&\text{of corresponding coordinates } x^1, \ldots, x^n \text{ for g}, x_1, \ldots, x_n \text{ for g}^*
\end{aligned}
\tag{183}
$$

This Berezin Bracket is given by:

$$
\begin{aligned}
&\{m, n\}_x = dm_x\left(ad^*_{dn(x)}(x)\right) = \left(ad^*_{dn(x)}(x)\right)(dm_x) \\
&\{m, n\}_x = \langle x, [dn_x, dm_x]\rangle = C_{ij}^k x_k \frac{\partial n}{\partial x^i}\frac{\partial m}{\partial x^j} \text{ with } dn_x = \frac{\partial n}{\partial x_i}e_i, dm_x = \frac{\partial m}{\partial x_j}e_j
\end{aligned}
\tag{184}
$$

By developping Berezin Bracket $\{m, n\} = -C_{ij}^k x_k \frac{\partial m}{\partial x^i}\frac{\partial n}{\partial x^j}$ with $[e_i, e_j] = C_{ij}^k e_k$, we can prove that the bracket verify jacoby identy $\{m, \{n, p\}\} + \{\{m, n\}, p\} + \{n, \{m, p\}\} = 0$ and then $d\omega = 0$.

We will see that differential equation for (semi-)invariants of the coadjoint representations could be established. We will note $An(g^*)$ the space of analytic function on the dual space of the Lie agebra g^*. A function $F^* \in An(g^*)$ is an invariant if for any $g \in G, X \in g^*$, we have $F^*(Ad_g^* X) = F^*(X)$, and is semi-invariant if $F^*(Ad_g^* X) = \chi(g)F^*(X)$ where $\chi(g)$ is a character of the Lie group G.

We have a representation of Lie algebras $\phi : g \to Vec(\Gamma)$ defined on basis $(e_1, e_2, \ldots, e_n)$ in g where $Vec(\Gamma)$ is the space of vector fields on Γ an open subset in g^*, given by:

$$
\phi(e_i) = C_{ij}^k X_k \frac{\partial}{\partial X_j}, i = 1, \ldots, n
\tag{185}
$$

where C_{ij}^k is the structure tensor of the Lie algebra g in the basis $(e_1, e_2, \ldots, e_n)$, while X_k are the coordinates in g^* in the basis $(e^1, e^2, \ldots, e^n)$ defined by $\langle e^j, e_i\rangle = \delta_{ij}$. The representation is not dependent of the choice of the basis, with the property: $[\phi(e_i), \phi(e_i)] = C_{ij}^k \phi(e_k)$.

We have the property, that:

$$
\left.\frac{d^n F^*\left(Ad_{e^{t\xi}}^* h\right)}{dt^n}\right|_{t=0} = \left[(-\phi(\xi))^n F^*\right](h)
\tag{186}
$$

This result is obtained by the following development:

$$
\begin{aligned}
\left.\frac{dF^*\left(Ad^*_{e^{t\xi}}h\right)}{dt}\right|_{t=0} &= \frac{\partial F^*}{\partial X_i}(h).\left.\frac{d\left\langle Ad^*_{e^{t\xi}}h,X_i\right\rangle}{dt}\right|_{t=0} = \frac{\partial F^*}{\partial X_i}(h).\left.\frac{d\left\langle h,Ad_{e^{-t\xi}}e_i\right\rangle}{dt}\right|_{t=0} \\
\left.\frac{d\left\langle h,Ad_{e^{-t\xi}}e_i\right\rangle}{dt}\right|_{t=0} &= \left\langle h,-[\xi,e_i]\right\rangle = -\left\langle h,C^k_{ij}\xi^j e_k\right\rangle \text{ with } [\xi,e_i] = \left[\xi^j e_j,e_i\right] = C^k_{ji}\xi^j e_k \\
\left.\frac{d\left\langle h,Ad_{e^{-t\xi}}e_i\right\rangle}{dt}\right|_{t=0} &= \left\langle h_k e^k,-C^k_{ji}\xi^j e_k\right\rangle = -C^k_{ji}\xi^j h_k \\
\left.\frac{dF^*\left(Ad^*_{e^{t\xi}}h\right)}{dt}\right|_{t=0} &= -C^k_{ji}\xi^j h_k\frac{\partial F^*}{\partial X_i}(\xi) = (-\phi(\xi)F^*)(h)
\end{aligned}
\tag{187}
$$

We use then Taylor expansion of $F^*\left(Ad^*_{e^{t\xi}}h\right)$ given by:

$$
F^*\left(Ad^*_{e^{t\xi}}h\right) = F^*(h) + \sum_{n=1}^{\infty}\frac{(-\phi(\xi))^n F^*}{n!}(h).t^n
\tag{188}
$$

We can observe that F^* is invariant if $F^*\left(Ad^*_{e^{t\xi}}h\right) = F^*(h)$ and then $(-\phi(\xi))^n F^* = 0$ or $\phi(\xi)F^* = 0$ that could be written $C^k_{ji}\xi^j h_k\frac{\partial F^*}{\partial X_i}(\xi) = 0$.

If F^* is semi-invariant of the coadjoint representation of group if and only if:

$$
\phi(e_i)F^* = -\lambda_i F^* \text{ with } \lambda_i = d\chi(e_i) \text{ (d}\chi: \text{ derivative of } \chi \text{ at the group } G \text{ identity element}
$$

$$
\begin{aligned}
&F^*\left(Ad^*_{e^{t\xi}}h\right) = \chi\left(e^{t\xi}\right)F^*(h) \text{ with } \chi\left(e^{t\xi}\right) = e^{t\chi_*(\xi)} \\
&\left[(-\phi(\xi))^n F^*\right](h) = \chi_*(\xi)F^*(h) \\
&\Rightarrow F^*\left(Ad^*_{e^{t\xi}}h\right) = \left[1 + \sum_{n=1}^{\infty}\frac{[\chi_*(\xi)]^n}{n!}t^n\right].F^*(h)
\end{aligned}
\tag{189}
$$

9. Conclusion: Lie Groups Thermodynamics for Machine Learning

With Lie groups Thermodynamics, we have presented Souriau tools to extend Gibbs density for Lie groups [95–107]. We can make reference to other explorations of Lie Group Representation theory to built exponential families [108–111] or Information Geometry in Quantum Physics [112–123]. Gibbs density estimation is a basic tool in statistical macine learning. Classically, we can associate to any posterior distribution an effective generalized geometric temperature, given by an element of the dual space of the Lie algebra, relating it to the Gibbs prior distribution. Classification rules could be introduced by Gibbs measures defined on parameter sets and depending on the observed sample value. A Gibbs measure is a special kind of probability measure used in statistical mechanics to describe the state of a particle system driven by a given energy function at some given temperature. Gibbs measures will be realized as minimizers of the average loss value under entropy constraints. In this extension for Lie groups, an important tool is the log-Laplace transform related to the Massieu Characteristic Function in Thermodynamics (a re-parameterization of the free energy by Planck temperature preserving Legendre transform with respect to Entropy). As we want to deal with Lie group data for Machine Learning, we will consider tools very similar to those used in statistical mechanics to describe particle systems with many degrees of freedom. Classification rules could be described by Gibbs measures defined on parameter sets and depending on the observed sample value. Comparing any posterior distribution with a Gibbs prior distribution make it possible to provide a way to build an estimator which can be proved to reach adaptively at the best possible asymptotic error rate (by temperature selection of a Gibbs posterior distribution built within a single parametric model). Estimators derived from Gibbs posteriors show excellent performance in diverse tasks, such as classification, regression and ranking. The usual recommendation is to sample from a Gibbs posterior

using MCMC (Markov chain Monte Carlo). With covariant Souriau Gibbs density, it is possible to extend MCMC and Gibbs sampler approach for Lie Groups Machine Learning.

More recently, the use of perturbation techniques has been proposed as an alternative to MCMC techniques for sampling. These results have been extended in conditional random fields loss, proving that the maximum in expectation with low-rank perturbations, provides an upperbound on the log partition (what we call Massieu characteristic function). New lower bounds on the partition function and new unbiased sequential sampler for the Gibbs distribution based on low-rank perturbations have been introduced. All these methods are based on sampling from the Gibbs distribution, upper-bounding the log partition function. All these results are synthetized in [124], where they also propose a new general method, with connections to the recently-proposed Fenchel-Young losses [125], using doubly stochastic scheme for minimization of these losses, for unsupervised and supervised learning. This is a generalization to the Gibbs distribution. Methods for learning parameters of a Gibbs distribution on data $(y_i)_{i=1,...,n}$ are based on maximization of the likelihood:

$$\hat{l}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n}\log p_{Gibbs,\theta}(y_i) = \frac{1}{n}\sum_{i=1}^{n}\langle y_i, \theta \rangle - \log \psi(\theta) \tag{190}$$

that is optimized by gradients methods using the empirical log-likelihood, given by:

$$\nabla_\theta \hat{l}_n(\theta) = \hat{y}_n - E_{Gibbs,\theta}[y] \tag{191}$$

For this method of moment-matching, the expectation of the Gibbs distribution is a challenge in some cases. This approach has been replaced by computing p_θ, with a method called "*perturb-and-MAP*" to learn the parameters in this model as a proxy for log-likelihood. This minimization is equivalent to maximizing previous equation by substituting the log-partition $\log \psi(\theta)$ with:

$$F_\varepsilon(\theta) = E[F(\theta + \varepsilon V)] = E\left[\max_{y \in C}\langle y, \theta + \varepsilon V \rangle\right] \text{ with a random noise vector}$$
$$\varepsilon V, \varepsilon > 0 \tag{192}$$

This approach could be linked with the use of Fenchel-Young losses [125]. In the perturbed model, the Fenchel-Young loss is given by:

$$L_\varepsilon(\theta; y) = F_\varepsilon(\theta) + \varepsilon \Omega(y) - \langle \theta, y \rangle = D_{\varepsilon\Omega}(y, \hat{y}_\varepsilon^*(\theta)) \tag{193}$$

with loss gradient $\nabla L_\varepsilon(\theta; y) = \nabla F_\varepsilon(\theta) - y = y_\varepsilon^*(\theta) - y$ where $y_\varepsilon^*(\theta) = E_{p_\theta(y)}[y] = E\left[\operatorname*{argmax}_{y \in C}\langle y, \theta + \varepsilon V \rangle\right]$ and $D_{\varepsilon\Omega}(y, \hat{y}_\varepsilon^*(\theta))$ Bregman divergence associated to $\varepsilon\Omega$. As F_ε generalizes the log-sum-exp function on the simplex, its dual Ω is a generalization of the negative entropy (which is the Fenchel dual of log-sum-exp).These connections have been studied in [126].

To conclude, we have seen that Lie group tools based on Representation Theory and Orbits Methods could be used with Souriau-Fisher Metric on Coadjoint Orbits as an extension of Fisher Metric for Lie group through homogeneous Symplectic Manifolds on Lie group Co-Adjoint Orbits.

We can then beneficiate of different tools based on Souriau Lie groups Thermodynamics and Kirillov Representation Theory, as illustrated in Figure 6, for:

- **Supervised Machine Learning**

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

- ○ Souriau Maximum 2164Entropy Density on Co-Adjoint Orbits: Covariant Maximum Entropy Probability Density for Lie groups defined with Souriau Moment Map, Co-Adjoint Orbits Method & Kirillov Representation Theory
- ○ Symplectic Integrator preserving Moment Map: Extension of Neural Network Natural Gradient to Geometric Integrators as Symplectic integrators that preserve moment map

- **Non-Supervised Machine Learning**

- ○ Souriau Exponential Map on Lie Algebra: Exponential Map for Geodesic Natural Gradient on Lie Algebra based on Souriau Algorithm for Matrix Characteristic Polynomial
- ○ Fréchet Geodesic Barycenter by Hermann Karcher Flow: Extension of Mean/Median on Lie group by Fréchet Definition of Geodesic Barycenter on Souriau-Fisher Metric Space, solved by Karcher Flow.
- ○ Mean-Shift on Lie groups with Souriau-Fisher Distance: Extension of Mean-Shift for Homogeneous Symplectic Manifold and Souriau-Fisher Metric Space.
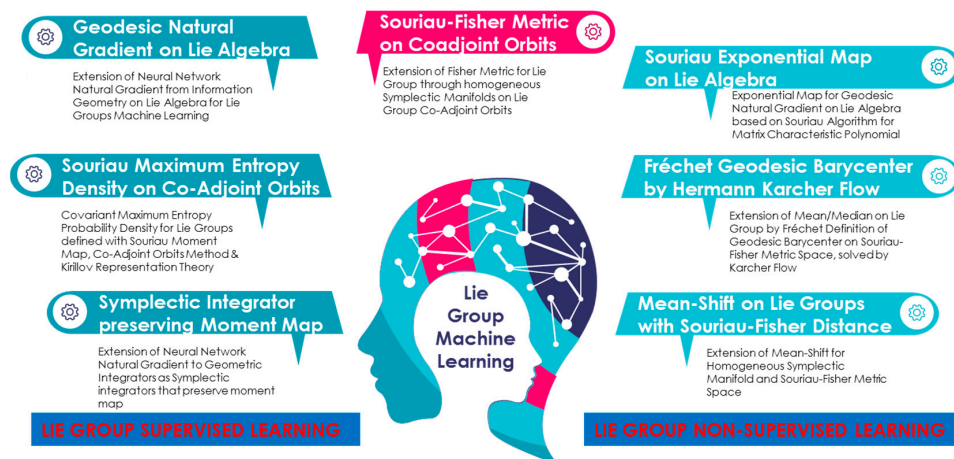
Figure 6. Supervised/Non-Supervised Machine Learning based on Lie Groups Thermodynamics.

[There is nothing more in physical theories than symmetry groups except the mathematical construction which allows precisely to show that there is nothing more] « Il n'y a rien de plus dans les théories physiques que les groupes de symétrie si ce n'est la construction mathématique qui permet précisément de montrer qu'il n'y a rien de plus ». **Jean-Marie Souriau** (see Figure 7)

La notion classique d'ensemble canonique de Gibbs est étendue au cas d'une variété symplectique sur laquelle un groupe de Lie possède une action symplectique ("groupe dynamique"). La définition rigoureuse donnée ici permet d'étendre un certain nombre de propriétés thermodynamiques classiques (la température est ici un élément de l'algèbre de Lie du groupe, la chaleur un élément de son dual), notamment des inégalités de convexité. Dans le cas de groupes non commutatifs, des propriétés particulières apparaissent: la symétrie est spontanément brisée, certaines relations de type cohomologique sont vérifiées dans l'algèbre de Lie du groupe [The classical notion of Gibbs' canonical ensemble is extended to the case of a symplectic manifold on which a Lie group has a symplectic action ("dynamic group"). The rigorous definition given here makes it possible to extend a certain number of classical thermodynamic properties (the temperature here is an element of the Lie group algebra, heat an element of its dual), notably inequalities of convexity. In the case of non-commutative groups, particular properties appear: the symmetry is spontaneously broken, certain relations of cohomological type are verified in the Lie algebra of the group]. **Jean-Marie Souriau, Mécanique Statistique, Groupes de Lie et Cosmologie, colloque CNRS n°237 – Géométrie Symplectique et physique mathématique**
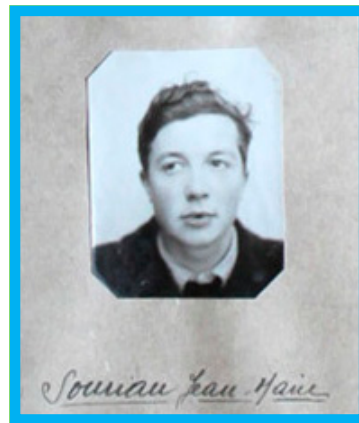
Figure 7. Jean-Marie Souriau, student at ENS Paris 1942.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Coadjoint orbits and Moment Map for SU(1,1)

We give more details developpements to obtain *SU(1,1)*/K coadjoint orbit and moment map from [127]. SU(1,1) has been intensively studied for Coherent states in Quantum Physics [128].

If we consider hyperbolic Group $SL(2,R)$

$$SL(2,R) = \left\{ \begin{pmatrix} m & p \\ q & n \end{pmatrix} \in GL(2,R) / mn - pq = 1 \right\} \tag{A1}$$

Elements of $SL(2,R)$ could be written with Iwasawa decomposition:

$$g = k_\theta.a_t.n_b \text{ with } k_\theta \in K, a_t \in A, n_b \in N \tag{A2}$$

$$
\begin{aligned}
K &= \left\{ k_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} / 0 \le \theta < 2\pi \right\} \text{ with } e^{-i\theta} = \frac{m-iq}{\sqrt{m^2+q^2}} \\
A &= \left\{ a_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} / t \in R \right\} \text{ with } e^{2t} = m^2 + q^2 \\
N &= \left\{ n_b = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} / b \in R \right\} \text{ with } b = \frac{mp+qn}{\sqrt{m^2+q^2}}
\end{aligned}
\tag{A3}
$$

K is a maximal compact sub-group of G_{SL}.

Appendix A.1. Group of Unit Disk Automorphisms

Consider $SU(1,1)$ sub-group of $SL(2,R)$ given by

$$SU(1,1) = \left\{ A = \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} / \det(A) = |a|^2 - |b|^2 = 1, \ a,b \in R \right\} \tag{A4}$$

The following interior automorphism that transforms $SL(2,R)$ to $SU(1,1)$, inducing an isomorphism between them:

$$SL(2,R) \to SU(1,1)$$
$$g \mapsto CgC^{-1} \text{ with } C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \tag{A5}$$

$SU(1,1)$ acts on the Poincaré unit disk $D = \{z \in C/|z| < 1\}$:

$$g^{-1}.z = \frac{az+b}{b^*+a^*} \text{ with } g^{-1} = \begin{pmatrix} a & b \\ b^* & a^* \end{pmatrix} \in SU(1,1) \text{ and } z \in D \tag{A6}$$

Valentine Bargmann has parameterized $SU(1,1)$:

$$\begin{cases} \gamma = \frac{b}{a} \\ \omega = \arg(\alpha) \bmod 2\pi \end{cases} \Rightarrow \begin{cases} |\gamma| < 1 \\ a = e^{i\omega}\left(1 - |\gamma|^2\right)^{-1/2} \\ b = e^{i\omega}\gamma\left(1 - |\gamma|^2\right)^{-1/2} \end{cases} \tag{A7}$$

Then $SU(1,1) = \left\{(\gamma,\omega)/|\gamma| < 1, \omega \in \,]-\pi, +\pi]\right\}$ with Group composition law given by:

$$\begin{aligned} &(\gamma,\omega).(\gamma',\omega') = (\gamma'',\omega'') \\ &\begin{cases} \gamma'' = \left[\gamma' + \gamma e^{-2i\omega''}\right]\left[1 + \gamma\gamma^{*'}e^{2i\omega'}\right]^{-1} \\ \omega'' = \omega + \omega' + \frac{1}{2i}\log\left(\left[1 + \gamma^{*'}\gamma e^{-2i\omega'}\right]\left[1 + \gamma'\gamma^* e^{-2i\omega'}\right]^{-1}\right) \bmod 2\pi \\ g = (\gamma,\omega) \Rightarrow g^{-1} = \left(-\gamma e^{2i\omega}, -\omega\right) \end{cases} \end{aligned} \tag{A8}$$

$SU(1,1)$ is topological product of unit disk and circle.

Appendix A.2. Universal covering of $SL(2,R)$

If we consider $G = \left\{(\gamma,\omega)/|\gamma| < 1, \omega \in R\right\}$, the following mapping:

$$\begin{aligned} \Theta : G &\to SU(1,1) \\ \Theta(\gamma,\omega) &= (\gamma, \omega \bmod 2\pi) \end{aligned} \quad \text{with } Ker\Theta = \{(0,2k\pi)/k \in Z\} \tag{A9}$$

Topological product of unit Disk D dans C and real straight line R is the universal covering of $SU(1,1)$.

A maximal compact subgroup of $SU(1,1)$ is $CKC^{-1} = \left\{\begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}/\theta \in R\right\}$ and the subgroup for G is $\Theta^{-1}\left(CKC^{-1}\right) = \{(0,\theta)/\theta \in R\}$.

Pukanszky and Sally have defined irreductible unitary representation of $\widetilde{SL}(2,R)$, classified in principal serie, discrete serie and complemantary serie.

The Lie algebra g of G and $SU(1,1)$ is given by:

$$l_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, l_2 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, l_0 = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{A10}$$

with the commutation relation: $[l_0,l_1] = l_2$, $[l_1,l_2] = -l_0$, $[l_2,l_0] = l_1$.

The dual space of the Lie algebra g* of g to g thanks to Killing form:

$$B\left(\sum_{i=1,2,0} x_i l_i, \sum_{i=1,2,0} y_i l_i\right) = 2(x_1 y_1 + x_2 y_2 - x_0 y_0) \tag{A11}$$

Appendix A.3. Coadjointes Orbits of $SU(1,1)$

Considering adjoint representation $Ad_g : g \to g$, and coadjointe representation as transpose linear mapping of $Ad_{g^{-1}}$, written by $Ad_g^* = \left(Ad_{g^{-1}}\right)^*$, $Ad_g^* : g^* \to g^*$.

Let $f = (\eta, 0, 0)$ in base $\{l_1, l_2, l_0\}$ with $\eta > 0$ coadjoint orbit of f is:

$$O_\eta = \left\{ gfg^{-1}/g = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in SL(2, R) \right\}$$

$$O_\eta = \{\eta(1 + 2bc), \eta(bd - ac), \eta(bd + ac)/a, b, c, d \in R \text{ with } ad - bc = 1\}$$

(A12)

Stabilizer of f is

$$G_1(f) = \{g \in G/g.f = f\} = A \cup \{-A\} \text{ with } A = \left\{ a_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} /t \in R \right\}$$

(A13)

$SL(2, R) = K'NA'$ with $A' = A \cup \{-A\}$ where $N = \left\{ n_b = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} /b \in R \right\}$ and $K' = \left\{ k_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} /0 \le \theta < \pi \right\}$ and as $SL(2, R)/G_1(f)$ is a bijection with O_η, then O_η is in bijection with $K'N$, diffeomorph to $K/\{I_d\} \times N$. Then all element $x \in O_\eta$ can be written through this bijection $x = k_\theta n_b, k_\theta \in K', n_b \in N$. O_η is set of points $l = (x_1, x_2, x_0) \in g^*$ such that: $x_1^2 + x_2^2 - x_0^2 = \eta^2 > 0$ a one sheet hyperboloid in g^*.

We will study discrete sequence.

Appendix A.4. Quantization of Kähler Manifold

Let (M, ω) a Kalherian manifold of dimension 2n (complex manifold with complex structure J) with 2-form $g(X, Y) = \omega(X, JY)$, a Riemannian structure on M.

We have seen that $SU(1, 1)$ is conjugate of $SL_2(R)$ in $SL(2, R)$

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \in SU(1, 1), \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(R)$$

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} = \frac{-i}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} i & i \\ -1 & 1 \end{pmatrix}$$

(A14)

$$\begin{cases} \alpha = \frac{1}{2}[(a + d) + i(b - c)] \\ \beta = \frac{1}{2}[(a - d) - i(b + c)] \end{cases} \Rightarrow \begin{cases} \alpha + \alpha^* = a + d \\ \alpha - \alpha^* = i(b - c) \\ \beta + \beta^* = a - d \\ \beta - \beta^* = -i(b + c) \end{cases} \Rightarrow \begin{cases} a = \frac{1}{2}[(\alpha + \alpha^*) + (\beta + \beta^*)] \\ b = \frac{1}{2i}[(\alpha - \alpha^*) - (\beta - \beta^*)] \\ c = \frac{-1}{2i}[(\alpha - \alpha^*) + (\beta - \beta^*)] \\ d = \frac{1}{2}[(\alpha + \alpha^*) - (\beta + \beta^*)] \end{cases}$$

(A15)

Let $f = \left(0, 0, h - \frac{1}{2}\right)$ in base $\{l_1, l_2, l_0\}$ where $l_1 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, l_2 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, l_0 = \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ with $h > \frac{1}{2}$ (respectively $f = \left(0, 0, h + \frac{1}{2}\right)$ with $h < -\frac{1}{2}$), the coadjoint orbit of f is:

$$O_h = \left\{ gfg^{-1}/g = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in SL(2, R) \right\}$$

$$O_h = \left\{ \left(|h| - \frac{1}{2}\right)\text{sign}h(ac + bd); \frac{(|h| - \frac{1}{2})\text{sign}h}{2}\left(d^2 + c^2 - \left(a^2 + b^2\right)\right); \frac{(|h| - \frac{1}{2})\text{sign}h}{2}\left(d^2 + c^2 + a^2 + b^2\right) \right\}$$

(A16)

O_h appears as points $l = (x_1, x_2, x_0) \in g^*$ such that:

$$x_1^2 + x_2^2 - x_0^2 = -\left(|h|^2 - \frac{1}{2}\right)^2 < 0 \text{ with } \begin{cases} x_0 > 0 \text{ if } h > \frac{1}{2} \\ x_0 < 0 \text{ if } h < -\frac{1}{2} \end{cases}$$

(A17)

O_h is one of hyperboloid sheets of g^*, associated to representation π_h of discrete serie of G.

$$\begin{cases} x_1 = h'(ac + bd) \\ x_2 = \frac{h'}{2}\left(d^2 - a^2 + c^2 - b^2\right) \quad \text{with } h' = \left(|h| - \frac{1}{2}\right)\text{sign}h \\ x_0 = \frac{h'}{2}\left(d^2 + a^2 + c^2 + b^2\right) \end{cases} \tag{A18}$$

Then, we obtain:

$$\begin{cases} x_1 = ih'(\alpha\beta - \alpha^*\beta^*) \\ x_2 = -h'(\alpha\beta + \alpha^*\beta^*) \\ x_0 = h'(\alpha\alpha^* + \beta\beta^*) \end{cases} \tag{A19}$$

If we set:

$$\begin{cases} \alpha = \frac{e^{i\theta}}{\sqrt{1-|z|^2}} \\ \beta = \frac{ze^{-i\theta}}{\sqrt{1-|z|^2}} \end{cases} \Rightarrow \begin{cases} x_1 = \frac{ih'(z-z^*)}{1-|z|^2} \\ x_1 = \frac{-h'(z+z^*)}{1-|z|^2} \\ x_0 = \frac{h'(1+|z|^2)}{1-|z|^2} \end{cases} \tag{A20}$$

We use the parametrization of O_h by unit disk D:

$$\begin{aligned} D = \{z \in C / |z| < 1\} &\to O_h \\ z &\mapsto (x_1, x_2, x_0) \end{aligned} \tag{A21}$$

$$\begin{cases} x_1 = -2\left(|h| - \frac{1}{2}\right)\text{sign}h\frac{\text{Im}(z)}{1-|z|^2} \\ x_1 = -2\left(|h| - \frac{1}{2}\right)\text{sign}h\frac{\text{Re}(z)}{1-|z|^2} \\ x_0 = \left(|h| - \frac{1}{2}\right)\text{sign}h\frac{1+|z|^2}{1-|z|^2} \end{cases} \tag{A22}$$

This parametrisation provides a kahlerian structure for O_h, inherit from D.

We have $O_h = G/\widetilde{K} = SL(2, R)/K$, and as $SL(2, R)$ is isomorph to $SU(1, 1)$, O_h is identified with $SU(1, 1)/K_0 = D$. Stabilizer of the origin is $K_0 = \left\{\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^* \end{pmatrix} / |\alpha| = 1\right\}$. Then O_h is globally diffeomorph to D for all h.

We can compute Liouville measure on O_h. This measure is $\omega_h = \frac{\omega_h'}{2\pi}$ where ω_h' is the canonical 2-form on orbit. Parameterization on O_h gives:

$$\omega_h = \frac{\omega_h'}{2\pi} = \frac{-i(2|h| - 1)}{2\pi\left(1 - |z|^2\right)}\text{sign}h|dz|^2 \tag{A23}$$

We can observe that:

$$\begin{cases} dx_1 = \frac{i}{(1-|z|^2)}\left[(z - z^*)(1 - zz^*)dh' + h'\left(1 - z^{*2}\right)dz - h'\left(1 - z^2\right)dz^*\right] \\ dx_2 = \frac{-1}{(1-|z|^2)}\left[(z + z^*)(1 - zz^*)dh' + h'\left(1 + z^{*2}\right)dz - h'\left(1 + z^2\right)dz^*\right] \\ dx_0 = \frac{1}{(1-|z|^2)}\left[(1 + zz^*)(1 - zz^*)dh' + 2z^*h'dz - 2zh'dz^*\right] \end{cases} \tag{A24}$$

$$dx_1 dx_2 dx_0 = \frac{-ih'^2\left(1 - |z|^2\right)}{\left(1 - |z|^2\right)^6}\det\begin{pmatrix} z - z^* & 1 - |z|^2 & -\left(1 - z^2\right) \\ z + z^* & 1 + z^{*2} & 1 + z^2 \\ 1 + |z|^2 & 2z^* & 2z \end{pmatrix} \tag{A25}$$

$$dx_1 dx_2 dx_0 = \frac{-2ih'^2}{\left(1 - |z|^2\right)}dh'|dz|^2 = \frac{2}{\pi}h'dh'\omega_h \tag{A26}$$

This the measure defined on open set of g^* given by $l_1^2 + l_2^2 - l_0^2 < 0$.

Appendix B. Bargman Parameterization of SU(1,1)

$SU(1,1)$ is isomorphic to $SL(2,R) = Sp(2,R)$ through the complex unitary matrix W:

$$SL(2,R) = \left\{ g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} / \det g = ad - bc = 1 \right\} \tag{A27}$$

$$Sp(2,R) = \left\{ g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} / gJg^T = J, J = \begin{pmatrix} 0 & +1 \\ -1 & 0 \end{pmatrix} \right\} \tag{A28}$$

$$W = \frac{1}{\sqrt{2}} \begin{pmatrix} \omega^{-1} & \omega^{-1} \\ -\omega & \omega \end{pmatrix} = \left(W^+ \right)^{-1} \text{ with } \omega = e^{i\pi/4} = \frac{1}{\sqrt{2}}(1+i) \tag{A29}$$

If we observe that $W^{-1}JW = -iM$, the isomorphism is given explicitly by:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = g(u) = WuW^{-1} = \begin{pmatrix} \text{Re}(\alpha + \beta) & -\text{Im}(\alpha - \beta) \\ \text{Im}(\alpha + \beta) & \text{Re}(\alpha - \beta) \end{pmatrix} \tag{A30}$$

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} = u(g) = W^{-1}gW = \frac{1}{2} \begin{pmatrix} (a+d) - i(b-c) & (a-d) + i(b+c) \\ (a-d) - i(b+c) & (a+d) + i(b-c) \end{pmatrix} \tag{A31}$$

We can also make also a link with $SO(2,1)$ of "1 + 2" pseudo-orthogonal matrices:

$$SO(2,1) = \left\{ \Gamma \in GL(3,3) / \det(\Gamma) = 1, \Gamma K \Gamma^T = \Gamma, K = \begin{pmatrix} +1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \right\} \tag{A32}$$

$$\Gamma(g) = \begin{pmatrix} \frac{1}{2}(a^2 + b^2 + c^2 + d^2) & \frac{1}{2}(a^2 - b^2 + c^2 - d^2) & -cd - ab \\ \frac{1}{2}(a^2 + b^2 - c^2 - d^2) & \frac{1}{2}(a^2 - b^2 - c^2 + d^2) & cd - ab \\ -bd - ac & bd - ac & ad + bc \end{pmatrix} \tag{A33}$$

with $\Gamma(g_1)\Gamma(g_2) = \Gamma(g_1 g_2), \Gamma(I) = I, \Gamma(g^{-1}) = \Gamma(g)^{-1}$

The $SO(2,1)$ matrix corresponds to any $SU(1,1)$:

$$\Gamma(u) = \begin{pmatrix} |\alpha|^2 + |\beta|^2 & 2\text{Re}\alpha\beta^* & 2\text{Im}\alpha\beta^* \\ 2\text{Re}\alpha\beta & \text{Re}(\alpha^2 + \beta^2) & \text{Im}(\alpha^2 - \beta^2) \\ -2\text{Im}\alpha\beta & -\text{Im}(\alpha^2 + \beta^2) & \text{Re}(\alpha^2 - \beta^2) \end{pmatrix} \tag{A34}$$

and $\alpha = \pm \sqrt{\frac{1}{2}(\Gamma_{11} + \Gamma_{12}) + i(\Gamma_{12} - \Gamma_{21})}, \beta = \frac{1}{2\alpha}(\Gamma_{10} - i\Gamma_{20})$

The properties of connectivity of $Sp(2,R)$ is described by its isomorphy with $SU(1,1)$. Using unimodular condition:

$$|\alpha|^2 - |\beta|^2 = 1 \Rightarrow \alpha_R^2 + \alpha_I^2 - \beta_R^2 = 1 + \beta_I^2 \geq 1 \text{ with } \alpha = \alpha_R + i\alpha_I \text{ and } \beta = \beta_R + i\beta_I$$

If β_I is fixed, $(\alpha_R, \alpha_I, \beta_R)$ are constrained to define a one-sheeted revolution hyperboloid, with its circular waist in the α plane.

To $SU(1,1)$, we can associate the simply-connected universal covering group, using the maximal compact subgroup $U(1)$ and corresponding to the Iwasawa decomposition (factorization of a noncompact semisimple group into its maximal compact subgroup times a solvable subgroup).

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} = \begin{pmatrix} e^{i\omega} & 0 \\ 0 & e^{i\omega} \end{pmatrix} \begin{pmatrix} \lambda & \mu \\ \mu^* & \lambda \end{pmatrix} \text{ with } \begin{cases} \omega = \arg\alpha = \frac{1}{2}i\ln(\alpha^*\alpha^{-1}) \\ \lambda = |\alpha| > 0 \\ \mu = e^{-i\omega}\beta = \sqrt{\frac{\alpha^*}{\alpha}}\beta \end{cases} \tag{A35}$$

$$\beta = e^{i\omega}\mu, |\alpha|^2 - |\beta|^2 = \lambda^2 - |\mu|^2 = 1 \text{ so } |\mu| < \lambda \tag{A36}$$

Bargmann has generalized this parameterization for $Sp(2N, R)$, more convenient but difficult to generalize to N dimensions. For $SU(1,1)$:, Bargmann has used (ω, γ):

$$\gamma = \frac{\mu}{\lambda} = \frac{\beta}{\alpha} \left(|\gamma| < 1\right), \lambda = \frac{1}{\sqrt{1 - |\gamma|^2}}, \mu = \frac{\gamma}{\sqrt{1 - |\gamma|^2}} \tag{A37}$$

For $SL(2,R) = Sp(2,R)$, the Bargman, parameterization is given by this decomposition of a non-singular matrix into the product of an orthogonal and a positive definite symmetric matrix:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix} \begin{pmatrix} \lambda + \text{Re}\mu & \text{Im}\mu \\ \text{Im}\mu & \lambda - \text{Re}\mu \end{pmatrix} \tag{A38}$$

Conversely: $\omega = \arg[(a + d) - i(b - c)], \mu = e^{-i\omega}[(a - d) + i(b + c)]$

ω is counted modulo 2π, $\omega \equiv \omega(\text{mod}2\pi)$.

$SU(1,1)$ and $SL(2,R) = Sp(2,R)$ are described when ω is counted modulo 2π, $\omega \equiv \omega(\text{mod}2\pi)$.

Valentine Bargmann has proposed the covering of the general symplectic group $Sp(2N, R)$:

$$Sp(2N,R) = \left\{ g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} / gJ_{2N}g^T = J_{2N}, J_{2N}^T = -J_{2N}, J_{2N} = \begin{pmatrix} 0 & I_N \\ -I_N & 0 \end{pmatrix} \right\} \tag{A39}$$

with relations:

$$AB^T = BA^T, AC^T = CA^T, BD^T = DB^T, CD^T = DC^T, AD^T - BC^T = I_N \tag{A40}$$

$$g \in Sp(2N,R) \Rightarrow g^{-1} = M_{2N}g^T M_{2N} = \begin{pmatrix} D^T & -B^T \\ -C^T & A^T \end{pmatrix} \tag{A41}$$

Bargmann has observed that although $Sp(2N, R)$ is not isomorphic to any pseudo-unitary group, its inclusion in $U(N, N)$ will display the connectivity properties through its unitary $U(N)$ maximal compact subgroup, generalizing the role of $U(1) = SO(2)$ in $Sp(2, R)$.

$W_N = W \otimes I_N$ a $2N \times 2N$ matrix where $W = W_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} \omega_{\pi/4}^{-1} & \omega_{\pi/4}^{-1} \\ -\omega_{\pi/4} & \omega_{\pi/4} \end{pmatrix}$ with $\omega = e^{i\pi/4} = \frac{1}{\sqrt{2}}(1 + i)$, which gives the $N \times N$ block coefficients.

$$u(g) = W_N^{-1} g W_N = \frac{1}{2} \begin{pmatrix} [A+D] - i[B-C] & [A-D] + i[B+C] \\ [A-D] - i[B+C] & [A+D] + i[B-C] \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \tag{A42}$$

with

$$\begin{cases} \alpha\alpha^+ - \beta\beta^+ = I_N, \alpha^+\alpha - \beta^T\beta^* = I_N \\ \alpha\beta^T - \beta\alpha^T = 0, \alpha^T\beta^* - \beta^+\alpha = 0 \end{cases} \tag{A43}$$

and

$$u^{-1} = M_{2N}u^+ M_{2N}^{-1} = \begin{pmatrix} \alpha^+ & -\beta^T \\ -\beta^+ & \alpha^T \end{pmatrix} \tag{A44}$$

The symplecticity property of g becomes:

$$uM_{2N}u^+ = M_{2N}, M_{2N} = iW_N^{-1}J_{2N}W_N = \begin{pmatrix} I_N & 0 \\ 0 & -I_N \end{pmatrix} \tag{A45}$$

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = g(u) = W_N u W_N^{-1} = \begin{pmatrix} \text{Re}(\alpha + \beta) & -\text{Im}(\alpha - \beta) \\ \text{Im}(\alpha + \beta) & \text{Re}(\alpha - \beta) \end{pmatrix} \tag{A46}$$

Valentine Bargmann has extended the well-know theorem that any real matrix R may be decomposed into the product of an orthogonal Q and a symmetric positive definite matrix S, uniquely as $R = QS$. Bargmann has shown that if $R \in Sp(2N, R)$, then $R = QS$ with $Q, S \in Sp(2N, R)$ where Q maps onto unitary matrix and S maps onto Hermitian positive definite matrix:

$$u(Q) = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^* \end{pmatrix}, \alpha\alpha^+ = I_N, \alpha \in U(N) \text{ and } u(S) = \exp\begin{pmatrix} 0 & \xi \\ \xi^* & 0 \end{pmatrix}, \xi = \xi^T \tag{A47}$$

We can generalize Bargmann parameterization of $SU(1,1)$ to $Sp(2N, R)$:

$$u\{\omega, \lambda, \mu\} = \begin{pmatrix} e^{i\omega}I_N & 0 \\ 0 & e^{-i\omega}I_N \end{pmatrix} \begin{pmatrix} \lambda & \mu \\ \mu^* & \lambda^* \end{pmatrix} \oplus, \det\lambda > 0 \tag{A48}$$

Then the Bargmann parameters are:

$$\omega = \frac{1}{N}\arg\det\alpha, \lambda = e^{-i\omega}\alpha, \mu = e^{-i\omega}\beta, e^{iN\omega} = \frac{\det\alpha}{|\det\alpha|}, \det\lambda = |\det\alpha| > 0 \tag{A49}$$

The $Sp(2N, R)$ matrices in terms of the Bargmann parameters are:

$$g\{\omega, \lambda, \mu\} = \begin{pmatrix} \cos\omega I_N & -\sin\omega I_N \\ \sin\omega I_N & \cos\omega I_N \end{pmatrix} \begin{pmatrix} \mathrm{Re}(\lambda + \mu) & -\mathrm{Im}(\lambda - \mu) \\ \mathrm{Im}(\lambda + \mu) & \mathrm{Re}(\lambda - \mu) \end{pmatrix} \tag{A50}$$

V. Bargmann has proposed the covering of the general symplectic group $Sp(2N, R)$:

$$Sp(2N, R) = \left\{ g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} / gJ_{2N}g^T = J_{2N}, J_{2N}^T = -J_{2N}, J_{2N} = \begin{pmatrix} 0 & I_N \\ -I_N & 0 \end{pmatrix} \right\} \tag{A51}$$

$$AB^T = BA^T, AC^T = CA^T, BD^T = DB^T, CD^T = DC^T, AD^T - BC^T = I_N \tag{A52}$$

Bargmann has observed that although $Sp(2N, R)$ is not isomorphic to any pseudo-unitary group, its inclusion in $U(N, N)$ will display the connectivity properties through its unitary $U(N)$ maximal compact subgroup, generalizing the role of $U(1) = SO(2)$ in $Sp(2, R)$:

$W_N = W \otimes I_N, 2N \times 2N$ matrix where $W = W_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} \omega_{\pi/4}^{-1} & \omega_{\pi/4}^{-1} \\ -\omega_{\pi/4} & \omega_{\pi/4} \end{pmatrix}$ with $\omega = e^{i\pi/4} = \frac{1}{\sqrt{2}}(1 + i)$.

$$u(g) = W_N^{-1}gW_N = \frac{1}{2}\begin{pmatrix} [A + D] - i[B - C] & [A - D] + i[B + C] \\ [A - D] - i[B + C] & [A + D] + i[B - C] \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \tag{A53}$$

with

$$\alpha\alpha^+ - \beta\beta^+ = I_N, \alpha^+\alpha - \beta^T\beta^* = I_N \text{ and } \alpha\beta^T - \beta\alpha^T = 0, \alpha^T\beta^* - \beta^+\alpha = 0 \tag{A54}$$

The symplecticity property of g becomes:

$$uM_{2N}u^+ = M_{2N}, M_{2N} = iW_N^{-1}J_{2N}W_N = \begin{pmatrix} I_N & 0 \\ 0 & -I_N \end{pmatrix} \tag{A55}$$

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = g(u) = W_N uW_N^{-1} = \begin{pmatrix} \mathrm{Re}(\alpha + \beta) & -\mathrm{Im}(\alpha - \beta) \\ \mathrm{Im}(\alpha + \beta) & \mathrm{Re}(\alpha - \beta) \end{pmatrix} \tag{A56}$$

Appendix C. Shirokov Method to build Casimir function

I. V. Shirokov [71–75] has proposed a method for constructing invariants of the coadjoint representation of Lie groups with an arbitrary dimension and structure based on local symplectic

coordinates on the coadjoint orbits. They also have made link with construction of the Lie algebra polarization

$B_\lambda(X, Y)$ be the skewsymmetric form on g defined by:

$$B_\lambda(X, Y) = \langle \lambda, [X, Y] \rangle \, , X, Y \in g \, , \lambda \in g^* \\ \text{with } B_\lambda(X, Y) = B_{ij} X^i Y^j \, , B_{ij} = C_{ij}^k \lambda_k \equiv C_{ij}(\lambda) \, , i, j, k = \dim g \tag{A57}$$

$$\text{Ker} B_\lambda = \{X \in g / B_\lambda(X, g) = 0\} \equiv g^\lambda \text{ annihilator of } \lambda \in g^* \tag{A58}$$

The polarization h $\equiv$ P(g) of the the covector λ is given by:

$$B_\lambda(h, h) = 0 \, , \dim h = \frac{1}{2} \left(\dim g + \dim g^\lambda \right) \tag{A59}$$

The polarization of a semisimple Lie algebra is the construction of polarization of the reductive Lie algebra g, where g $= C \oplus B$, where C is the center in g, and B is a semisimple subalgebra, given by:

$$h \equiv P(C) \oplus P(B) = C \oplus P(B) \tag{A60}$$

Let g be an arbitrary Lie algebra, R its radical and $R^\perp = \{X \in g / B_\lambda(X, R) = 0\}$ the orthogonal complement of R in g with respect to the form $B_\lambda(X, Y)$, then the polarization is given by:

$$P(g) \equiv P(R) \oplus P(R^\perp) = R(R) \oplus P(R^\perp) \tag{A61}$$

Oleg L. Kurnyavko and Igor V. Shirokov idea of the method of constructing coadjoint invariants is to construct the canonical transition to the Darboux coordinates on the orbits of the dual Lie algebra g* of maximal dimension dual to the Lie algebra g of the Lie group G.

If we consider a Lie group G, its Lie algebra g and its dua Lie algebra g*, and let each element $f \in g^*$ corresponds to the coordinates $(f_1, f_2, \ldots, f_n)$ where $n = \dim g^*$. The action of the group G on g* splits the dual Lie algebra into orbits O_λ where $\lambda \in g^*$ is some element of the given orbit. As with Souriau Theorem, these orbits is a symplectic manifold, with Darboux theorem, there is a special coordinates (p, q) on them such that the corresponding symplectic form can be written:

$$\omega_\lambda = dp_i \wedge dq^i \, , i = 1, \ldots, \frac{1}{2} \dim O_\lambda \tag{A62}$$

The transition to canonical coordinates is determined by the functions:

$$f_j = f_j(q, p, \lambda) \, , j = 1, 2, \ldots, \dim g^* \, , \text{ with } f_j(0, 0, \lambda) = \lambda_j \tag{A63}$$

Eliminating the variables q and p in these transition functions, we deduce the equation of the orbit:

$$\psi_k(f) = \omega_k \, , k = 1, \ldots, n - \dim O_\lambda \tag{A64}$$

We obtain $n - \dim O_\lambda$ functions such that:

$$\psi_k \left(Ad_g^* f \right) = \psi_k(f) \tag{A65}$$

Each nondegenerate orbit is determined by the set of numbers $\omega = \left\{ \omega_1, \omega_2, \ldots, \omega_{indg} \right\}$, that is fixed by the choice of the covector λ such that rank $C_{ij}(\lambda)$ is maximal, and then:

$$\omega_k = \psi_k(\lambda_1) = \psi_k(\lambda_2) \, , \text{ if } \lambda_1, \lambda_2 \in O_\lambda \tag{A66}$$

We can then consider $\lambda(j)$ orbits O_λ with $j = \left(j_1, j_2, \ldots, j_{indg} \right)$, with the constraint:

$$\psi_k(\lambda(j)) = \omega_k(j) \, , \det \frac{\partial \omega_k}{\partial j} \neq 0 \tag{A67}$$

If we set $\lambda_i(j) = a_i^l j_l$ and X_k^i the components of the vectors that form the basis of the annihilator of the covector $\lambda(j)$ such that:

$$\det \frac{\partial \omega_k(j)}{\partial j_k} = \det \frac{\partial \psi_k}{\partial f_i} \left. \frac{\partial f_i(j)}{\partial j_k} \right|_{f=\lambda(j)} = \det X_k^i \frac{\partial f_i(j)}{\partial j_k} = \det X_k^i (\lambda(j)) a_i^k \neq 0 \tag{A68}$$

Finally, we obtain the parametrization, considering with Darboux coordiantes:

$$f_i = f_i(q, p, j) \, , i = 1, \ldots, indg^* \tag{A69}$$

By eliminating the Darboux coordinate (q, p) in previous equation, we have new relations:

$$\varphi_k(f, j) = 0 \text{ and } j_k = j_k(f) \, , k = 1, \ldots, indg \tag{A70}$$

These relations provide invariants of the coadjoint representation of the Lie group *G*.

Igor V. Shirokov has proved that the desired linear transition to the Darboux coordinates on a nondegenerate orbit O_λ is defined by

$$f_i(p, q) = \xi_i^a(q) p_a + \xi_i^A(0, q) \lambda_A \text{ where } g = \{l_A, l_a\} \text{ with } l_A \text{ polarization basis of g} \tag{A71}$$

References

1. Duhem, P. L'intégrale des forces vives en thermodynamique. *JMPA* **1898**, *4*, 5–19.
2. Duhem, P. Sur l'équation des forces vives en thermodynamique et les relations de la thermodynamique avec la mécanique classique 23 December 1897. In *Procès-verbaux des Séances de la Société des Sciences Physiques et Naturelles de Bordeaux*; PVSScPhNB (1897-98): Bordeaux, France, 1898; pp. 23–27.
3. Duhem, P. *Sur deux Inégalites Fondamentales de la Thermodynamique, CR 156*; Gauthier-Villars: Paris, France, 1913; pp. 421–425.
4. Duhem, P. *Traité d'Énergetique ou Thermodynamique Générale. Tome Conservation de l'Énergie. Mécanique Rationelle. Statique Générale. Déplacement de l'Équilibre—Tome II. Dynamique Générale. Conductibilité de la Chaleur. Stabilité de l'Équilibre*; GauthierVillars: Paris, France, 1911; pp. 504–528.
5. Ollivier, Y.; Maceau-Caron, G. Practical Riemannian Neural Networks. *arxiv* **2016**, arXiv:abs/1602.08007.
6. Ollivier, Y. Riemannian metrics for neural networks I: Feedforward networks. *Inf. Inference J. IMA* **2015**, *4*, 108–153. [CrossRef]
7. Marceau-Caron, G.; Ollivier, Y.; Nielsen, F.; Barbaresco, F. Natural Langevin Dynamics for Neural Networks. In *Computer Vision*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2017; Volume 10589, pp. 451–459.
8. Ollivier, Y. True Asymptotic Natural Gradient Optimization. *arXiv* **2017**, arXiv:1712.08449.
9. Amari, S.; Karakida, R.; Oizumi, M. Fisher information and natural gradient learning in random deep networks. In Proceedings of the AISTATS 2019, Okinawa, Japan, 16–18 April 2019.
10. Zhang, G.; Martens, J.; Grosse, R. Fast convergence of natural gradient descent for over parametrized neural networks. *arXiv* **2019**, arXiv:1905.10961v1.
11. Amari, S.-I. Any Target Function Exists in a Neighborhood of Any Sufficiently Wide Random Network: A Geometrical Perspective. *arXiv preprint* **2020**, arXiv:2001.06931. [CrossRef]
12. Kirillov, A.A. *Elements of the Theory of Representations*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 1976; Volume 220.
13. Kosmann-Schwarzbach, Y. *Groupes et Symétries Groupes Finis, Groupes et Algèbres de Lie, Representations*; Ecole Polytechnique: Palaiseau, France, 2006.

14. Kosmann-Schwarzbach, Y. *Siméon-Denis Poisson, Les Mathématiques au Service de la Science*; Ecole Polytechnique: Palaiseau, France, 2013.

15. Laurent-Gengoux, C.; Pichereau, A.; Vanhaecke, P. *Poisson Structures, Grundlehren der Mathematischen Wissenschaften*; Springer: Berlin/Heidelberg, Germany, 2013.

16. Cartier, P. Groupoïdes de Lie et leurs algébroïdes. Séminaire BOURBAKI. ASTÉRISQUE. 2009, 326. Available online: http://preprints.ihes.fr/2008/M/M-08-20.pdf (accessed on 6 June 2020).

17. Barbaresco, F.; Gay-Balmaz, F. Lie Group Cohomology and (Multi)Symplectic Integrators: New Geometric Tools for Lie Group Machine Learning Based on Souriau Geometric Statistical Mechanics. *Entropy* **2020**, *22*, 498. [CrossRef]

18. HYPERBOLIC DEEP LEARNING: A Nascent and Promising Field. Available online: http://hyperbolicdeeplearning.com/ (accessed on 2 June 2020).

19. Fakhri, H. Su(1, 1)-Barut–Girardello coherent states for Landau levels. *J. Phys. A Math. Gen.* **2004**, *37*, 5203–5210. [CrossRef]

20. Gazeau, J.-P. Coherent States in Quantum Optics: An Oriented Overview. *arXiv preprint* **2018**, arXiv:1810.06473.

21. Citti, G.; Sarti, A. A Cortical Based Model of Perceptual Completion in the Roto-Translation Space. *J. Math. Imaging Vis.* **2006**, *24*, 307–326. [CrossRef]

22. Miolane, N. Geometric Statistics for Computational Anatomy. Ph.D. Thesis, INRIA Sophia Antipolis, Valbonne, France, Stanford University, Stanford, CA, USA, 2016.

23. Verblunsky, S. On Positive Harmonic Functions: A Contribution to the Algebra of Fourier Series. *Proc. Lond. Math. Soc.* **1935**, *38*, 125–157. [CrossRef]

24. Verblunsky, S. On positive harmonic functions (second paper). *Proc. Lond. Math. Soc.* **1936**, *40*, 290–320. [CrossRef]

25. Trench, W.F. An Algorithm for the Inversion of Finite Toeplitz Matrices. *J. Soc. Ind. Appl. Math.* **1964**, *12*, 515–522. [CrossRef]

26. Barfoot, T.D.; Furgale, P. Associating Uncertainty with Three-Dimensional Poses for Use in Estimation Problems. *IEEE Trans. Robot.* **2014**, *30*, 679–693. [CrossRef]

27. Barrau, A.; Bonnabel, S. Intrinsic Filtering on Lie Groups With Applications to Attitude Estimation. *IEEE Trans. Autom. Control* **2014**, *60*, 436–449. [CrossRef]

28. Chirikjian, G.S. *Stochastic Models, Information Theory, and Lie Groups, Volume 1: Classical Results and Geometric Methods. Applied and Numerical Harmonic Analysis*; Birkhäuser: Basel, Switzerland, 2009. [CrossRef]

29. Gregory, S.C. *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2011.

30. Gromov, M. In a Search for a Structure, Part 1: On Entropy. In Proceedings of the European Congress of Mathematics, Kraków, Poland, 2–7 July 2012; pp. 51–78.

31. Gromov, M. Six Lectures on Probabiliy, Symmetry, Linearity. *Jussieu* **2014**, preprint.

32. Arnaudon, A.; De Castro, A.L.; Holm, D. Noise and Dissipation on Coadjoint Orbits. *J. Nonlinear Sci.* **2017**, *28*, 91–145. [CrossRef]

33. Benenti, S.; Tulczyjew, W.M. Cocycles of the coadjoint representation of a Lie group interpreted as differential forms. *Mem. Accad. Sci. Torino* **1986**, *10*, 117–138.

34. Benenti, S.; Tulczyjew, W.M. A geometrical interpretation of the 1-cocycles of a Lie group. In *Geometrodynamics*; Prastaro, A., Ed.; World Scientific Publishing Co.: Singapore, 1985; pp. 3–24.

35. Nencka, H.; Streater, R.F. INFORMATION GEOMETRY FOR SOME LIE ALGEBRAS. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **1999**, *2*, 441–460. [CrossRef]

36. Souriau, J.-M. Mécanique statistique, groupes de Lie et cosmologie. In *Géométrie symplectique et physique mathématique*; Éditions du C.N.R.S: Aix-en-Provence, France, 1974; pp. 59–113.

37. Davis, M.S. (Under the Direction of Francois Ziegler); *Homogeneous Symplectic Manifolds of the Galilei Group*; Georgia Southern University: Statesboro, GA, USA, 2012.

38. Marle, C.M. *Géométrie Symplectique et Géométrie de Poisson*; Calvage & Mounet: Paris, France, 2018.

39. Vandebogert, K. *Notes on Symplectic Geometry*; University of South Carolina: Columbia, SC, USA, 3 September 2017.

40. Cartier, P. Some Fundamental Techniques in the Theory of Integrable Systems, IHES/M/94/23, SW9421. 1994. Available online: https://cds.cern.ch/record/263222/files/P00023319.pdf (accessed on 31 May 2020).

41. Bismut, J.-M.; Labourie, F. Symplectic geometry and the Verlinde Formulas. *Surv. Differ. Geom.* **1999**, *5*, 97–311. [CrossRef]

42. Koszul, J.-L.; Zou, Y.M. *Introduction to Symplectic Geometry*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2019.

43. Knapp, A.W. *Representation Theory of Semisimple Groups*; Walter de Gruyter GmbH: Berlin, Germany, 1986.

44. Frenkel, I.B. Orbital theory for affine Lie algebras. *Inven. Math.* **1984**, *77*, 301–352. [CrossRef]

45. Libine, M. Introduction to Representations of Real Semisimple Lie Groups. *arXiv* **2014**, arXiv:1212.2578v2.

46. Clerc, J.-L.; Ørsted, B. The Maslov index revisited. *Transform. Groups* **2001**, *6*, 303–320. [CrossRef]

47. Foth, P.; Lamb, M. The Poisson geometry of SU(1,1). *J. Math. Phys.* **2010**, *51*, 092701. [CrossRef]

48. Perelomov, A.M. Coherent states for arbitrary Lie group. *Commun. Math. Phys.* **1972**, *26*, 222–236. [CrossRef]

49. Hashimoto, T.; Ogura, K.; Okamoto, K.; Sawae, R.; Yasunaga, Y. Kirillov-Konstant theory and Feynman path integrals on coadjoint orbits I. *Hokkaido Math. J.* **1991**, *20*, 353–405. [CrossRef]

50. Hashimoto, T.; Ogura, K.; Okamoto, K.; Sawae, R. Kirillov Konstant theory and Feynman path integrals on coadjoint orbits of SU(2) and SU(1, 1). *Int. J. Mod Phys.* **1992**, *A7* (Suppl. 1A), 377–390. [CrossRef]

51. Hashimoto, T.; Ogura, K.; Okamoto, K.; Sawae, R. Borel-Weil theory and Feynman path integrals on flag manifolds. *Hiroshima Math. J.* **1993**, *23*, 231–247. [CrossRef]

52. Hashimoto, T. KirillovKonstant theory and Feynman path integrals on coadjoint orbits of a certain real semisimple Lie group. *Hiroshima Math. J.* **1993**, *23*, 607–627. [CrossRef]

53. Cishahayo, C.; De Bièvre, S. On the contraction of the discrete series of SU(1,1). *Ann. Inst. Fourier* **1993**, *43*, 551–567. [CrossRef]

54. Cahen, B. Contraction de SU(1,1) vers le Groupe de Heisenberg. *Travaux Mathématiques* **2004**, *XV*, 19–43.

55. Cahen, M.; Gutt, S.; Rawnsley, J. Quantization on Kähler manifolds I, Geometric interpretation of Berezin quantization. *J. Geom. Phys.* **1990**, *7*, 45–62.

56. Marle, C.-M. Projection Stéréographique et Moments, Hal-02157930, Version 1. June 2019. Available online: https://hal.archives-ouvertes.fr/hal-02157930/ (accessed on 31 May 2020).

57. Guichardet, A. La methode des orbites: Historiques, principes, résultats. In *Leçons de Mathématiques d'Aujourd'hui*; Cassini: Paris, France, 2010; Volume 4, pp. 33–59.

58. Vergne, M. Representations of Lie Groups and the Orbit Method. In *Emmy Noether in Bryn Mawr*; Springer: Berlin/Heidelberg, Germany, 1983; pp. 59–101. [CrossRef]

59. Duflo, M.; Heckman, G.; Vergne, M. Projection d'orbites, formule de Kirillov et formule de Blattner. *Mémoires Société Mathématique Fr.* **1984**, *1*, 65–128. [CrossRef]

60. Pukanszky, L. The Plancherel formula for the universal covering group of SL(R, 2). *Math. Ann.* **1964**, *156*, 96–143. [CrossRef]

61. Pukanszky, L. *Leçons sur les Représentations des Groupes, Monographies de la Société Mathématique de France*; Dunod: Paris, France, 1967.

62. Bernat, P. *Représentations des Groupes de Lie, Monographie de la Société Mathématique de France*; Dunod: Paris, France, 1972.

63. Dixmier, J. *Enveloping Algebras*; American Mathematical Society: Providence, RI, USA, 1996.

64. Duflo, M. Construction De Representations Unitaires D'un Groupe De Lie. In *Harmonic Analysis and Group Representation*; Talamanca, A.F., Ed.; C.I.M.E. Summer Schools 1980; Springer: Berlin/Heidelberg, Germany, 2010; Volume 82.

65. Guichardet, A. Théorie de Mackey et méthode des orbites selon M. Duflo. *Expo. Math.* **1985**, *3*, 303–346.

66. Mnemné, R.; Testard, F. *Groupes de Lie Classiques*; Hermann: Paris, France, 1985.

67. Rais, M. *Orbites Coadjointes et Représentations des Groupes, Cours*; C.I.M.P.A.: Blagnac/Toulouse, France, 1980.

68. Rais, M. La représentation coadjointe du groupe affine. *Ann. Inst. Fourier* **1978**, *28*, 207–237. [CrossRef]

69. De Micheli, E. On the Connection between Spherical Laplace Transform and Non-Euclidean Fourier Analysis. *Mathematics* **2020**, *8*, 287. [CrossRef]

70. Marsden, J.E.; Misiolek, G.; Ortega, J.-P.; Perlmutter, M.; Ratiu, T.S. *Hamiltonian Reduction by Stages*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2007.

71. Kurnyavko, O.L.I.; Shirokov, V. Algebraic method for construction of infinitesimal invariants of Lie groups representations. *arXiv* **2017**, arXiv:1710.07977.

72. Kurnyavko, O.L.; Shirokov, I.V. Construction of invariants of the coadjoint representation of Lie groups using linear algebra methods. *Theor. Math. Phys.* **2016**, *188*, 965–979. [CrossRef]

73. Shirokov, I.V. Differential invariants of the transformation group of a homogeneous space. *Sib. Math. J.* **2007**, *48*, 1127–1140. [CrossRef]

74. Goncharovskii, M.M.; Shirokov, I.V. Differential invariants and operators of invariant differentiation of the projectable action of Lie groups. *Theor. Math. Phys.* **2015**, *183*, 619–636. [CrossRef]

75. Shirokov, I.V. Darboux coordinates onK-orbits and the spectra of Casimir operators on lie groups. *Theor. Math. Phys.* **2000**, *123*, 754–767. [CrossRef]

76. Holm, D. Variational principles for stochastic fluid dynamics. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **2015**, *471*, 20140963. [CrossRef] [PubMed]

77. Gay-Balmaz, F.; Holm, D. Selective decay by Casimir dissipation in inviscid fluids. *Nonlinearity* **2013**, *26*, 495–524. [CrossRef]

78. Gay-Balmaz, F.; Holm, D. A geometric theory of selective decay with applications in MHD. *Nonlinearity* **2014**, *27*, 1747–1777. [CrossRef]

79. Casimir, H.G.B. Uber die konstruktion einer zu den irreduziblen darstellungen halbeinfacher kontinuierlicher gruppen gehörigen differentialgleichung. *Proc. R. Soc. Amst.* **1931**, *34*, 844–846.

80. Racah, G. Sulla caratterizzazione delle rappresentazioni irriducibili dei gruppi semisemplici di Lie. *Atti Accad. Naz. Lincei. Rend. Cl. Sci. Fis. Mat. Nat.* **1950**, *8*, 108–112.

81. Oller, P.J. Oller Moving frames and differential invariants in centro-affine geometry. *Lobachevskii J. Math.* **2010**, *31*, 77–89. [CrossRef]

82. Berezin, D.V. Invariants of the co-adjoint representation for Lie algebras of a special form. *Russ. Math. Surv.* **1996**, *51*, 137–139. [CrossRef]

83. Abellanas, L. A general setting for Casimir invariants. *J. Math. Phys.* **1975**, *16*, 1580. [CrossRef]

84. Beltrametti, E.; Blasi, A. On the number of Casimir operators associated with any lie group. *Phys. Lett.* **1966**, *20*, 62–64. [CrossRef]

85. Pecina-Cruz, J.N. An algorithm to calculate the invariants of any Lie algebra. *J. Math. Phys.* **1994**, *35*, 3146–3162. [CrossRef]

86. Dixmier, J. *Algèbres enveloppantes (Cahiers Scientifiques. No. 37)*; Gauthier-Villars: Paris, France, 1974.

87. Mikheyev, V.V.; Shirokov, I.V. Application of coadjoint orbits in the thermodynamics of non-compact manifolds. *Electron. J. Theor. Phys.* **2005**, *2*, 1–10.

88. Mikheyev, V.; Nielsen, F.; Barbaresco, F. Method of Orbits of Co-Associated Representation in Thermodynamics of the Lie Non-compact Groups. In *Applications of Evolutionary Computation*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2017; Volume 10589, pp. 425–431.

89. Fomenko, A.T.; Trofimov, V.V. *Integrable Systems on Lie Algebras and Symmetric Spaces*; Gordon and Breach Science Publishers: New York, NY, USA, 1988.

90. Trofimov, V.V. *Introduction to Geometry of Manifolds with Symmetry*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 1994.

91. Machon, T. The Godbillon–Vey invariant as a restricted Casimir of three-dimensional ideal fluids. *J. Phys. A Math. Theor.* **2020**, *53*, 235701. [CrossRef]

92. Casimir, H.B.G. On Onsager's Principle of Microscopic Reversibility. *Rev. Mod. Phys.* **1945**, *17*, 343. [CrossRef]

93. Thiffeault, J.-L.; Morrison, P. Classification and Casimir invariants of Lie–Poisson brackets. *Phys. D Nonlinear Phenom.* **2000**, *136*, 205–244. [CrossRef]

94. Arnaudon, A.; De Castro, A.L.; Holm, D.D. Noise and dissipation in rigid body motion. *arXiv* **2016**, arXiv:1606.06308v1.

95. Bargmann, V. Irreducible Unitary Representations of the Lorentz Group. *Ann. Math.* **1947**, *48*, 568. [CrossRef]

96. Souriau, J.-M. Mécanique Statistique, Groupes de Lie et Cosmologie. 2020. Available online: https://www.academia.edu/42630654/Statistical_Mechanics_Lie_Group_and_Cosmology_1_st_part_Symplectic_Model_of_Statistical_Mechanics (accessed on 20 April 2020).

97. Souriau, J.-M. *Structure des Systèmes Dynamiques*; Dunod: Paris, France, 1969.

98. Souriau, J.M. *Mécanique Classique et Géométrie Symplectique, Rapport CNRS CPT-84/PE.1695*; Université de Provence et Centre de Physique Théorique CNRS: Marseille, France, 1984.

99. Souriau, J.M. Equations Canoniques et Géométrie Symplectique. *Pub. Sci. Univ. Alger. Sér. A* **1954**, *1*, 239–265.

100.   Souriau, J.M. *Géométrie de l'Espace des Phases, Calcul des Variations et Mécanique Quantique, Tirage Ronéotypé*; Faculté des Sciences: Marseille, France, 1965.

101.   Souriau, J.-M. Realisations d'algèbres de Lie au moyen de variables dynamiques. *Il Nuovo Cim. A* **1967**, *49*, 197–198. [CrossRef]

102.   Marle, C.-M. From Tools in Symplectic and Poisson Geometry to J.-M. Souriau's Theories of Statistical Mechanics and Thermodynamics. *Entropy* **2016**, *18*, 370. [CrossRef]

103.   Barbaresco, F. Higher Order Geometric Theory of Information and Heat Based on Poly-Symplectic Geometry of Souriau Lie Groups Thermodynamics and Their Contextures: The Bedrock for Lie Group Machine Learning. *Entropy* **2018**, *20*, 840. [CrossRef]

104.   Barbaresco, F. *Souriau Exponential Map Algorithm for Machine Learning on Matrix Lie Groups*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2019; pp. 85–95.

105.   Barbaresco, F. Geometric Theory of Heat from Souriau Lie Groups Thermodynamics and Koszul Hessian Geometry: Applications in Information Geometry for Exponential Families. *Entropy* **2016**, *18*, 386. [CrossRef]

106.   Barbaresco, F. Lie Group Machine Learning and Gibbs Density on Poincaré Unit Disk from Souriau Lie Groups Thermodynamics and SU(1,1) Coadjoint Orbits. In *Applications of Evolutionary Computation*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2019; Volume 11712, pp. 157–170.

107.   Barbaresco, F. Application exponentielle de matrice par l'extension de l'algorithme de Jean-Marie Souriau, utilisable pour le tir géodésique et l'apprentissage machine pour les groupes de Lie. In Proceedings of the Colloque GRETSI 2019, Lille, France, 26–29 August 2019.

108.   Ishi, H.; Kolodziejek, B. Characterization of the Riesz Exponential Familly on Homogeneous Cones. *arXiv* **2018**, arXiv:1605.03896.

109.   Tojo, K.; Yoshino, T. On a Method to Construct Exponential Families by Representation Theory. *arXiv* **2018**, arXiv:1811.01394.

110.   Tojo, K.; Yoshino, T. On a Method to Construct Exponential Families by Representation Theory. In *Applications of Evolutionary Computation*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2019; pp. 147–156.

111.   Tojo, K.; Yoshino, T. Harmonic exponential families on homogeneous spaces. preprint. 2020.

112.   Arnold, V. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier* **1966**, *16*, 319–361. [CrossRef]

113.   Arnold, V.I.; Givental, A.B. Symplectic Geometry. In *Dynamical Systems IV: Symplectic Geometry and Its Applications*; Encyclopaedia of Mathematical Sciences; Arnol'd, V.I., Novikov, S.P., Eds.; Springer: Berlin, Gemany, 1990; Volume 4, pp. 1–136.

114.   Balian, R.; Alhassid, Y.; Reinhardt, H. Dissipation in many-body systems: A geometric approach based on information theory. *Phys. Rep.* **1986**, *131*, 1–146. [CrossRef]

115.   Balian, R.; Balazs, N. Equiprobability, inference, and entropy in quantum theory. *Ann. Phys.* **1987**, *179*, 97–144. [CrossRef]

116.   Balian, R. On the principles of quantum mechanics and the reduction of the wave packet. *Am. J. Phys.* **1989**, *57*, 1019. [CrossRef]

117.   Balian, R. *From Microphysics to Macrophysics*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 1991; Volume 1–2.

118.   Balian, R. Incomplete descriptions and relevant entropies. *Am. J. Phys.* **1999**, *67*, 1078–1090. [CrossRef]

119.   Balian, R.; Valentin, P. Hamiltonian structure of thermodynamics with gauge. *Eur. Phys. J. B* **2001**, *21*, 269–282. [CrossRef]

120.   Balian, R. Entropy, a Protean Concept. In *Poincaré Seminar 2003*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2004; pp. 119–144.

121.   Balian, R. Information in statistical physics. *Stud. Hist. Philos. Sci. Part B Stud. Hist. Philos. Mod. Phys.* **2005**, *36*, 323–353. [CrossRef]

122.   Balian, R. The Entropy-Based Quantum Metric. *Entropy* **2014**, *16*, 3878–3888. [CrossRef]

123.   Balian, R. *François Massieu et les Potentiels Thermodynamiques, Évolution des Disciplines et Histoire des Découvertes*; Académie des Sciences: Paris, France, 2015.

124.   Berthet, Q.; Blondel, M.; Teboul, O.; Cuturi, M.; Vert, J.-P.; Bach, F. Learning with Differentiable Perturbed Optimizers. *arXiv preprint* **2002**, arXiv:2002.08676.

125. Blondel, M.; Martins, A.F.T.; Niculae, V. Learning with Fenchel-Young Losses. *J. Mach. Learn. Res.* **2020**, *21*, 1–69.

126. Wainwright, M.J.; Jordan, M.I. Graphical Models, Exponential Families, and Variational Inference. *Found. Trends Mach. Learn.* **2007**, *1*, 1–305. [CrossRef]

127. Yahyai, M. Représentations Étoile du Revêtement Universel du Groupe Hyperbolique et Formule de Plancherel. Ph.D. Thesis, Université de Metz, Metz, France, 23 June 1995.

128. Bertrand, J.; Irac-Astaud, M. Characterization of su(1,1) coherent states in terms of affine wavelets. *J. Phys. A Math. Gen.* **2002**, *35*, 7347–7357. [CrossRef]