



entropy



Article

Thermodynamics à la Souriau on Kähler Non-Compact Symmetric Spaces for Cartan Neural Networks

Pietro G. Fré, Alexander S. Sorin and Mario Trigiante

Topic Collection

Feature Papers in Information Theory

Edited by

Prof. Dr. Raúl Alcaraz, Prof. Dr. Luca Faes, Prof. Dr. Leandro Pardo and Prof. Dr. Boris Ryabko



<https://doi.org/10.3390/e28040365>

Article

Thermodynamics à la Souriau on Kähler Non-Compact Symmetric Spaces for Cartan Neural Networks

Pietro G. Fré ^{1,2,*}, Alexander S. Sorin ³  and Mario Trigiante ^{4,5} 

¹ Dipartimento di Fisica, Università di Torino, Via P. Giuria 1, I-10125 Torino, Italy

² Additati & Partners Consulting s.r.l., Via Filippo Pacini 36, I-51100 Pistoia, Italy

³ Center for Quantum Science and Technology, Tel-Aviv University, Tel Aviv 69978, Israel; asorin@tauex.tau.ac.il

⁴ Dipartimento DISAT, Politecnico di Torino, Corso Duca degli Abruzzi 24, I-10129 Torino, Italy; mario.trigiante@polito.it

⁵ Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Torino, I-10125 Torino, Italy

* Correspondence: pietro.fre@unito.it

Abstract

In this paper, we clarify several issues concerning the abstract geometrical formulation of thermodynamics on non-compact symmetric spaces U/H that are the mathematical model of hidden layers in the new paradigm of Cartan Neural Networks. We introduce a clear-cut distinction between the generalized thermodynamics associated with Integrable Dynamical Systems and the challenging proposal of Gibbs probability distributions on U/H provided by generalized thermodynamics à la Souriau. Our main result is the proof that U/H s supporting such Gibbs distributions are only the Kähler ones. Furthermore, for the latter, we solve the problem of determining the space of temperatures, namely, of Lie algebra elements for which the partition function converges. The space of generalized temperatures is the orbit under the adjoint action of U of a positivity domain in the Cartan subalgebra $C_c \subset \mathbb{H}$ of the maximal compact subalgebra $\mathbb{H} \subset U$. We illustrate how our explicit constructions for the Poincaré and Siegel planes might be extended to the whole class of Calabi–Vesentini manifolds utilizing Paint Group symmetry. Furthermore, we claim that Rao’s, Chentsov’s, and Amari’s Information Geometry and the thermodynamical geometry of Ruppeiner and Lychagin are the very same thing. In particular, we provide an explicit study of thermodynamical geometry for the Poincaré plane. The key feature of the Gibbs probability distributions in this setup is their covariance under the entire group of symmetries U . The partition function is invariant against U transformations, and the set of its arguments, namely the generalized temperatures, can always be reduced to a minimal set whose cardinality is equal to the rank of the compact denominator group $H \subset U$.

Keywords: generalized thermodynamics; Lie groups; symplectic geometry; Cartan neural networks; contact geometry; Shannon entropy; Cartan Hadamard manifolds; non compact symmetric spaces; partition function; Siegel plane



Academic Editor: Boris Ryabko

Received: 18 December 2025

Revised: 4 March 2026

Accepted: 10 March 2026

Published: 24 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The purpose of the present paper is to clarify all the proper relations, identifications, and, when necessary, clear-cut distinctions among several mathematical constructions that have been recently introduced by different researchers into the mathematical formulation of Machine Learning and that admit, as their own conceptual pivot, the notions of Lie group, Hamiltonian dynamical system and the geometrical rephrasing of thermodynamical equilibrium states. This is particularly relevant in view of the new paradigm of Cartan

Neural Networks [1–4], intrinsically characterized by the identification of the network hidden layers with as many non-compact symmetric spaces U_i/H_i , each metrically equivalent to an appropriate solvable Lie group. The generalized notion of Gibbs state provides the proper way to introduce Gaussian-like probability distributions on non-compact symmetric spaces that, in Cartan Neural Network architectures, constitute the hidden layers. Hence, we start with a short summary of the new CaNN paradigm, highlighting its geometrical and group-theoretical strategic aspects. After that, we introduce the other mathematical ingredients of our discussion and clarification plan.

Let us remark since the very beginning that the generalized Gibbs probability distributions, whose structure, properties and appropriate construction principles are the main target of the present investigation, have been introduced into Machine Learning about a decade ago as a preferred method of Machine Learning analysis of electromagnetic signals such as those involved in radar technologies [5] (see also the review [6]). More generally, these probability distributions, which are covariant with respect to the full group of isometries of the manifold on which they are defined and fit into geometrical thermodynamics (alias information geometry), are promising tools for Machine Learning architectures concerned with all kinds of electromagnetic signals and also all kinds of temporal sequences.

1.1. Cartan Neural Networks: A New Paradigm

In [1], whose authors' list includes two of us, a new mathematical paradigm was introduced for the engineering of neural network architectures under the name of **PGTS** (PGTS is an acronym for Paint Group Tits–Satake) **theory of non-compact symmetric spaces**. The essential points of this paradigm are:

1. The systematic substitution of the Euclidean $\mathbb{R}^n$ space with a non-compact symmetric coset manifold U/H , where U is a simple non-compact Lie group, and H is the maximal compact subgroup of U . All these manifolds are Cartan–Hadamard manifolds and are metrically equivalent to a specific **solvable Lie group manifold** $S_{U/H}$.
2. The grouping of these manifolds into Tits Satake universality classes, which provides an ideal mathematical definition of neural layers.
3. The systematic *suppression of point-wise activation functions* like the sigmoid and its close relatives, the necessary non-linearity being universally provided by *generalized exponential maps* from Lie Algebras to the corresponding Lie Groups and the *generalized logarithm maps* that are the inverse of the former.

In a twin pair of papers [2,3], it was shown how a generic multi-layer neural network can be cast into a form that implements points (1), (2), and (3) of the above paradigm. This class of neural network architectures based on the above principles was named **Cartan neural networks (CaNNs)** in honor of the monumental achievement of Èlie Cartan, who obtained the complete classification of all symmetric spaces and their one-to-one correspondence with the classification of real forms of simple complex Lie algebras [7–10]. In [2,3], the general scheme of *supervised learning* for a *classification task* within **CaNNs** was addressed. In [4], many additional mathematical general conceptions and constructions were included in the toolbox for the engineering of **CaNNs**, encompassing, in particular, the general theory of codimension one separators, harmonic analysis on non-compact U/H , Tits Satake fibre bundles and tessellation groups. We refer the reader to [4] for further details on such topics.

Covariance of CaNNs

In [2], an algorithm was described in which each *datum* is linearly mapped (with a matrix W_0 , target of learning) to the solvable coordinate vector $\mathbf{Y}$ labeling a point in a non-compact symmetric space U_1/H_1 . The latter is the first layer in a sequence of similar layers U_i/H_i , each a non-compact symmetric space with (a priori different) dimension d_i . As discussed at length in [2], the general scheme also allows the non-compact rank and the type of the various U_i/H_i to be different; yet, thanks to the fundamental property of metric equivalence with a suitable solvable group $\mathcal{S}_i$, we can constrain the map from one layer to the next one to be a group homomorphism derived from a linear homomorphism of the corresponding solvable Lie algebras. More specifically, denoting by K_i the map from the i th-layer, described by the space U_i/H_i , to the subsequent one U_{i+1}/H_{i+1} , this map is described by a group homomorphism between $\mathcal{S}_i$ and $\mathcal{S}_{i+1}$ while the corresponding *push forward* map K_{i*} is a linear homomorphism between the solvable Lie algebras $Solv_i$, $Solv_{i+1}$, generating $\mathcal{S}_i$ and $\mathcal{S}_{i+1}$, respectively:

$$\forall X, Y \in Solv_i : K_{i*}[X, Y] = [K_{i*}X, K_{i*}Y] \in Solv_{i+1}.$$

If $d_i \leq d_{i+1}$, K_i can be characterized as an *isometric inclusion* [11] having the defining property that, if g_i and g_{i+1} denote the Riemannian metrics on U_i/H_i and U_{i+1}/H_{i+1} , respectively,

$$\forall X, Y \in Solv_i : g_i(X, Y) = g_{i+1}(K_{i*}X, K_{i*}Y).$$

As shown in [11], the mapping K_{i*} between the tangent spaces at corresponding points on the two manifolds, separately isomorphic to $Solv_i$ and $Solv_{i+1}$, is injective. If, on the other hand, $d_i > d_{i+1}$, as a mapping between a linear space $Solv_i$ and a lower-dimensional one $Solv_{i+1}$, K_{i*} has a non-trivial kernel. If we define the metric $g_i^{(0)}$ on $\mathcal{S}_i$ by the property:

$$\forall X, Y \in Solv_i : g_i^{(0)}(X, Y) = g_{i+1}(K_{i*}X, K_{i*}Y),$$

$g_i^{(0)}$ is singular and thus does not coincide with g_i . Nevertheless, being K_{i*} a homomorphism between Lie algebras, $\text{Ker}(K_{i*})$, of dimension $d_i - d_{i+1}$, is an ideal of $Solv_i$ consisting of the *zero-norm* vectors with respect to $g_i^{(0)}$, orthogonal to all the other vectors with respect to the same singular metric. Restricted to the solvable Lie algebra $Solv'_i \equiv Solv_i \ominus \text{Ker}(K_{i*})$, $g_i^{(0)}$ coincides with g_i and, therefore, when $d_i > d_{i+1}$, K_i can be characterized as an *isometry* between $\mathcal{S}'_i \equiv \exp(Solv'_i)$, with metric g_i restricted to $Solv'_i \times Solv'_i$, and $\mathcal{S}_{i+1}$.

This general characterization of $K_i : U_i/H_i \rightarrow U_{i+1}/H_{i+1}$ as an isometric mapping implies its general covariance with respect to the transformations of both the U_i and the U_{i+1} group. The action of the two groups on K_i can be formally described as follows:

$$K_i \rightarrow U_{i+1} \circ K_i \circ U_i. \quad (1)$$

The Relevance of Covariance

Covariance, as expressed in Equation (1), is the fundamental conceptual and architectural advance provided by CaNN.s that, as explained in [2], are not just one proposal among others, rather they constitute the unique available scheme, allowed by basic theorems of differential geometry, able to **dispose off** the *point-wise activation functions* (contradictory with any sort of covariance), to maintain the existence of a unique distance function on each network layer and to preserve indispensable non-linearity.

1.2. The Mathematical Basis of CaNN

Having introduced the new paradigm, we summarize the mathematical key items that constitute its foundation.

The strategic metric equivalence with solvable groups

As discussed at length in the foundational paper [1] and already recalled above, the strategic element that allows the construction of CaNN.s, with all the properties mentioned above, is the *metric equivalence* of all non-compact symmetric spaces U/H with a suitable solvable Lie subgroup $S_{U/H} \subset U$ which is a generalization, in each different U -case, of the Borel subgroup, applying to the case $U = \mathrm{SL}(N, \mathbb{R})$. This metric equivalence amounts to the statement that all non-compact symmetric spaces are *Alekseevskian Normal Riemannian Manifolds*.

Alekseevsky Normal Manifolds and Solvable Lie Groups

Following the original viewpoint of Alekseevsky [12,13], we say that a Riemannian manifold $(\mathcal{M}, g)$ is *normal* if it admits a completely solvable Lie group $\mathcal{S}_{\mathcal{M}} \simeq \exp[\mathrm{Solv}_{\mathcal{M}}]$ of isometries that acts on the manifold in a simply transitive manner (i.e., for every 2 points in the manifold there is one and only one group element connecting them). The group $\mathcal{S}_{\mathcal{M}}$ is then generated by a so-called *normal metric Lie algebra*, that is, a completely solvable Lie algebra $\mathrm{Solv}_{\mathcal{M}}$ endowed with an Euclidean, positive definite, *Solv*-invariant, symmetric form $<, >$. The main tool to classify and study the normal homogeneous spaces is provided by the theorem [8,14] that states that if a Riemannian manifold $(\mathcal{M}, g)$ admits a transitive normal solvable group of isometries $\exp[\mathrm{Solv}_{\mathcal{M}}]$, then it is metrically equivalent to this solvable group manifold

$$\begin{aligned} \mathcal{M} &\simeq \exp[\mathrm{Solv}_{\mathcal{M}}], \\ g|_{e \in \mathcal{M}} &= <, > \end{aligned} \quad (2)$$

where $<, >$ is the Euclidean metric defined on the normal solvable Lie algebra $\mathrm{Solv}_{\mathcal{M}}$.

The original conjecture of Alekseevsky was just restricted to quaternionic Kähler manifolds and stated that any such manifold $\mathcal{M}$ that was also homogeneous and of negative Ricci curvature should be normal, in the over-mentioned sense, namely a transitive solvable group of isometries $\exp[\mathrm{Solv}_{\mathcal{M}}]$ should exist, that could be identified with the manifold itself. Note that the actual group of isometries U of $\mathcal{M}$ could be much larger than the solvable group,

$$U \supset \exp[\mathrm{Solv}_{\mathcal{M}}], \quad (3)$$

as it is, for instance, the case for all symmetric spaces $\mathcal{M} = \frac{U}{H}$; yet, the solvable normed Lie algebra $(\mathrm{Solv}_{\mathcal{M}}, <, >)$ had to exist. The problem of classifying the considered manifolds was turned in this way into the problem of classifying the *normal metric solvable Lie algebras* $(\mathrm{Solv}, <, >)$. Note that in Alekseevsky's case, the symmetric form $<, >$ was not only required to be positive definite but also quaternionic Kähler. Alekseevsky's conjecture actually applies to much more general homogeneous Riemannian manifolds than the quaternionic ones: For instance, it applies to all those endowed with a special Kähler geometry or with a real special one, as the classification of de Wit et al. [15–17] demonstrated. It also applies to the symmetric spaces appearing in the scalar sector of extended supergravities with more than eight supercharges. For all these manifolds, there exists the corresponding normal metric algebra $(\mathrm{Solv}, <, >)$; in other words, they are *normal*. Actually, by explicit construction, as discussed in [1], all non-compact symmetric spaces U/H with U a simple non-compact Lie group and $H \subset U$ its maximal compact subgroup are normal Alekseevskian manifolds.

1.3. The Link with Symplectic Geometry and Generalized Thermodynamics

Analyzing the mathematical foundations of CaNNs, we are naturally led to observe a natural link with symplectic geometry and generalized thermodynamics. In this subsection, we unveil, also historically, such a conceptual path.

Integrability of geodesic equations, Poisson, and symplectic manifolds

The mathematical theory that links *non-compact symmetric manifolds* to *Normed Solvable Lie Algebras* was pioneered by mathematicians but then it was extensively developed within the context of supergravity, since all scalar manifolds of extended supergravity Lagrangians happen to be non-compact symmetric spaces and their solvable representations played a decisive role in the systematic resolution of several problems, in particular the construction of cosmic billiards and extremal black-hole solutions. It was in this framework that the integrability of geodesic equations for such manifolds was analyzed by Fré and Sorin in terms of a Lax pair equation [18] in 2006, and also a Poisson manifold viewpoint underlying such integrability was introduced by the same authors in 2009 [19]. The complete integrability and the explicit integration of geodesic equations in U/H symmetric spaces is an essential brick in the construction of Cartan Neural Network architectures and it is discussed at length in the foundational paper [1], where it is shown that the explicit result for the solvable coordinates, as functions $Y(t)$ of the affine parameter t , can be obtained directly in terms of the initial data, namely the starting point and the so named matrix of conserved Noether charges Q , bypassing the solution of Lax equation. It might seem from this that the Poisson manifold viewpoint [19] is interesting yet unnecessary in the context of Machine Learning, but such a conclusion is too hasty and incorrect for the following reason. The virtue of the *Poisson/symplectic approach* to the geodesic problem is that it puts it into the perspective of dynamical systems and at the same time of geometric thermodynamics, creating a triple link among the geodesics on U/H, the symplectic/contact geometry of thermodynamics, and the contact structures of fluid-dynamics [20,21]. As we are going to see, a recent research line in Machine Learning introduces, in a different setup and under the name of *Gibbs States for Lie Groups*, basic structures of geometric thermodynamics, so that a conceptual clarification of all the relations is quite appropriate and useful in order to combine Cartan Neural Networks with statistical conceptions as those advocated in the mentioned research line.

Generalized Geometric Thermodynamics

A process of fundamental importance in Chemistry is the separation of different substances that are present in gas mixtures and multi-component liquids. From a conceptual point of view, any separation method is based on the thermodynamics of mixtures of different components coexisting in different phases. Since the 19th century, this phenomenon has been carefully conceptualized by great chemists, physicists, and mathematicians, and its study has become a focal point of statistical mechanics and classical thermodynamics. Gibbs' rule of phases and the use, at the level of statistical mechanics, of the grand canonical ensemble (see Appendix C.3), with the introduction of the chemical potential are two fundamental junctures in this affair. However, the discouraging and critical aspect in this area of knowledge is that the exact calculation of canonical and grand canonical partition functions is of extreme difficulty when the particles forming the chemicals of our interest interact with each other, i.e., always, in the case of real and non-ideal substances. The cases of exact computation of the partition function are isolated and rare, reducing essentially to those of the classical ideal gas, the quantum free gases of bosons or fermions, and the two-dimensional Ising model for ferromagnetism. In all other cases, there is a plethora of approximation methods and sophisticated perturbative or approximate computational techniques. The object of primary interest for thermodynamic calculations is the

equation of state, i.e., the relation between both extensional (such as volume V , entropy S , and internal energy U) and intensive (such as temperature T and chemical potentials μ_i) thermodynamic quantities that is valid in equilibrium states (see Appendix C for a summary of classical thermodynamics). Equations of state can be derived exactly from the partition function if one knows the latter. Thus, alternating with attempts at direct calculations of certain partition functions, there has been, over the last century, a great deal of modeling activity, both theoretical and experimental, aimed at constructing mathematical formulations of equations of state in want of the missing partition function. However, such equations of state are just phenomenological models, and a deeper understanding of their rationale is necessary. Thanks to the work, hitherto little known outside a small community of specialists, of an even smaller number of low-temperature physicists and mathematical physicists, there exists a surprisingly innovative geometric view of classical thermodynamics that provides a more intrinsic view of thermodynamic states and seems able, by classical geometric means to provide mesoscopic information about real gases and liquids, while also defining a conceptual frame of reference in which phenomenological equations of state can be evaluated and possibly modified in a more systematic and profound way, particularly taking into account the possible isometries of the Riemannian metric surprisingly associated with the space of thermodynamic variables.

The problem is quite general, also in systems of different nature; for example, in Big Data sets, if one arrives at a thermodynamical limit description, one can wonder about the advantages of a geometrical formulation of thermodynamics.

In view of applications to the research compound of Geometric Deep Learning, which implies the use of predetermined metrics, it is a stimulating perspective to compare the metric setup of Geometrical Thermodynamics and that of Information Geometry, in particular focusing on the symplectic structure that can migrate from Thermodynamics to Data Science.

The small group of low-temperature physicists and mathematical physicists to whom we owe the entire body of developments related to the geometric view of thermodynamics we have referred to consists of three *senior founders* the American *George Ruppeiner* affiliated with the New College of Florida in Sarasota and the two Russians *Valentin Lychagin* and *Mikhail Roop*, plus a small cohort of adherents consisting of their occasional collaborators, Ph.D. students, postdocs, and so on. It is very interesting to read Ruppeiner's autobiographical article [22] written in April 2016 for the commemorative volume in honor of Horst Meyer's 90th birthday, who unfortunately passed away a few months later. In this article, the author recounts how, in the years 1975–1980 when he was a Ph.D. student at Duke University conducting low-temperature gas experiments in Meyer's laboratory, devoting most of his efforts to perfecting himself in low-temperature experimental physics, he nonetheless had a drive to take General Relativity Courses and deepen his knowledge of differential geometry. A spark was ignited in his mind, he recounts, when he read in *Physics Today* an article by Frank Weinhold [23] in which a metric form was introduced in the context of thermodynamic variables, something hitherto considered peregrine and absurd. Ruppeiner, on the other hand, regarded it as a serious suggestion and deemed that a Riemannian view of thermodynamics could be constructed and could also be useful in the analysis of critical phenomena. Gradually, continuing on the path he had taken, he arrived at constructing two-dimensional metrics in the temperature-density plane that were consistent with the principles of thermodynamics and went so far as to calculate the curvature scalar R of such metrics. A salient moment in the development of his thought was when he arrived at the physical interpretation of R :

$$|R| \propto \zeta^3 \quad (4)$$

where ξ is the statistical correlation length, which, as everyone knows, tends to infinity in the vicinity of critical points and phase transitions. Of course, ideal gases correspond to flat metrics with zero curvature $R = 0$ and no critical points. Thus, thermodynamic curvature became a classical indicator of molecular interactions at the mesoscopic level, and in the first two decades of the 21st century, Ruppeiner contributed a series of very interesting papers on the use of Riemannian geometry in the study of thermodynamics and its critical phenomena: [22,24–28]. More recently, but following reflections developed over the years and expounded in his 2018 lectures at Wisla [29], Valentin Lychagin, a Russian mathematician for a long time professor at Tromsø University in Norway, identified and systematically expounded within the framework of information theory, an interesting connection between contact geometry and thermodynamics, characterizing possible thermodynamic equilibrium states as Legendrian subvarieties of contact varieties. Because of the complex and general relationships between contact varieties and symplectic varieties (see Appendix A for a summary of contact and symplectic geometry), thermodynamic states can also be interpreted as Lagrangian subvarieties of symplectic varieties, and the canonical symplectic form on them naturally connects to a Riemannian metric, which is the one hypothesized and studied by Ruppeiner.

1.4. Gibbs States and Lie Group Generalized Thermodynamics

In a series of papers of which we quote only a small selection [30–34], that is most informative about the main idea, a group of French authors, including Charles-Michel Marle, Frédéric Barbaresco, Yann Cabanes and Pierre-Yves Lagrave, relying on old ideas of late Jean-Marie Souriau, have introduced the notion of *Gibbs States of Mechanical Systems with Symmetries* and of *Lie Group Thermodynamics* which bears a close similarity with the geometrical formulation of thermodynamics as expounded in Lychagin’s lectures [29], yet is quite distinct from it. As we explain in the main body of the present article, the original distinctive idea of *Lie Group Thermodynamics* is the definition of a subspace $\Omega \subset \mathbb{G}$ of the Lie algebra of symplectic Killing vector fields $X \in \mathbb{G}$ that leave invariant a symplectic manifold $(\mathcal{M}, \omega)$ in the sense that the Lie derivative along them of the symplectic form vanishes

$$\mathcal{L}_X \omega = 0 \quad (5)$$

(compare Equation (5) with Definition A11 in Appendix A.7 of Liouville vector fields) such that the following integral (the partition function) is convergent:

$$\forall \beta \in \Omega \quad : \quad Z(\beta) \equiv \int_{\mathcal{M}} \exp[-\beta \cdot \mathfrak{P}(\Phi)] d\lambda(\Phi) < \infty \quad (6)$$

where Φ are the coordinates on the $2n$ -dimensional differentiable manifold $\mathcal{M}$;

$$d\lambda(\Phi) \equiv \underbrace{\omega \wedge \omega \wedge \cdots \wedge \omega}_{n\text{-times}} \quad (7)$$

is its Liouville integration measure; and $\mathfrak{P}(\Phi)$ is the **moment map** (see below for the discussion of such a concept):

$$\forall k \in \mathbb{G} \quad \mathfrak{P}(\Phi) \quad : \quad k \longrightarrow \mathfrak{P}_k(\Phi) \in \mathbb{C}^\infty(\mathcal{M}) \quad (8)$$

The subspace $\Omega \subset \mathbb{G}$ of the symmetry Lie algebra is named the space of **generalized temperatures**.

1.4.1. Symplectic Moment Map

On a d -dimensional Riemannian space $\mathcal{M}$, admitting an isometry Lie algebra $\mathbb{U}$, moment maps can be defined as linear mappings between $\mathbb{U}$ and smooth functions on $\mathcal{M}$, valued in the holonomy algebra $\text{Hol}(\mathcal{M})$:

$$\begin{aligned} X \in \mathbb{U} &\longrightarrow \mathfrak{P}_X \in \text{Hol}(\mathcal{M}) \times \mathbb{C}^\infty(\mathcal{M}), \\ \forall \mathfrak{k}_X^i \text{ Killing vector} &\longrightarrow (\mathfrak{P}_X)_i{}^j \equiv \nabla_i \mathfrak{k}_X^j \in \text{Hol}(\mathcal{M}). \end{aligned} \quad (9)$$

satisfying certain equivariance conditions. (For a general characterization of the moment maps, see, for instance, <https://arxiv.org/abs/1605.05559>, accessed on 17 December 2025). For several applications, one considers moment maps with values in a specific subalgebra $\mathbb{H}_0$ of $\text{Hol}(\mathcal{M})$. For instance, if $\mathcal{M}$ is Kähler, Special Kähler, or Hyper-Kähler, $\mathbb{H}_0 = \mathfrak{u}(1)$, Lie algebra of the $U(1)$ group of Kähler transformations. In this case, the compact $\mathfrak{u}(1)$ generator has an invertible action on the tangent space to $\mathcal{M}$. The curvature $\mathcal{K}$ of the corresponding $U(1)$ -connection is a non-singular closed 2-form which provides a symplectic structure on the manifold, namely a maximal rank, closed 2-form $\omega = \mathcal{K}$. In general, the existence of a symplectic 2-form is not related to the metric properties of the manifold, or to the very existence of a metric. On a Riemannian manifold, endowed with a symplectic 2-form, we require the latter to be consistent with the Riemannian structure and, in particular, with its isometries. This is only possible if the Lie derivative of ω , with respect to all Killing vectors, vanishes. The only possibility for this is that the manifold be of Kähler type and ω is proportional to the Kähler 2-form $\mathcal{K}$, since, in this case, the $\mathfrak{u}(1)$ subalgebra, defining $\mathcal{K}$, is in the center of the Holonomy algebra. Kähler, Special Kähler, and Hyper-Kähler manifolds are important ingredients in supergravity/supersymmetric gauge theories. Indeed, they play a very important role in the whole landscape of supersymmetric field theories: in particular, they are the basic building blocks in the construction of scalar potentials (see, for instance, the “Physics Report” [35], the book [36], and the general paper [37]). Moment maps also play a fundamental role in the resolution of singularities via Kähler and Hyper-Kähler quotients à la Kronheimer (for a review, see, for instance, the lecture notes [38] and all the vast literature quoted therein). In the series of papers [30–34], the authors rely on moment maps as a fundamental ingredient in the construction of partition functions for symplectic manifolds and the explicit examples they present, namely the cases of the hyperbolic plane and of the Siegel plane (see also [1] for the role that the latter might play in Machine Learning), the symplectic structure utilized to define the moment maps is the one provided by the Kähler 2-form; hence, it applies to the very manifold $\mathcal{M}$, which in the considered examples is indeed Kählerian, rather than to its tangent bundle $\mathcal{T}\mathcal{M}$ which, instead, is the geometrical substratum of the geodesic dynamical system that can be defined on every Riemannian manifold $\mathcal{M}$.

1.4.2. Coadjoint Orbits

In the discussion of the thermodynamics that might be associated with symmetric spaces U/H , another source of possible conceptual confusions is given by the symplectic structure, named after Kirillov–Kostant–Souriau, that can be defined on coadjoint orbits of any Lie group G . This matter is presented in a crystal clear form in chapter 5 of the book [39].

As we just anticipated, and we show systematically in Sections 2 and 3 we can define generalized temperatures and partition functions on a smooth manifold $\mathcal{M}$, whenever the latter is endowed with a bona fide symplectic structure, namely a closed antisym-

metric 2-form ω of maximal rank, and we have a Lie group G acting on $\mathcal{M}$ by means of diffeomorphisms:

$$\forall g \in G \quad \mathcal{D}(g) : \mathcal{M} \longrightarrow \mathcal{M}; \quad \forall g_1, g_2 \in G \quad \mathcal{D}(g_1 \cdot g_2) = \mathcal{D}(g_1) \circ \mathcal{D}(g_2) \quad (10)$$

which are generated by Hamiltonian vector fields, namely Killing vector fields $\mathfrak{k}_A$ (see Definition (5) of symplectic Killing vector fields), satisfying its Lie Algebra $\mathbb{G}$:

$$[\mathfrak{k}_B, \mathfrak{k}_C] = f_{BC}^A \mathfrak{k}_A; \quad A, B, C = 1, 2, \dots, \dim \mathbb{G} \quad (11)$$

Indeed, each symplectic Killing vector field $\mathfrak{k}$ on the symplectic manifold $\mathcal{M}$ is associated with a moment map $\mathfrak{P}_{\mathfrak{k}}$, as already anticipated in Equation (8), which is a function on the manifold $\mathcal{M}$:

$$\mathfrak{P}_{\mathfrak{k}} : \mathcal{M} \longrightarrow \mathbb{R} \quad (12)$$

locally satisfying the condition:

$$i_{\mathfrak{k}} \cdot \omega = d\mathfrak{P}_{\mathfrak{k}} \quad (13)$$

where $i_X \cdot$ is the contraction operation along the vector field X acting on any p -form and d is the exterior derivative also acting on any p -form.

The moment maps are Hamiltonians and can be used to define partition functions as in Equation (6), where a candidate generalized temperature is any element of the Lie algebra $\mathbb{G}$.

Keeping this essential point in mind, we turn to coadjoint orbits of a Lie group G .

As explained in [39], for any Lie group G , one can define the dual $\mathbb{G}^*$ of its Lie algebra $\mathbb{G}$, which is also a vector space of the same dimension, and from the adjoint action of the group on $\mathbb{G}$:

$$\forall g \in G, \forall \mathbf{X} \in \mathbb{G} : \text{Adj}_g(\mathbf{X}) \equiv g^{-1} \mathbf{X} g \in \mathbb{G} \quad (14)$$

We obtain the coadjoint action of G on $\mathbb{G}^*$ as follows:

$$\forall g \in G, \forall \mathbf{X} \in \mathbb{G}, \forall \lambda \in \mathbb{G}^*; \quad \text{CoAdj}_g(\lambda)(\mathbf{X}) \equiv \lambda(\text{Adj}_g(\mathbf{X})) = \lambda(g^{-1} \mathbf{X} g) \quad (15)$$

Fixing a particular element $\lambda \in \mathbb{G}^*$, the coadjoint orbit $\mathcal{O}_\lambda$ is defined as the subset of elements of $\mathbb{G}^*$ that are images of λ under the coadjoint action of some group element of G :

$$\forall \lambda \in \mathbb{G}^*; \quad \mathcal{O}_\lambda = \left\{ \mu \in \mathbb{G}^* \mid \exists g \in G / \text{CoAdj}_g(\lambda) = \mu \right\} \quad (16)$$

Equation (16) can be decoded in a more friendly and usable way if, on the Lie algebra $\mathbb{G}$, which is a vector space, we introduce a non-degenerate positive definite symmetric scalar product $\langle, \rangle$ so that any element of $\mathbb{G}^*$, by definition a linear functional on $\mathbb{G}$, can be described as follows:

$$\forall \lambda \in \mathbb{G}^*, \forall \mathbf{X} \in \mathbb{G} : \quad \lambda(\mathbf{X}) = \langle \lambda^\dagger, \mathbf{X} \rangle \quad \text{where } \lambda^\dagger \in \mathbb{G} \quad (17)$$

Given a set of generators T_A that form a basis for the vector space $\mathbb{G}$, we have the symmetric, invertible, positive definite matrix κ_{AB} defined below, together with its inverse:

$$\kappa_{AB} \equiv \langle T_A, T_B \rangle; \quad \kappa^{AB} \equiv \left(\kappa^{-1} \right)^{AB} \quad (18)$$

and Equation (17) becomes:

$$\lambda^\dagger = \lambda^A T_A; \quad \mathbf{X} = X^B T_B; \quad \lambda(\mathbf{X}) = \lambda^A X^B \kappa_{AB} \quad (19)$$

The adjoint representation of the Lie group is explicitly given by the adjoint matrix defined below:

$$g^{-1} T_A g = \mathcal{A}(g)_A{}^B T_B \quad (20)$$

and the coadjoint representation is defined by the position:

$$\langle \mathcal{CA}(g)_A{}^P T_P, T_R \rangle = \langle T_A, \mathcal{A}(g)_R{}^Q T_Q \rangle \quad (21)$$

which matrix-wise implies:

$$\mathcal{CA}(g) = \kappa \cdot \mathcal{A}(g) \cdot \kappa^{-1} \quad (22)$$

If λ^A are the coefficients of the element $\lambda^\dagger \in \mathbb{G}$ that defines the orbit, the latter is formed by all those elements of $\mathbb{G}$ that have the following form:

$$\mu^\dagger(g) = \mu^A(g) T_A \equiv \underbrace{\lambda^B \mathcal{CA}(g)_B{}^A}_{\mu^A(g)} T_A \in \mathbb{G}; \quad g \in G \quad (23)$$

One might think that, as g varies in the group G for a generic choice of $\lambda^\dagger$, the element $\mu^\dagger(g)$ will span the entire Lie algebra $\mathbb{G}$. If this were true, the coadjoint orbit would be diffeomorphic to the Lie group G . However, this is not true, since there is always a non-trivial Lie subgroup $\mathfrak{S}(\lambda^\dagger) \subset G$ for which the adjoint action on $\lambda^\dagger$ is trivial:

$$g^{-1} \lambda^\dagger g = \lambda^\dagger \quad \text{iff } g \in \mathfrak{S}(\lambda^\dagger) \quad (24)$$

That in Equation (24) is the very definition of the stabilizer subgroup of the Lie algebra element $\lambda^\dagger$, and such a subgroup is never trivial since it includes at least the one-dimensional subgroup generated by $\lambda^\dagger$ itself; for special choices of $\lambda^\dagger$, the stabilizer can be much larger. This means that the coadjoint orbit $\mathcal{O}_\lambda$ is always diffeomorphic to a coset manifold, namely $G/\mathfrak{S}(\lambda^\dagger)$. The non-degenerate symplectic form ω of Kirillov–Kostant–Souriau is not defined on G ; rather, it is defined on each coadjoint orbit labeled by a Lie algebra element λ , namely on a coset manifold $G/\mathfrak{S}(\lambda^\dagger)$. Hence, the symplectic manifolds that admit a Hamiltonian action of the group G are already all captured by the scan of all coset manifolds G/H where the subgroup $H \subset G$ stabilizes some non-trivial element $\lambda^\dagger$ of the Lie algebra $\mathbb{G}$.

Just for completeness, let us mention that the Kirillov–Kostant–Souriau symplectic 2-form ω^{KKS} defined on a coadjoint orbit is very simply given. Let $\mathbf{t}_A^\sharp$ be the realization on the orbit $\mathcal{O}_\lambda$, namely on the coset manifold $G/\mathfrak{S}(\lambda^\dagger)$ of the invariant vector fields $\mathbf{t}_A$ spanning the Lie algebra $\mathbb{G}$. They form a basis of sections of the tangent bundle $\mathcal{TO}_\lambda$. Hence, a 2-form ω is completely defined if we give its value on any pair of such vector fields. The Kirillov–Kostant–Souriau form is defined by setting

$$\omega_\lambda^{KKS}(\mathbf{t}_A^\sharp, \mathbf{t}_B^\sharp) = f_{AB}{}^C \kappa_{CE} \lambda^E \quad (25)$$

We will restrict ourselves to the case in which $G = U$ is a semisimple, isometry group of a symmetric space U/H . In this case, κ_{AB} is non-singular, and H is the maximal compact subgroup of U . The KKS symplectic form is invariant under U only if the element $\lambda^\dagger$ is central in $\mathbb{H}$, the Lie algebra of H , namely if U/H is Kähler and $\lambda^\dagger$ corresponds to the Kähler $\mathfrak{u}(1)$ generator.

1.5. Clearcut Distinctions

Notwithstanding whether the manifold $\mathcal{M}$ has a symplectic structure or not, its tangent bundle $\mathcal{T}\mathcal{M}$ always has the symplectic structure associated with the Hamiltonian description of geodesic equations on $\mathcal{M}$. Here comes the first essential distinction. Whenever we have a canonical dynamical system on a symplectic manifold $(\mathcal{SM}_{2n}, \omega)$, like the geodesic one where $\mathcal{SM}_{2n} = \mathcal{TM}_n$, we can construct standard thermodynamics in geometrical formulation, starting from the minimization of the Shannon entropy functional and arriving at Gibbs states of the form:

$$\begin{aligned} G(\lambda, \Phi) &= \frac{\exp[-\lambda \cdot \mathfrak{H}(\Phi)]}{Z(\lambda)} \\ Z(\lambda) &\equiv \int_{\mathcal{SM}} \exp[-\lambda \cdot \mathfrak{H}(\Phi)] d\lambda(\Phi) \end{aligned} \quad (26)$$

where $\mathfrak{H}(\Phi)$ is the multiplet of k Hamiltonians in involution admitted by the dynamical system (if the dynamical system is Liouville integrable, one has $k = n$, the dimension of the symplectic manifold being $2n$. In general, $1 \leq k < n$, and it is just 1 for a generic dynamical system without conserved charges; including the standard one defined by the Legendre transform of the Lagrangian):

$$\mathfrak{H}(\Phi) = \{\mathfrak{H}_1(\Phi), \dots, \mathfrak{H}_k(\Phi)\}; \quad \underbrace{\{\mathfrak{H}_i, \mathfrak{H}_j\}}_{\text{Poisson bracket}} = 0 \quad \forall i, j \quad (27)$$

$\lambda \in \mathbb{R}^k$ is a vector of generalized temperatures, and $Z(\lambda)$ is the partition function.

Following the conception reviewed in Section *Conditional Minimalization of Information and the Partition Function* and making reference to Equation (39) and following ones, we should also note that the stochastic vector variable $\mathbf{X}(q)$, of which we fix the average value in order to define a probability distribution that extremizes the Shannon functional with constraints, does not need to be a set of Hamiltonians in involution and there is no need of integrability of any dynamical system in order to introduce a **generalized thermodynamics**. In this case, the space of events (see Appendix B for the basic definitions of probability theory) is a symplectic manifold endowed with Hamiltonian vector fields. We can use the moment-maps of the latter as a convenient set of stochastic variables in order to introduce a generalized thermodynamics by fixing their average values. Yet this is only a subclass of examples in a general class.

The Lie group thermodynamics advocated in [30–34] leads to Gibbs states of the form:

$$\begin{aligned} G(\beta, Y) &= \frac{\exp[-\beta \cdot \mathfrak{P}(Y)]}{Z(\beta)} \\ Z(\beta) &\equiv \int_{\mathcal{M}} \exp[-\beta \cdot \mathfrak{P}(Y)] dg[Y] \end{aligned} \quad (28)$$

where Y denotes the coordinates of the Riemannian manifold $(\mathcal{M}, g)$ and the generalized temperature $\beta \in \Omega$ is an element of the Lie algebra $\mathbb{G}$ of the isometry group G such that the integral defining the partition function is convergent. Let us observe that in Equation (28) $dg[Y]$ is the Riemannian integration measure which coincides with the Liouville measure (7) if $(\mathcal{M}, g)$ is a Kähler manifold and the Kähler 2-form $\mathcal{K}$ is utilized to define the symplectic structure on $\mathcal{M}$.

A third possibility, which is the conceptual framework underlining [30–34], is to use the setup of Equation (28) using, however, as substratum manifold $\mathcal{M}$ some coadjoint orbit of $\mathcal{O}_{\lambda^+}$ under the action of a group G of some special Lie algebra element $\lambda^+ \in \mathbb{G}$. As extensively discussed in the previous subsection, coadjoint orbits are, anyhow, coset

manifolds, and it is conceptually much more economic to start from the coset manifold structure, asking oneself the question: *Given G/H , what is the Lie algebra element $\lambda^\dagger \in \mathbb{G}$ that is stabilized by the chosen subgroup H ?* The answer is fairly simple. In view of the discussion of the previous section, it is clear that $\lambda^\dagger \in \mathbb{H} \subset \mathbb{G}$ since the one-parameter group generated by $\lambda^\dagger$ must be contained in H . Therefore, the condition is just:

$$[\mathbb{H}, \lambda^\dagger] = 0 \quad (29)$$

The solution of the constraint (29) is immediate. The Lie algebra $\mathbb{H}$ must have the following structure:

$$\mathbb{H} = \mathbb{H}' \oplus \mathbb{H}_0; \quad \mathbb{H}_0 = \text{span}[\lambda^\dagger] \quad (30)$$

and the unidimensional Lie algebra $\mathbb{H}_0$ is either $\mathbb{R}$ or $\mathfrak{u}(1)$ depending on whether the Lie algebra element $\lambda^\dagger$ is non-compact or compact.

1.5.1. Kähler Non-Compact Symmetric Spaces

In view of Cartan Neural Networks, where the relevant manifolds are non-compact symmetric spaces U/H , with $H \subset U$, the maximal compact subgroup of a non-compact simple Lie group, it follows that $\mathbb{H}$ is compact and $\mathbb{H}_0 = \mathfrak{u}(1)$. This has a universal and simple interpretation: the presence in the isotropy subgroup H of a factor $U(1)$ simply means that U/H is a Kähler manifold, and that the symplectic structure is provided by the Kähler 2-form.

1.5.2. Hence Two Cases

Summarizing the previous discussion, we conclude that there are just two distinct cases of geometrical thermodynamics related to non-compact symmetric spaces U/H

- (A) The thermodynamics associated with the Geodesic Dynamical System (GDS) on U/H , where the symplectic structure is that provided by the phase-space of the GDS, existing for all manifolds and in particular for all symmetric spaces U/H .
- (B) Kähler thermodynamics on the symmetric spaces U/H defined by

$$G_K(\beta, Y) = \frac{\exp[-\beta \cdot \mathfrak{P}(Y)]}{Z_K(\beta)} \quad (31)$$

$$Z_K(\beta) \equiv \int_{U/H} \exp[-\beta \cdot \mathfrak{P}(Y)] \underbrace{\mathcal{K} \wedge \mathcal{K} \wedge \dots \wedge \mathcal{K}}_{n\text{-times}} \quad (32)$$

$$\dim \frac{U}{H} = 2n; \quad n \in \mathbb{N} \quad (33)$$

$$\mathcal{K} = \text{Kähler 2-form} \quad (34)$$

$$i_{\mathfrak{k}_A} \mathcal{K} = d\mathfrak{P}_A \quad (35)$$

where $\mathfrak{P}(Y)$ denotes the vector of moment maps $\mathfrak{P}_A(Y)$ associated with a basis of Killing vectors $\mathfrak{k}_A$ that correspond to a basis T_A of generators of the U Lie algebra and $\beta = \beta^A T_A \in \Omega \subset U$ is a generalized temperature vector such that the partition function integral (32) converges.

1.6. Relevance for Cartan Neural Networks

In the new paradigm of Cartan Neural Networks, all the manifolds that model the hidden layers and to which data are injected are diffeomorphic to as many solvable Lie groups $\mathcal{S}$. Furthermore, they have a simple group U of isometries, which gives rise to a Lie algebra $\mathbb{U}$. For all these manifolds, the geodesic dynamical system is completely integrable, and one can construct a nice algebraic resetting of the corresponding Hamiltonian setup liable to be

used in the study of Gibbs states of the conventional type defined by Equation (26). The use of such Gibbs states in Machine Learning algorithms based on the CaNNs paradigm is a perspective to be investigated with care. However, it must be noted that, as we show in the sequel, the Gibbs probability distribution depends only on the momenta (velocities) and not on the positions in the manifold U/H . Hence, if one is interested in probability distributions on the very manifold to which data are mapped, the Geodesic Dynamical System thermodynamics seems to be of little use.

On the other hand, in the organization of non-compact symmetric spaces into Tits Satake universality classes, there are entire classes consisting of Kähler manifolds, for instance, the $r = 2$ class (see [1] for details on the classification):

$$\mathcal{M}^{[2,q]} \equiv \frac{SO(2, 2+q)}{SO(2) \times SO(2+q)} \quad (36)$$

Hence, for such cases, the possibility of defining Kähler thermodynamics and corresponding Gibbs states à la Souriau as in Equation (31) arises. Once again, the use of such generalized Gibbs states in Machine Learning algorithms has to be studied, yet its viability is guaranteed by the already existent applications to radar signal analysis and to other time series discussed in the thesis [40] and in all references quoted therein (in particular [5]). Indeed, such Gibbs states provide a Gaussian-like probability distribution on the very manifold U/H , rather than on the fibres of its tangent bundle.

The perspective of Gibbs states à la Souriau for non-compact symmetric spaces U/H that are Kählerian requires a systematic theoretical study, which seems to be so far missing, namely that of an intrinsic characterization of the subspace Ω of generalized temperatures inside the relevant $\mathbb{U}$ algebras. We consider such a study an interesting priority and address it in the present paper, both in its general form and in the case study of two examples. We found a general answer that was so far missing: the space of generalized temperature is the adjoint orbit of the positivity domain in the space of the Cartan subalgebra of the compact subalgebra $\mathbb{H}$ (see Section 7).

Furthermore, let us also recall that in [31], Barbaresco has claimed that the geometry of Lie Group thermodynamics is to be identified with the **Riemannian Geometry of Information** introduced several years ago by Rao [41] and Chentsov [42] (see the review paper [43] and the book [44] whose author is frequently credited for the introduction of Information Geometry in the Data Science community). Once the conceptual framework is clarified, as we hope to have done with the present paper, the relation between Kähler thermodynamics and the Riemannian metric naturally associated with equilibrium states, via the general setup of **generalized geometrical thermodynamics**, becomes clear and universal, as we are going to show.

1.7. Outline of This Paper

The present paper is organized as follows. First, we briefly recall the basic principles of generalized thermodynamics and of its link with Shannon's information functional. Next, we analyze the general structure of the Geodesic Dynamical System and its specialization to the case of symmetric spaces U/H . This is instrumental in completing the Poissonian structure on dual solvable Lie algebras into a full symplectic structure on the tangent bundle of non-compact symmetric spaces. In this perspective, we can study examples of generalized thermodynamics associated with integrable dynamical systems and show that they are too simple and of little interest for Machine Learning applications. We come next to discuss generalized thermodynamics à la Souriau, and it is in this context that we obtain our most relevant results that are summarized in the conclusive Section 8. We do not anticipate them here. We just say that, according to our opinion, thanks to a strategic use of

the metric equivalence of non-compact symmetric spaces U/H with appropriate solvable Lie groups, we have established generalized thermodynamics à la Souriau on clear general principles for all non-compact symmetric spaces that are Kähler manifolds, introducing, in this way, a new powerful weapon for Machine Learning algorithms.

We have equipped our paper with several mathematical and physical appendices in order to make it self-contained and readable to a larger audience.

2. Shannon Information Entropy and the Partition Function

In his celebrated 1948 paper [45], Claude Elwood Shannon introduced what is called the entropy of information relative to a probability density ρ defined on some measurable space Ω (see Appendix B for a summary of the fundamental principle and concepts of probability theory as exposed in standard textbooks like [46]).

Let $\mathbf{q} \in \Omega$ be a point in the stochastic space we consider; let $d\mu(\mathbf{q})$ be the integration measure on Ω ; and let $\rho(\mathbf{q}) \in [0, 1]$ be the value in $\mathbf{q}$ of the probability density that is obviously normalized as follows:

$$N[\rho] \equiv \int_{\Omega} \rho(\mathbf{q}) d\mu(\mathbf{q}) = 1 \quad (37)$$

The measure of information contained in the probability distribution ρ was defined by Shannon by means of the following functional:

$$\mathcal{I}[\rho] \equiv - \int_{\Omega} \rho(\mathbf{q}) \log [\rho(\mathbf{q})] d\mu(\mathbf{q}) \quad (38)$$

Conditional Minimalization of Information and the Partition Function

The precise conceptual connection between Information Theory and Statistical Mechanics and thus with Thermodynamics can be made through the notion of conditional minimization introduced by Jaynes in 1957 who, in the papers [47,48], clarified transparently and definitively the logical relationship between Shannon's theory and Statistical Thermodynamics. Thanks to recent works [29,49–52], this relationship is further clarified in geometric terms and completes the design of the conceptual framework in which the association of a Riemannian metric with Thermodynamics obtains a solid foundation.

We pose the following problem: determine the probability distribution that extremizes the functional $\mathcal{I}[\rho]$ under the following two conditions:

- (A) The correct normalization (37) should hold true.
- (B) The average value of a certain stochastic vector $\mathbf{X}$ should be fixed to a certain precise vector $\mathbf{x} \in \mathbb{V}$:

$$\langle \mathbf{X} \rangle \equiv \int_{\Omega} \mathbf{X}(\mathbf{q}) \rho(\mathbf{q}) d\mu(\mathbf{q}) = \mathbf{x} \in \mathbb{V} \quad (39)$$

The classical way to solve this problem is to use variational calculus in the presence of Lagrange multipliers. One introduces $r + 1$ multipliers: λ_0 associated with the normalization constraint (37) and $r = \dim \mathbb{V}$ multipliers λ^i that we can regard as the components of a vector in $\boldsymbol{\lambda} \in \mathbb{V}^*$ that are associated with the constraints (39). Thus, the new functional to be extremized is as follows:

$$\mathcal{F}[\rho] = -\mathcal{I}[\rho] - \lambda_0 (N[\rho] - 1) + \boldsymbol{\lambda} \cdot (\langle \mathbf{X} \rangle - \mathbf{x}) \quad (40)$$

The variation of the functional in $\delta\rho$ yields

$$\frac{\delta \mathcal{F}[\rho]}{\delta \rho} = \log[\rho] + 1 - \lambda_0 + \boldsymbol{\lambda} \cdot \mathbf{X} = 0 \quad (41)$$

which implies:

$$\rho(\mathbf{q}) = \exp[\lambda_0 - 1 - \boldsymbol{\lambda} \cdot \mathbf{X}(\mathbf{q})] \quad (42)$$

Imposing the normalization constraint (37) fixes the value of λ_0 so that the final expression of the extremal probability distribution is the following :

$$\rho_{ex}(\mathbf{q}) = \frac{\exp[-\boldsymbol{\lambda} \cdot \mathbf{X}(\mathbf{q})]}{Z(\boldsymbol{\lambda})} \quad (43)$$

where

$$Z(\boldsymbol{\lambda}) \equiv \int_{\Omega} \exp[-\boldsymbol{\lambda} \cdot \mathbf{X}(\mathbf{q})] d\mu(\mathbf{q}) \quad (44)$$

is the **Partition Function** and, for reasons that will become immediately clear, the following object

$$\mathcal{H}^{stoch}(\boldsymbol{\lambda}) = -\log[Z(\boldsymbol{\lambda})] \quad (45)$$

is named the **stochastic Hamiltonian**. As a consequence of the definition (45), the value $\mathbf{x}$ imposed to the stochastic vector $\mathbf{X}$ is obtained from the Hamiltonian by means of a derivative:

$$\mathbf{x} = d_{\boldsymbol{\lambda}} \mathcal{H}^{stoch}(\boldsymbol{\lambda}) \Rightarrow \text{short hand for } x_i = \frac{\partial}{\partial \lambda^i} \mathcal{H}^{stoch}(\lambda_1, \dots, \lambda_r) \quad (46)$$

Calculating Shannon Entropy Functional (38) on the extremal probability distribution (43) with elementary algebra, we obtain

$$-\mathcal{I}[\rho_{ex}] = \mathcal{H}^{stoch}(\boldsymbol{\lambda}) - \boldsymbol{\lambda} \cdot \mathbf{x} = \mathcal{H}^{stoch}(\boldsymbol{\lambda}) - \lambda^i \frac{\partial}{\partial \lambda^i} \mathcal{H}^{stoch}(\boldsymbol{\lambda}) \quad (47)$$

which has the form of a Legendre transform. Hence, the Shannon functional plays the same role as that of a Lagrangian; the stochastic Hamiltonian is indeed a Hamiltonian; the intensive variables of Thermodynamics (i.e., the Lagrange multipliers $\boldsymbol{\lambda}$) are the momenta; and the average values x^i are the coordinates.

Next, we turn to classical thermodynamics. We refer to Appendix C for a recollection of its basic concepts and constructions, which are presented in order to fix notation and also for the benefit of those readers who are not physicists by education. In the next Section 3, we illustrate the geometric reformulation of classical thermodynamics in the context of contact and symplectic geometry, which leads to the introduction of the new notion of thermodynamical curvature.

Note that what is named metric of Information Geometry in the Machine Learning literature is the following Hessian obtained from a parameterized by a vector $\boldsymbol{\lambda} = \{\lambda_1, \dots, \lambda_r\}$ probability distribution $\rho_{\boldsymbol{\lambda}}(\mathbf{X}(\mathbf{q}))$ of the stochastic variable $\mathbf{X}(\mathbf{q})$ over the manifold of events:

$$ds_{info}^2 \equiv \frac{\partial^2}{\partial \lambda^i \partial \lambda^j} \log[\rho_{\boldsymbol{\lambda}}(\mathbf{X}(\mathbf{q}))] d\lambda^i \times d\lambda^j \quad (48)$$

When the probability distribution is the generalized Gibbs one of Equation (43), we find

$$ds_{info}^2 \equiv \frac{\partial^2}{\partial \lambda^i \partial \lambda^j} \mathcal{H}^{stoch}(\boldsymbol{\lambda}) d\lambda^i \times d\lambda^j \quad (49)$$

As we will show in the next section, the metric (49) coincides with the thermodynamics metric introduced in geometrical thermodynamics, much before the advent of Machine Learning contributions.

3. Geometrical Structure of Thermodynamics

In this section, we present the reformulation of classical thermodynamic laws in geometrical terms, based on what we explained in Section 2, where we elucidated the relation between information theory and statistical mechanics. It is now time to show how classical thermodynamic laws are linked with the notion of a contact manifold $(\mathcal{M}^{2n+1}, \xi_\alpha)$, defined by a suitable contact 1-form α and certain Legendrian submanifolds $\mathcal{L}_n \subset \mathcal{M}^{2n+1}$ of the latter, specifically defined and identifiable with Lagrangian submanifolds $\mathcal{L}_n \subset \mathcal{S}_{2n}$ of a symplectic manifold $(\mathcal{S}_{2n}, \omega)$ that is canonically associated with the contact one according to the scheme (A43). The Lagrangian vision leads to the definition of a canonical Riemannian metric induced on $\mathcal{L}_n$, which is the most relevant novelty of the introduced conceptual framework.

Indeed, calculating the *thermodynamical curvature* is a new powerful investigation tool in all applications.

The intuition of the relevance of *thermodynamical curvature* as a probe of molecular interactions at the mesoscopic level is indeed, as we stressed in the introductory Section 1, particularly due to Ruppeiner.

There is, however, something even more pertinent that should be stressed. Recalling the fundamental relation between Information Theory and Statistical Mechanics outlined in Section 2 and contextually illustrated in Appendix C it appears that the geometric reformulation of classical thermodynamics has a much wider scope than physical or chemical systems. Indeed, any conditioned probability distribution describing whatever phenomena and fitting to whatever Big Data system defines a thermodynamical setup and would lead to equations of state if we knew the probability distribution and were able to calculate the partition function. The geometrical formulation of the equations of state as embedding functions of Lagrangian submanifolds is a scheme that can be utilized in an inverse engineering procedure to work out the probability distribution and possibly learn it from Data Behavior. This is a challenging possibility for Deep Learning, completely unexplored at the present moment.

3.1. The Geometric Reformulation

To develop the program announced above, we need only to collect the ideas already introduced, focusing on the standard Darboux expression of a contact form given in Equation (A34) and on Equation (47), which shows that the Functional measuring Information $\mathcal{I}[\rho_{ex}]$ is related to the stochastic Hamiltonian $\mathcal{H}(\lambda)$ by a Legendre transform. Summarizing, we can say that in thermodynamics, we have $n + 1 \geq 3$ extensive variables collectively denoted x_0, x_i ($i = 1, \dots, n$) which explicitly might be identified as follows:

1. Internal Energy U ;
2. Entropy S ;
3. Volume V ;
4. Molar Fractions N_ℓ ($\ell = 1, \dots, n - 3$);

and n -intensive variables; collectively denoted λ^i and explicitly identified as:

1. Temperature T ;
2. Pressure P ;
3. Chemical Potentials μ^ℓ ($\ell = 1, \dots, n - 3$).

The first principle of thermodynamics, combined with the second, can be formulated by stating that the following differential form vanishes:

$$0 \approx \tilde{\alpha} \equiv dU - TdS + PdV - \sum_{\ell=1}^{n-3} \mu^\ell dN_\ell = dx_0 + \sum_{i=1}^n \lambda^i dx_i \quad (50)$$

The last form à la Darboux of $\tilde{\alpha}$ follows from the identification of x_0 with the internal energy U and the remaining coordinates as follows $\lambda = \{-T, P, -\mu^\ell\}$ e $x^i = \{S, V, N_\ell\}$. Obviously, because of its form, à la Darboux $\tilde{\alpha}$ satisfies the defining condition in order to be a contact 1-form, namely:

$$\underbrace{d\tilde{\alpha} \wedge d\tilde{\alpha} \wedge \cdots \wedge d\tilde{\alpha}}_{n \text{ times}} \wedge \tilde{\alpha} \neq 0 \quad (51)$$

To better conjugate the emerging contact geometry underlying thermodynamics with the Equation (47) that identifies, minus a multiplicative factor, the information measure with thermodynamic entropy (see Equation (A93)), it is convenient to multiply the form $\tilde{\alpha}$ introduced in Equation (50) times a factor $1/(k_B T)$, obtaining in this way:

$$\begin{aligned} \alpha &= (k_B T)^{-1} \tilde{\alpha} = -k_B dS + (k_B T)^{-1} dU + (k_B T)^{-1} P dV - \sum_{\ell=1}^{n-3} (k_B T)^{-1} \mu^\ell dN_\ell \\ &= d\mathcal{I} - \sum_{i=1}^n \lambda^i dx_i \end{aligned} \quad (52)$$

where the new definition of the $2n + 1$ coordinates is the following one:

$$\lambda = \left\{ -\frac{1}{k_B T}, -\frac{P}{k_B T}, \frac{\mu^\ell}{k_B T} \right\}; \quad \mathbf{x} = \{U, V, N_\ell\}; \quad x_0 = \mathcal{I} \quad (53)$$

Obviously, the 1-form α satisfies the same condition (51) as $\tilde{\alpha}$ and, therefore, it is also a contact 1-form. Thus, we have defined a contact manifold $(\mathcal{M}_{2n+1}, \xi)$, where $\xi = \ker \alpha$ is the contact structure $\mathcal{M}_{2n+1} = \mathbb{R}^{2n+1}$ has $2n + 1$ coordinates $\{I, \lambda, \mathbf{x}\}$, the variable I being, at this stage, a free coordinate, just as λ and $\mathbf{x}$.

3.1.1. Legendrian Submanifolds

Definition A7 of Legendrian submanifolds given in Appendix A states that they are isotropic submanifolds of maximal dimension n of a contact manifold $\mathcal{M}_{2n+1}$. On the other hand, we recall that an isotropic submanifold is a submanifold such that its tangent bundle is in the kernel of the contact form, namely, the 1-form α vanishes on each Legendrian submanifold. The great intuition of the authors of [49,50] has been that of identifying, independently from the utilized coordinates and hence in an intrinsic way, the **thermodynamic equilibrium states** with the points of specific **Legendrian submanifolds** of the ambient space. In terms of the theory of conditional minimization discussed in Section *Conditional Minimalization of Information and the Partition Function*, it is very simple to define the Legendrian submanifolds that represent the thermodynamic equilibrium states.

Definition 1. Any admissible thermodynamic state can be identified with a point in the following n -dimensional submanifold of the contact manifold:

$$\mathcal{L}_n = \left\{ I = \mathcal{I}(\lambda, \mathbf{x}), x_i = \frac{\partial}{\partial \lambda^i} \mathcal{H}(\lambda) \right\} \subset \mathcal{M}_{2n+1} \quad (54)$$

Theorem 1. The submanifold $\mathcal{L}_n$ defined by means of Equation (54) is isotropic and hence Legendrian.

Proof. The proof is extremely simple. It suffices to recall Equation (47). Using that relation, we can evaluate the total differential dI , as follows:

$$dI = d\mathcal{I}(\boldsymbol{\lambda}, \mathbf{x}) = d(H(\boldsymbol{\lambda}) - \boldsymbol{\lambda} \cdot \mathbf{x}) = \left(\frac{\partial}{\partial \lambda^i} H(\boldsymbol{\lambda}) - x^i \right) d\lambda^i - \lambda^i dx_i = -\lambda^i dx_i \quad (55)$$

Hence, on the submanifold (54), we have $dI + \lambda^i dx_i = 0 \quad \square$

3.1.2. The Lagrangian Submanifold and Its Metric

Given the original contact variety $\mathcal{M}^{2n+1}$ with the contact 1-form given by the presentation (52), we see at once that the Reeb vector field is

$$\mathbf{R} = \frac{\partial}{\partial \mathcal{I}} \quad (56)$$

In fact, it satisfies the two conditions:

$$\alpha(\mathbf{R}) = 1; \quad d\alpha(\mathbf{R}, \mathbf{X}) = 0 \quad \forall \mathbf{X} \in \Gamma[\mathcal{T}\mathcal{M}^{2n+1}, \mathcal{M}^{2n+1}] \quad (57)$$

On the other hand, from the general discussion in Appendix A.8, we know that every $2n$ -dimensional submanifold of a contact manifold $\mathcal{M}^{2n+1}$ that is transverse to the Reeb vector field of the latter is a symplectic manifold $\mathcal{S}^{2n}$ whose symplectic 2-form is the restriction to $\mathcal{S}^{2n}$ of the exterior differential of the contact 1-form:

$$\omega = d\alpha|_{\mathcal{S}^{2n}} \quad (58)$$

Hence, applying these general notions to the case at hand, we see that the symplectic variety transverse to Reeb's vector (56) is given by the following projection map:

$$\pi : \mathcal{M}^{2n+1} \rightarrow \mathcal{S}^{2n}; \quad \pi(\mathcal{I}, \boldsymbol{\lambda}, \mathbf{x}) = (\boldsymbol{\lambda}, \mathbf{x}) \quad (59)$$

and the symplectic 2-form is as follows:

$$\omega = - \sum_{i=1}^n d\lambda^i \wedge dx_i \quad ; \quad d\alpha = \pi^*(\omega) \quad (60)$$

Let us now consider the Legendrian submanifold $\mathcal{L}_n \subset \mathcal{M}^{2n+1}$ which contains the thermodynamic equilibrium states introduced in Definition 1. It is obvious that we can consider its image through the projection map (59):

$$\mathcal{S}^{2n} \supset \mathfrak{L}_n \equiv \pi(\mathcal{L}_n) \quad (61)$$

The important result is that the submanifold $\mathfrak{L}_n$ thus defined is a Lagrangian submanifold, namely one on which the symplectic form completely vanishes. The demonstration of this fact follows immediately from the definition. In fact, we can translate Equation (61) into the following constructive definition:

$$\mathfrak{L}_n = \left\{ x^i = \frac{\partial}{\partial \lambda^i} \mathcal{H}(\boldsymbol{\lambda}) \right\} \quad (62)$$

Using the latter, we find:

$$\omega|_{\mathfrak{L}_n} = - \sum_{i,j=1}^n d\lambda^i \wedge d\lambda^j \partial_i \partial_j \mathcal{H}(\boldsymbol{\lambda}) = 0 \quad (63)$$

which follows because of the commutativity of the partial derivatives. We can therefore conclude that the thermodynamic equilibrium states are immersed in a Lagrangian submanifold of the thermodynamic symplectic space $(\mathcal{S}^{2n}, \Omega)$, the coordinates of this latter

being λ and x in terms of traditional thermodynamic variables, specified by relations in (53).

3.1.3. The Canonical Riemannian Metric on the Lagrangian Submanifold

The Lagrangian submanifold $\mathfrak{L}_n$ is naturally equipped with a Riemannian canonical metric, which is the image

$$ds_{\mathfrak{L}_n}^2 = \iota^*(ds_{can}^2) \quad (64)$$

through the pull-back ι^* of the immersion map:

$$\iota : \mathfrak{L}_n \xhookrightarrow{\iota} \mathcal{S}^{2n} \quad (65)$$

of the canonical flat metric on the symplectic ambient manifold:

$$ds_{can}^2 = \frac{1}{2} \sum_{i=1}^n (d\lambda^i \otimes dx^i + dx^i \otimes d\lambda^i) \quad (66)$$

The canonical Riemannian metric (66) on the ambient symplectic space corresponds to assuming the standard complex structure and standard symplectic 2-form matrices

$$\mathfrak{J} = \left(\begin{array}{c|c} -i\mathbf{1}_{n \times n} & 0 \\ \hline 0 & i\mathbf{1}_{n \times n} \end{array} \right); \quad \mathfrak{K} = \frac{1}{2} \left(\begin{array}{c|c} 0 & i\mathbf{1}_{n \times n} \\ \hline -i\mathbf{1}_{n \times n} & 0 \end{array} \right) \quad (67)$$

so that, according to the general theory, the canonical Hermitian metric is indeed given by the matrix:

$$\mathfrak{G} = \mathfrak{K} \cdot \mathfrak{J} \quad (68)$$

The Riemannian metric (64) is the one that was promoted to an investigation tool of mesoscopic physical chemistry of critical phenomena, in particular by Ruppeiner and collaborators. This metric would be perfectly defined and calculable if we explicitly knew the stochastic Hamiltonian $\mathcal{H}(\lambda)$, since by means of the immersion equations we get:

$$ds_{\mathfrak{L}_n}^2 = -\mathcal{H}_{ij}(\lambda) d\lambda^i \otimes d\lambda^j; \quad \mathcal{H}_{ij}(\lambda) \equiv \partial_i \partial_j \mathcal{H} \quad (69)$$

where $\mathcal{H}_{ij}$ is named the Hessian. As we already anticipated above, the metric (69) exactly coincides with the metric (49) and, hence, with the metric (48) named the Information Geometry metric in the Machine Learning literature, when the probability distribution is the generalized Gibbs distribution (43). This is also reminiscent of the AMSY symplectic formulation of Toric Kähler Geometry in the action-angle coordinates, for which we refer the reader to [53,54] and, for a review, to [38] and to the original references there cited. The problem is that the stochastic Hamiltonian is by definition the negative of the logarithm of the partition function $Z(\lambda)$ and generally beyond the reach of explicit analytical computation, as we have repeatedly pointed out. In the absence of this generally inaccessible tool, there was an extensive research activity in devising and proposing phenomenological equations of state, each of which provides an explicit recipe for calculating the metric (64). In the next subsection, we will examine this methodology in general for the case where $n = 2$, corresponds to the canonical ensemble and is the most frequently utilized in the field of phenomenological equations of state.

3.1.4. The Lagrangian Submanifold in the Two-Dimensional Case and Its Riemannian Structure

In order to illustrate the general concepts and as a term of comparison with the subsequent examples of geometrical thermodynamics on Riemannian manifolds and, in particular, on symmetric spaces, we provide here a brief sketch of the geometrical treatment for physical thermodynamics. In the simplest situation, we deal with a 4-dimensional symplectic space with coordinates T, P, U, V , where the former two are intensive quantities (temperature and pressure), and the latter two are extensive ones: internal energy and volume (we recall that the fifth coordinate to complete the odd-dimensional contact manifold is the entropy S). In the even-dimensional space, the symplectic 2-form ω is the following:

$$\omega \equiv d[T^{-1}] \wedge dU + d[T^{-1}P] \wedge dV \quad (70)$$

The following two are assumed to be the embedding equations of the two-dimensional Lagrangian variety:

(A) The thermic equation

$$f(P, T, V, U) \equiv P - \mathcal{A}(T, V) \quad (71)$$

(B) The caloric equation

$$g(P, T, V, U) \equiv U - \mathcal{B}(T, V) \quad (72)$$

The first condition is the true equation of state. The second must be found in agreement with the constraint that the hypersurface cut out by the two constraints in the symplectic manifold should be Lagrangian (namely, should make the 2-form ω vanish). Introducing the shorthand $\mathfrak{w} \equiv \{P, T, V, U\}$, we can write

$$\mathcal{SM}_4 \supset \mathfrak{L}_2 = \{\mathfrak{w} \in \mathcal{SM}_4 \mid f(\mathfrak{w}) = 0 \text{ and } g(\mathfrak{w}) = 0\} \quad (73)$$

The surface $\mathfrak{L}_2$ is Lagrangian if the symplectic 2-form vanishes when restricted to it, which is equivalent to saying that the Poisson bracket of the two embedding functions is zero:

$$\omega|_{\mathfrak{L}_2} = 0 \quad \Leftrightarrow \quad \{f, g\} = 0 \quad (74)$$

Substituting the equations of state (71) and (72) into the symplectic form (70), we obtain:

$$\omega|_{\mathfrak{L}_2} = \frac{dT \wedge dV (T\partial_T \mathcal{A}(T, V) - \mathcal{A}(T, V) - \partial_V \mathcal{B}(T, V))}{T^2} \quad (75)$$

Hence, the constraint to be satisfied by the two immersion functions $\mathcal{A}(T, V), \mathcal{B}(T, V)$ in order, for The immersed submanifold to be Lagrangian is that the image through the projection π of a Legendrian submanifold $\mathcal{L}_2$ immersed in the contact manifold $\mathcal{M}_5$ should be the following:

$$T\partial_T \mathcal{A}(T, V) - \mathcal{A}(T, V) - \partial_V \mathcal{B}(T, V) = 0 \quad (76)$$

Therefore, the canonical Riemannian metric on the Lagrangian submanifold $\mathfrak{L}_2$ is the following:

$$ds_{\mathfrak{L}_2}^2 = d[T^{-1}] \otimes d\mathcal{B}(T, V) + d[T^{-1} \mathcal{A}(T, V)] \otimes dV \quad (77)$$

which makes sense if and only if the constraint (76) is satisfied. We can verify Equation (76) in the familiar case of Ideal Gases, recalling Equations (A112)–(A114) from which we see that in the ideal gas case, we have:

$$\mathcal{A}(T, V) = \frac{k_B N T}{V}; \quad \mathcal{B}(T, V) = \frac{3}{2} k_B T \quad (78)$$

In this case, the thermodynamical metric (77) becomes:

$$ds_{IG}^2 = -k_B \left(\frac{3}{2} \left(\frac{dT}{T} \right)^2 + N \left(\frac{dV}{V} \right)^2 \right) \quad (79)$$

which is obviously a flat metric. Indeed, it suffices to change variables ($T = \sqrt{\frac{2}{3}} \log[x]$, $V = \frac{1}{\sqrt{N}} \log[y]$) and (79) becomes proportional to the standard Euclidean metric on $\mathbb{R}^2$.

In Appendix C.5, as an illustrative counterexample, we briefly discuss the van der Waals model of a real gas equation of state, and we show that the immersion functions of the Lagrangian equilibrium submanifold (78) are substituted by the new immersion functions (A122), which also satisfy the Lagrangian constraint (76), as they should. Correspondingly, the flat thermodynamical metric (79) is replaced by its van der Waals equivalent, shown in Equation (A123), which is not flat. Its curvature, shown in Figure A2, has a non-trivial behavior and displays a singularity along the critical curve separating the gas from the liquid phase.

3.2. General Conclusion of This Section

What we have shown above is the use of the geometrical definition of equilibrium states as Lagrangian submanifolds in the context of equations of state for conventional thermodynamical systems. Yet we should keep in mind that the identification of the thermodynamical metric with the Hessian of the stochastic Hamiltonian displayed in Equation (69) is general and applies to any Gibbs state probability distribution of whatever type. Indeed, given any dynamical system, we can define the partition function as in Equation (26), and we obtain the stochastic Hamiltonian from Equation (45). This is also true for the geodesic dynamical system that we discuss in the next section and for Kähler thermodynamics of non-compact symmetric spaces discussed in later sections. The thermodynamical curvature, as we are going to see, also exists in this case and might display singular behaviors signaling critical phenomena.

4. The Geodesic Dynamical System

As recalled in the introduction and fully explained in [2], Cartan Neural Networks are based on the scheme:

$$\mathbb{V}_{input} \xrightarrow[\text{injection}]{l_{[Q, \Lambda]}} \underbrace{U_1/H_1 \xrightarrow{\hat{K}_{[W_1, \Psi_2]}^1} U_2/H_2 \xrightarrow{\hat{K}_{[W_2, \Psi_3]}^2} \dots \xrightarrow{\hat{K}_{[W_{N-1}, \Psi_N]}^{N-1}} U_N/H_N}_{\text{hidden layers}} \xrightarrow[\text{output map}]{\Rightarrow} \mathbb{V}_{output} \quad (80)$$

where the hidden layers U_i/H_i are non-compact symmetric spaces, metrically equivalent to as many solvable Lie groups $S_i \subset U_i$, and the maps $K_i : U_i/H_i \rightarrow U_{i+1}/H_{i+1}$ are isometric mappings endowed with general covariance with respect to the transformations of both the U_i and the U_{i+1} group. The initial injection map and the final output map depend on the task for which the network architecture is designed and its categorical type. For instance in the frequent case of the classification task (e.g., for images) implemented with a simple CaNN, the *initial datum* is regarded as a *single vector* Ξ_i (e.g., the list of all pixels) and the *injection map* is a linear relation between the *solvable coordinates* (see [1,2] for the general definition of solvable coordinates on U/H manifolds) Y_A of a point in the first hidden layer and the components of the datum vector:

$$Y_A = Q_A^i \Xi_i + \Lambda_A \quad (81)$$

where Q_A^i is a $\dim[U_1/H_1] \times \dim[V_{input}]$ matrix and Λ_A a $\dim[U_1/H_1]$ -dimensional vector, both being targets of the learning process (see section 6.3 of [2]). In the same case of the classification task and in the same categorical type of a simple CaNN, the output map is schematically described below (see Equation (6.11) of [2]):

$$\underbrace{\Rightarrow V_{output}}_{\text{output map}} = \underbrace{\xrightarrow{\mathfrak{S}} \mathcal{M}_+ \cup \mathcal{M}_-}_{\text{partition}} \xrightarrow{\sigma} \underbrace{[0,1]_{out}}_{\text{log. regr.}} \quad (82)$$

where $\xrightarrow{\mathfrak{S}}$ is the partition map of the last layer into two or more disjointed components, induced by one or more **separator submanifolds**, defined in Definition 6.1 of [2], whose general constructive theory is instead presented in [4]. After separation, the final classification output is obtained with the probabilistic setup of the *logistic regression* or of its multi-component generalization, i.e., the *softmax*.

In the case of the Convolutional Cartan Neural Networks outlined in section 1.4 of [4], the hidden layers are chosen as $U_i/H_i = \mathcal{M}^{[r,q_i]}$, where

$$\mathcal{M}^{[r,q]} \equiv \frac{\text{SO}(r, r+q)}{\text{SO}(r) \times \text{SO}(r+q)} \quad (83)$$

Fixing r once for all, the hidden layers have to be regarded as the total spaces of Tits Satake vector bundles sharing the same base manifold, namely the Tits Satake submanifold $\mathcal{M}^{[r,1]}$:

$$\begin{aligned} \mathcal{M}^{[r,q]} &= \text{tot}[\mathcal{E}^{[r(q-1)]}] \\ \mathcal{E}^{[r(q-1)]} &\xrightarrow{\pi_{TS}} \mathcal{M}^{[r,1]} \end{aligned} \quad (84)$$

that have structural group

$$G_{\text{struc}} = \text{SO}(r) \times \text{SO}(q-1) \quad (85)$$

and having as standard fibre a vector space of dimension $r \times (q-1)$

$$F = \mathbb{V}^{(r|q-1)} \quad (86)$$

in the mentioned direct product representation of G_{struc} .

In any case, the important feature of the manifolds corresponding to all the inner layers of the network is what we already emphasized, namely that they are Cartan–Hadamard manifolds, hence diffeomorphic to $\mathbb{R}^n$ and each metrically equivalent to a suitable solvable Lie subgroup manifold $\mathcal{S} \subset U$. As such, all the U_i/H_i admit a uniquely defined distance function $d(u, v)$ between any two points $u, v \in U_i/H_i$ and such a distance function is the length of the unique geodesic arc starting at u and ending in v . The existence of a distance function is essential for all Machine Learning algorithms, and it is for this reason that the properties of the geodesics, their general construction for U_i/H_i symmetric manifolds, and the structure of the distance function were carefully analyzed in [1].

In the present section, we reconsider the problem of geodesics from the point of view of Hamiltonian mechanics. This is a necessary step in order to transform the space of geodesics into a symplectic manifold and introduce the thermodynamical structures discussed in Sections 2 and 3. We begin with a general framework by recasting the geodesic problem on a generic Riemannian manifold $(\mathcal{M}, g)$ into the form of a Hamiltonian dynamical system. Then we specialize such a general setup to the case of U/H non-compact symmetric spaces, metrically equivalent to appropriate solvable Lie groups, $\mathcal{S}_{U/H}$, and show how the general framework specializes to such a case, revealing additional relevant properties.

4.1. The Geodesic Dynamical System in General

Let $(\mathcal{M}, g)$ be a generic finite-dimensional Riemannian space $\dim_{\mathbb{R}}(\mathcal{M}) = d < \infty$ and let us consider the problem of deriving the second-order differential equations whose solutions are its geodesic curves:

$$\gamma : \mathbb{R} \longrightarrow \mathcal{M} \quad (87)$$

In each coordinate patch $x^\alpha = \{x^1, \dots, x^d\}$, every geodesic is described by d functions $x^\alpha(t)$, where $t \in \mathbb{R}$ is the affine parameter. As explained in [55] (volume 1, page 145), the well-known geodesic differential equations can be derived as the Euler–Lagrange equations of a mechanical system whose Lagrangian is the following:

$$\mathcal{L}(x, \dot{x}) = \frac{1}{2} g_{\alpha\beta}(x) \dot{x}^\alpha \dot{x}^\beta \quad (88)$$

having denoted by $\dot{x}^\alpha \equiv dx^\alpha/dt$ the generalized velocities and by $g_{\alpha\beta}(x)$ the metric tensor. We can easily convert the Lagrangian system (88) into a Hamiltonian one, introducing, as usual, the canonical momenta:

$$p_\alpha \equiv \frac{\partial \mathcal{L}}{\partial \dot{x}^\alpha} = g_{\alpha\beta}(x) \dot{x}^\beta \Rightarrow \dot{x}^\alpha = g^{\alpha\beta}(x) p_\beta \quad (89)$$

where $g^{\alpha\beta}(x)$ is the inverse metric, and defining the Hamiltonian as the Legendre transform of the Lagrangian:

$$\mathcal{H}(p, x) \equiv \dot{x}^\alpha p_\alpha - \mathcal{L}(x, \dot{x}) = \frac{1}{2} g^{\alpha\beta}(x) p_\alpha p_\beta \quad (90)$$

According to the definitions and conventions of Appendix A.7, we define the canonical symplectic manifold by introducing the $2d$ canonical coordinates

$$Z^\Lambda = \{p_\alpha, x^\beta\}; \quad \alpha, \beta = 1, \dots, d \quad (91)$$

and the symplectic 2-form

$$\omega \equiv \omega_{\Lambda\Sigma} dZ^\Lambda \wedge dZ^\Sigma; \quad \omega_{\Lambda\Sigma} = \frac{1}{2} \begin{pmatrix} \mathbf{0}_{d \times d} & \mathbf{1}_{d \times d} \\ -\mathbf{1}_{d \times d} & \mathbf{0}_{d \times d} \end{pmatrix} \quad (92)$$

which agrees with Definition A9, being closed, non-degenerate, and of maximal rank as in Equation (A35). In terms of canonical momenta and canonical coordinates, one has:

$$\omega = \sum_{\alpha=1}^d dp_\alpha \wedge dx^\alpha \quad (93)$$

and according to the Definition A9, the total space $\text{tot}[\mathcal{TM}]$ of the tangent bundle

$$\mathcal{TM} \xrightarrow{\pi} \mathcal{M} \quad (94)$$

to the Riemannian manifold $(\mathcal{M}, g)$ equipped with the 2-form ω , locally defined in each coordinate patch U , by the expression (93), becomes a symplectic $2d$ -dimensional manifold $(\text{tot}[\mathcal{TM}], \omega)$, namely, the phase-space of the geodesic dynamical system. Since ω is non-degenerate, the symplectic manifold $(\text{tot}[\mathcal{TM}], \omega)$ is also a Poisson manifold. Indeed, in every coordinate patch, it suffices to invert the matrix $\omega_{\Lambda\Sigma}$ so as to obtain the **Poissonian bivector**:

$$\pi^{\Lambda\Sigma} = \left(\omega^{-1}\right)^{\Lambda\Sigma} \quad (95)$$

and given any function $f(Y) \in \mathcal{C}^\infty(\text{tot}[\mathcal{TM}])$, according to Equation (A39), we can associate with it the corresponding Hamiltonian vector field:

$$X_f \equiv \pi^{\Lambda\Sigma} \partial_\Lambda f \partial_{\Sigma^i} \quad \partial_\Gamma \equiv \frac{\partial}{\partial Z^\Gamma} \quad (96)$$

and we obtain the Poisson bracket fulfilling the properties listed in its Definition A10 by setting:

$$\forall f, g \in \mathcal{C}^\infty(\text{tot}[\mathcal{TM}]) \quad : \quad \{f, g\} = \omega(X_f, X_g) \quad (97)$$

like in Equation (A40).

The standard geometric geodesic equation:

$$\ddot{x}^\alpha + \Gamma_{\beta\gamma}^\alpha(x) \dot{x}^\beta \dot{x}^\gamma = 0 \quad (98)$$

where

$$\Gamma_{\beta\gamma}^\alpha(x) = \frac{1}{2} g^{\alpha\mu} (\partial_\beta g_{\gamma\mu} + \partial_\gamma g_{\beta\mu} - \partial_\mu g_{\beta\gamma}) \quad (99)$$

denotes the Christoffel symbols, i.e., the components of the Levi Civita connection on $(\mathcal{M}, g)$, and is retrieved from the Hamiltonian equations:

$$\begin{aligned} \dot{x}^\alpha &= \{\mathcal{H}, x^\alpha\} = g^{\alpha\nu} p_\nu \\ \dot{p}_\alpha &= \{\mathcal{H}, p_\alpha\} = -\frac{1}{2} \partial_\alpha g^{\rho\sigma} p_\rho p_\sigma \end{aligned} \quad (100)$$

4.2. The Geodesic Dynamical System for Non-Compact Symmetric Spaces

In the case of non-compact symmetric spaces U/H , thanks to their metric equivalence with a suitable solvable group manifold $\mathcal{S}_{U/H}$, the structure of the geodesic dynamical system can be recast into a more algebraic and very convenient form. To this effect, we observe that the unique Einstein U -invariant metric on U/H can be written as:

$$ds^2 \equiv \mathbf{g} = \kappa_{AB} e^A \times e^B \quad (101)$$

where $\kappa_{AB} = \kappa_{BA}$ is a **constant symmetric matrix** and

$$e^A = e(\mathbf{Y})^A_\alpha dY^\alpha \quad (102)$$

are the **left-invariant Maurer–Cartan 1-forms** on the group manifold $\mathcal{S}_{U/H}$, satisfying the Maurer–Cartan equations of the solvable Lie algebra $\text{Solv}_{U/H}$:

$$de^A + \frac{1}{2} f_{BC}^A e^B \wedge e^C = 0 \quad (103)$$

having denoted by $\mathbf{Y}$ the solvable coordinates, namely the parameters of the solvable Lie group $\mathcal{S}_{U/H} \subset U$, and by f_{BC}^A , the solvable Lie algebra structure constants (see [1,2] for further details). Metric equivalence is nothing more than Equations (101) and (103). Indeed, let

$$\mathbf{t}_A = e(\mathbf{Y})^A_\alpha \frac{\partial}{\partial Y^\alpha} \quad (104)$$

be a basis of sections of the tangent bundle $\mathcal{T}(U/H)$ made of *left-invariant vector fields dual to the left-invariant one-forms* e^A :

$$e^A(\mathbf{t}_B) \equiv e(\mathbf{Y})^A_\alpha e(\mathbf{Y})^\alpha_B = \delta^A_B; \quad [\mathbf{t}_B, \mathbf{t}_C] = f_{BC}^A \mathbf{t}_A \quad (105)$$

The linear combinations with constant coefficients of the $\mathbf{t}_A$ vector fields constitute a vector space endowed with a Lie bracket that is the very definition of the Lie algebra $Solv_{U/H}$ of the solvable group $\mathcal{S}_{U/H}$ (see, for instance, [9]):

$$\mathbf{X} \in Solv_{U/H} \Leftrightarrow \mathbf{X} = X^A \mathbf{t}_A; \quad X^A \in \mathbb{R} \quad (106)$$

and we obtain:

$$\forall \mathbf{X}, \mathbf{Y} \in Solv_{U/H} : \quad \mathbf{g}(\mathbf{X}, \mathbf{Y}) = \kappa_{AB} X^A Y^B \equiv \kappa(\mathbf{X}, \mathbf{Y}) \quad (107)$$

In this way, we see that once reduced to the **left-invariant vector fields**, the metric is a scalar product on the solvable Lie Algebra, and the symmetric matrix κ_{AB} provides the coefficients of the corresponding symmetric quadratic form. Conversely, any quadratic form $k(,)$ on $Solv_{U/H}$ induces a metric on the solvable Lie group $\mathcal{S}_{U/H}$:

$$ds_k^2 = k_{AB} e^A \times e^B \quad (108)$$

that is $\mathcal{S}_{U/H}$ -invariant but not necessarily U -invariant, nor Einstein. So any positive definite quadratic bilinear $k(,)$ equips $\mathcal{S}_{U/H}$ with the structure of an *Alekseevskian normal Riemannian space* [2,12,13,56], yet there is only one quadratic form that corresponds to the unique Einstein metric (The Einstein metric is unique up to a homothety, namely up to a overall constant rescaling of the metric tensor or of the vielbein and such is the corresponding invariant quadratic form on the solvable Lie algebra) of the symmetric space U/H .

The structure of Equation (101) being clarified, we reconsider the generic form of the geodesic dynamical system Lagrangian (88), and we rewrite it as:

$$\mathcal{L}_{geoUH} = \frac{1}{2} \kappa_{AB} e(\mathbf{Y})^A_{\alpha} \dot{Y}^{\alpha} e(\mathbf{Y})^B_{\beta} \dot{Y}^{\beta} \quad (109)$$

Next, we introduce the *anholonomic Lagrangian velocities*. The anholonomic Lagrangian velocities are defined as the velocities in an anholonomic basis and, as such, should not be intended as time derivatives of coordinates. By an abuse of notation, we still denote them using an upper dot.

$$\dot{q}^A \equiv e(\mathbf{Y})^A_{\alpha} \dot{Y}^{\alpha} = i_{\partial_t} e^A \quad (110)$$

which are just the contraction of the Maurer–Cartan 1-forms with the time derivative. Similarly, we introduce the *anholonomic Hamiltonian momenta* as:

$$\mathfrak{p}_A \equiv \frac{\partial \mathcal{L}_{geoUH}}{\partial \dot{q}^A} = \kappa_{AB} \dot{q}^B \quad (111)$$

Then the Hamiltonian is defined as usual by the Legendre transform

$$\mathcal{H}_{geoUH} = \dot{q}^A \mathfrak{p}_A - \mathcal{L}_{geoUH} = \frac{1}{2} \kappa^{AB} \mathfrak{p}_A \mathfrak{p}_B \quad (112)$$

where, according to standard conventions, κ^{AB} is the inverse of the quadratic form matrix κ_{AB} :

$$\kappa^{AB} \kappa_{BC} = \delta^A_C \quad (113)$$

4.2.1. The Symplectic 2-Form

Next, we have to convert the standard symplectic form of Equation (93) to the new basis of Hamiltonian coordinates:

$$\Phi^\Lambda = \{p_A, Y^\alpha\} \quad (114)$$

We write the identification:

$$\omega = \omega_{\Lambda\Sigma}(Z) dZ^\Lambda \wedge dZ^\Sigma = \hat{\omega}_{\Lambda\Sigma}(\Phi) d\Phi^\Lambda \wedge d\Phi^\Sigma \quad (115)$$

so that our convention is that we name $\omega_{\Lambda\Sigma}$ the components of ω in the old Hamiltonian basis and $\hat{\omega}_{\Lambda\Sigma}$ the components of the same 2-form in the new Hamiltonian basis. The derivation of $\hat{\omega}_{\Lambda\Sigma}$ is the simple and direct calculation sketched below:

$$\begin{aligned} -\omega &= dY^\alpha \wedge dp_\alpha = dY^\alpha \wedge d[g_{\alpha\beta} \dot{Y}^\beta] = dY^\alpha \wedge d[e_\alpha^P(Y) \kappa_{PQ} e_\beta^Q(Y) \dot{Y}^\beta] \\ &= dY^\alpha \wedge dY^\mu \partial_\mu e_\alpha^P(Y) (\kappa_{PQ} e_\beta^Q(Y) \dot{Y}^\beta) + dY^\alpha e_\alpha^P(Y) \wedge d[\kappa_{PQ} e_\beta^Q(Y) \dot{Y}^\beta] \\ &= -de^A p_A + e^A \wedge dp_A \\ &= \frac{1}{2} f_{BC}^A e^B \wedge e^C + e^A \wedge dp_A \end{aligned} \quad (116)$$

Summarizing, we have the simple and elegant formula:

$$\omega = -\frac{1}{2} f_{BC}^A e^B \wedge e^C p_A - e^A \wedge dp_A \quad (117)$$

which is defined on the total space of the tangent bundle of the symmetric space, coinciding with the total space of the tangent bundle of the corresponding solvable Lie group manifold: $\mathcal{T}(U/H) \simeq \mathcal{T}(\mathcal{S}_{U/H})$. The 2-form ω is closed and of maximal rank:

$$\begin{aligned} d\omega &= 0 \\ \underbrace{\omega \wedge \omega \wedge \cdots \wedge \omega}_{d\text{-times}} &\neq 0; \quad d \equiv \dim_{\mathbb{R}} \left[\frac{U}{H} \right] = \dim_{\mathbb{R}} [\mathcal{S}_{U/H}] \end{aligned} \quad (118)$$

The first line in Equation (118) follows from the consistency of the Maurer–Cartan Equation (103) while the second is evident since we get

$$\omega \wedge \omega \wedge \cdots \wedge \omega = \underbrace{\text{const}}_{\neq 0} \times \underbrace{e^1 \wedge e^2 \wedge \cdots \wedge e^d \wedge dp_1 \wedge \cdots \wedge dp_d}_{\text{Vol}(\mathcal{M}_{2d})} \quad (119)$$

where by $\text{Vol}(\mathcal{M}_{2d})$ we have denoted the top $2d$ -form on the manifold $\mathcal{M}_{2d}$ defined as the total space of the tangent bundle $\mathcal{T}(\mathcal{S}_{U/H})$:

$$\mathcal{M}_{2d} = \text{tot}[\mathcal{T}(\mathcal{S}_{U/H})] \quad (120)$$

Hence, the pair $(\mathcal{M}_{2d}, \omega)$ as defined by Equations (117) and (120) is a bona-fide symplectic manifold as defined and illustrated in Appendix A.7 and such a statement is true for the solvable Lie group $\mathcal{S}_{U/H}$ singled out by any non-compact symmetric space U/H with U simple as thoroughly discussed in [1,2].

4.2.2. The Poissonian Bi-Vector

According to the general theory discussed in Appendix A.7 and partially already recalled above in Equations (95)–(97), every symplectic manifold is also a Poissonian manifold, although the reverse is not true. Indeed, the *Poissonian bivector* $\pi^{\Lambda\Sigma}$ can be obtained from the symplectic 2-form components as in Equation (95). Therefore, our next task is that of retrieving the Poissonian bivector starting from the symplectic 2-form in Equation (117). This is easily done. In the coordinate basis (114) the $2d \times 2d$ matrix $\hat{\omega}_{\Lambda\Sigma}$ has the following structure:

$$\begin{aligned}\hat{\omega}_{\Lambda\Sigma} &= \left(\begin{array}{c|c} \mathbf{0}_{d \times d} & \hat{\omega}^M_\beta \\ \hline \hat{\omega}^N_\alpha & \hat{\omega}_{\alpha\beta} \end{array} \right) \\ \hat{\omega}_{\alpha\beta} &= -\frac{1}{2} f_{BC}{}^A \mathfrak{p}_A e^B_\alpha e^C_\beta; \quad \hat{\omega}^M_\beta = \frac{1}{2} e^M_\beta; \quad \hat{\omega}^N_\alpha = -\frac{1}{2} e^N_\alpha\end{aligned}\quad (121)$$

The Poissonian bivector is an antisymmetric matrix $\pi^{\Lambda\Sigma}$ such that:

$$\omega_{\Lambda\Sigma} \pi^{\Sigma\Delta} = \delta^\Delta_\Lambda \quad (122)$$

We immediately find

$$\begin{aligned}\pi_{\Lambda\Sigma} &= \left(\begin{array}{c|c} \pi_{MN} & \pi_M{}^\beta \\ \hline \pi^\alpha{}_N & \mathbf{0}_{d \times d} \end{array} \right) \\ \pi^{MN} &= -2 f_{MN}{}^A \mathfrak{p}_A; \quad \pi_M{}^\beta = -2 e^\beta_M; \quad \pi^\alpha{}_N = 2 e^\alpha_N\end{aligned}\quad (123)$$

where e^N_α are the components of the **left-invariant Maurer–Cartan forms** e^M and e^β_M , the components of their dual **left-invariant vector fields** on the solvable Lie group manifold:

$$e^A = e^A_\alpha dY^\alpha; \quad \mathbf{t}_B = e^\beta_B \frac{\partial}{\partial Y^\beta}; \quad e^A(\mathbf{t}_B) = e^A_\alpha e^\alpha_B = \delta^A_B \quad (124)$$

that generate right-translations.

4.2.3. Hamiltonian Vector Fields and the Poisson Bracket

Having determined the Poissonian bivector, we can write the explicit form of the Hamiltonian vector field associated with any function $\varphi(\Phi) \in \mathbb{C}^\infty(\text{tot}[\mathcal{T}(\mathcal{S}_{U/H})])$. Recalling Equation (96), we set:

$$\begin{aligned}\varphi(\Phi) \rightarrow \mathbf{X}_\varphi &= \pi^{\Lambda\Sigma} \partial_\Lambda \varphi \frac{\partial}{\partial \Phi^\Sigma} \\ &= -2 \mathfrak{p}_A f_{BC}{}^A \frac{\partial \varphi}{\partial \mathfrak{p}_B} \frac{\partial}{\partial \mathfrak{p}_C} - 2 \frac{\partial \varphi}{\partial \mathfrak{p}_M} \mathbf{t}_M + 2 \mathbf{t}_M \varphi \frac{\partial}{\partial \mathfrak{p}_M}\end{aligned}\quad (125)$$

and we define the Poisson bracket as:

$$\begin{aligned}\forall \varphi(\Phi), \psi(\Phi) \in \mathbb{C}^\infty(\text{tot}[\mathcal{T}(\mathcal{S}_{U/H})]) \quad : \quad \{\varphi, \psi\} &\equiv \omega(\mathbf{X}_\varphi, \mathbf{X}_\psi) \\ &= -2 \left(\mathfrak{p}_A f_{BC}{}^A \frac{\partial \varphi}{\partial \mathfrak{p}_B} \frac{\partial \psi}{\partial \mathfrak{p}_C} + \frac{\partial \varphi}{\partial \mathfrak{p}_M} \mathbf{t}_M \psi - \mathbf{t}_M \varphi \frac{\partial \psi}{\partial \mathfrak{p}_M} \right)\end{aligned}\quad (126)$$

4.2.4. Symplectic Moment Map

Given a vector field $\mathbf{k} \in \Gamma[\mathcal{T}\mathcal{M}, \mathcal{M}]$, namely a section of the tangent bundle to (Here and in the following lines we are always talking about the solvable Lie group $\mathcal{S}$ metrically equivalent to the symmetric space U/H , and for notation simplicity we drop the subscript U/H).

$$\mathcal{M} = \text{tot}[\mathcal{T}\mathcal{S}] \quad (127)$$

which has the symplectic structure specified by the 2-form (117); we define the moment map:

$$\mu \quad : \quad \mathbf{k} \longrightarrow \mu_{\mathbf{k}}(\Phi) \in \mathbb{C}^\infty(\mathcal{M}) \quad (128)$$

by imposing the condition:

$$\forall f(\Phi) \in \mathbb{C}^\infty \quad : \quad \mathbf{k}f = \{\mu_{\mathbf{k}}, f\} = \omega(\mathbf{k}, \mathbf{X}_f) \quad (129)$$

where $\mathbf{X}_f$ is the Hamiltonian vector field associated with the function f (see Equation (125)). Hence, the moment map $\mu_{\mathbf{k}}$ is a solution to the following differential equation:

$$\mathbf{k} = -2f_{BC}{}^A \mathfrak{p}_A \frac{\partial \mu_{\mathbf{k}}}{\partial \mathfrak{p}_B} \frac{\partial}{\partial \mathfrak{p}_C} - 2 \frac{\partial \mu_{\mathbf{k}}}{\partial \mathfrak{p}_M} \mathbf{t}_M + 2 \mathbf{t}_M \mu_{\mathbf{k}} \frac{\partial}{\partial \mathfrak{p}_M} \quad (130)$$

Consider, in particular, the vector fields

$$\mathbf{k}_N \equiv \mathbf{t}_N + \mathbf{v}_N \quad (131)$$

where $\mathbf{t}_N$ are the purely **horizontal**, right-invariant vector fields defined over the solvable group manifold $\mathcal{S}$ that generate its solvable Lie algebra $Solv$ since

$$[\mathbf{t}_N, \mathbf{t}_R] = f_{NR}{}^A \mathbf{t}_R \quad (132)$$

while

$$\mathbf{v}_N \equiv f_{NB}{}^A \mathfrak{p}_A \frac{\partial}{\partial \mathfrak{p}_B} \quad (133)$$

are purely **vertical** vector fields that, as a consequence of the Jacobi identities, satisfy the commutation relations of $Solv$:

$$[\mathbf{v}_N, \mathbf{v}_R] = f_{NR}{}^A \mathbf{v}_R \quad (134)$$

and commute with the horizontal partners:

$$[\mathbf{v}_N, \mathbf{t}_R] = 0 \quad (135)$$

Hence, the vector fields $\mathbf{k}_N$ also satisfy the solvable Lie algebra commutation relations:

$$[\mathbf{k}_N, \mathbf{k}_R] = f_{NR}{}^A \mathbf{k}_R \quad (136)$$

and are the infinitesimal generators for the action of the solvable Lie group $\mathcal{S}$ on the total space of its tangent bundle, namely the phase-space of the geodesic dynamical system. We claim that for the vector fields $\mathbf{k}_N$, the appropriate moment map is the following

$$\mu_N \equiv \mu_{\mathbf{t}_N} = -\frac{1}{2} \mathfrak{p}_N \quad (137)$$

4.2.5. Relation with the Nomizu Operator

Once the metric form is given, the construction of geometry and of the associated geodesic equations follows uniquely. The issue is just that of calculating the Levi-Civita connection of the metric g induced on the manifold by the form $\langle, \rangle$ defined on the solvable Lie algebra $Solv$. One way of describing this Levi-Civita connection is by means of the so called *Nomizu operator* acting on $Solv$. The latter is defined as follows:

$$\begin{aligned} \mathbb{L} &: Solv \otimes Solv \rightarrow Solv, \\ \forall X, Y, Z \in Solv &: 2 \langle \mathbb{L}_X Y, Z \rangle = \langle [X, Y], Z \rangle - \langle X, [Y, Z] \rangle - \langle Y, [X, Z] \rangle \end{aligned} \quad (138)$$

The *Riemann curvature operator* on $Solv$ can be expressed as

$$\text{Riem}(X, Y) = [\mathbb{L}_X, \mathbb{L}_Y] - \mathbb{L}_{[X, Y]} \quad (139)$$

If we introduce a basis of abstract generators $\{T_A\}$ for $Solv$ and the corresponding structure constants

$$[T_A, T_B] = f_{AB}{}^C T_C \quad (140)$$

that are the same as those appearing in Equations (103) and (105) together with the metric tensor

$$\langle T_A, T_B \rangle = \kappa_{AB} \quad (141)$$

the connection defined by Equation (138) leads to the following connection coefficients:

$$\begin{aligned} \mathbb{L}_A T_B &= \Gamma_{AB}^C T_C \\ \Gamma_{AB}^C &= f_{AB}{}^C - \kappa_{AD} \kappa^{CE} f_{BE}{}^D - \kappa_{BD} \kappa^{CE} f_{AE}{}^D \end{aligned} \quad (142)$$

which are constant numbers.

Given the connection coefficients, the differential geodesic equations can be written immediately. In the chosen basis, the tangent vector to the geodesic is described by n fields

$$\Pi^A(t) \equiv \kappa^{AB} \mathfrak{p}_B(t) \quad (143)$$

which depend on the affine parameter t along the curve. The geodesic equation is given by the following first-order differential system:

$$\frac{d}{dt} \Pi^A + \Gamma_{BC}^A \Pi^B \Pi^C = 0 \quad (144)$$

The above equation contains two pieces of data:

- (1) the structure constants of the solvable Lie algebra $f_{AB}{}^C$;
- (2) the constant tensor κ_{AB} defining the norm on the solvable Lie algebra.

Equation (144) can be obtained as Hamiltonian equations from the definition (112) of the Hamiltonian $\mathcal{H}_{geoUH}$ and the explicit expression of the Poisson bracket (126):

$$\partial_t \Pi^A = \left\{ \mathcal{H}_{geoUH}, \Pi^A \right\} \quad (145)$$

$$\partial_t \dot{Y}^\alpha = \left\{ \mathcal{H}_{geoUH}, Y^\alpha \right\} = e_M^\alpha \Pi^M \quad (146)$$

Note that the first Equation (145) yielding Equation (144) was obtained in [18] using the definition (126) of Poisson brackets reduced to functions only of the momenta $\mathfrak{p}_A$, namely:

$$\{\varphi(\mathfrak{p}), \psi(\mathfrak{p})\}_{red} = -2 \mathfrak{p}_A f_{BC}{}^A \frac{\partial \varphi}{\partial \mathfrak{p}_B} \frac{\partial \psi}{\partial \mathfrak{p}_C} \quad (147)$$

It was observed in [18] that Equation (147) equips the dual vector space $Solv^*$ to any solvable Lie algebra $Solv$ with the structure of a Poisson manifold, which is not a symplectic manifold since the bivector provided by the solvable structure constants $f_{BC}{}^A$ is not invertible. Yet, as recalled in [18], when the solvable Lie algebra is the Borel subalgebra of the special linear group

$$Solv = \mathbb{B}_N \equiv \mathbb{B}[\mathfrak{sl}(N, \mathbb{R})] \quad (148)$$

namely the space of $N \times N$ triangular traceless matrices, it was proved by Arkhangel'skii in [57] that the Hamiltonian system based on the Poisson bracket (147) is always Liouville integrable since it possesses the required number of Hamiltonians in involution. Specifically, distinguishing the even case $\mathbb{B}_{2\nu}$ from the odd one $\mathbb{B}_{2\nu+1}$ (where $\nu \in \mathbb{N}$), we have that for $\mathbb{B}_{2\nu}$ there are $\nu^2 + \nu - 1$ Hamiltonians in involution, while for $\mathbb{B}_{2\nu+1}$ the number of such objects is $\nu^2 + 2\nu$. Furthermore, in both cases, among the Hamiltonian in involution there

are respectively $r = \nu - 1$ and $r = \nu$ Casimirs $C_i(\mathbf{p})$ ($i = 1, \dots, r$), namely such functions of the momenta $\mathbf{p}$ that have vanishing Poisson bracket with all of them:

$$\forall i, A \quad : \quad \{\mathbf{p}_A, C_i(\mathbf{p})\} = 0 \quad (149)$$

According to the general theory of dynamical systems, one can define level surfaces of the Casimirs:

$$\mathcal{L}_{k_1, \dots, k_r} \subset \text{Sol}\nu^*; \quad \forall \mathbf{p} \in \mathcal{L}_{k_1, \dots, k_r} \quad : \quad C_i(\mathbf{p}) = k_i = \text{const} \in \mathbb{R} \quad (150)$$

and, as explained in [18], such level surfaces always have an even dimension and become symplectic manifolds since when restricted to them the Poissonian bi-vector becomes invertible.

Although this is very interesting, as one sees from the above discussion, it is only one aspect of the full story. Indeed, the complete space that includes not only the momenta $\mathbf{p}_A$ but also the coordinates Y^α is always symplectic, and the full definition of the Poisson bracket is that given in Equation (126). Yet the nice point about Arkhangel'skii Hamiltonians is that those functions of the momenta $\mathbf{p}$ that are in involution with respect to the reduced Poisson bracket (147) remain in involution also with respect to the full Poisson bracket (126), so that they are constant along any geodesic.

5. A Master Example for the Geodesic Dynamical System: $\text{SL}(3, \mathbb{R})/\text{SO}(3)$

As a concrete illustration of the above-discussed concepts and constructions we choose the 5-dimensional symmetric space:

$$\mathcal{M}_5 \equiv \frac{\text{SL}(3, \mathbb{R})}{\text{SO}(3)} \quad (151)$$

The reasons for such a choice are several:

1. $\mathcal{M}_5$ is the smallest symmetric space with a non-compact rank $r > 1$ and a non-trivial solvable Lie algebra $\text{Sol}\nu_5$.
2. $\mathcal{M}_5$ belongs to the mother series of non-compact symmetric spaces $\frac{\text{SL}(N, \mathbb{R})}{\text{SO}(N)}$ that, thanks to the triangular embedding, explained in [1,2], contains all the members of the other series as submanifolds.
3. $\mathcal{M}_5$ is not a Kähler manifold, yet its 10-dimensional tangent bundle $\mathcal{T}\mathcal{M}_5$ has the symplectic structure discussed in Sections 4.2 and 5 as any other tangent bundle. This allows to illustrate the distinction among the symplectic manifold of the GDS utilized here with respect to what is done in paper [32], where, as we extensively stressed in the introduction, the moment maps and the thermodynamical states are defined with respect to the Souriau symplectic 2-form (25) constructed on coadjoint orbits and also with respect to what was done in the above mentioned papers [18,57], where the Poisson structure is defined only on the standard fibre of the tangent bundle to $\mathcal{T}\mathcal{U}/\mathcal{H}$. As we stressed in the introduction, the Souriau case corresponds to thermodynamics on Kähler manifolds.

The Solvable Lie Algebra Generators

An explicit basis for the solvable Lie algebra $\text{Sol}\nu_5$ of traceless, upper triangular matrices in 3-dimension is the following one:

$$\begin{aligned}
T_1 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}; \quad T_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}; \quad T_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; \quad T_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\
T_5 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix};
\end{aligned} \tag{152}$$

where the diagonal $T_{1,2}$ are Cartan generators, while $T_{3,4}$ correspond to the simple roots $\alpha_{1,2}$ and T_5 is associated with the highest root $\alpha_1 + \alpha_2$.

The solvable Lie group generic element

Next, it is very simple to write the generic element of the solvable Lie group $S_5 \cong \exp[\text{Sol}v_5]$ obtained according to the rules of the exponential map Σ as defined in [1,2].

$$S_5 \ni \mathbb{L}(\mathbf{Y}) \equiv \Sigma[Y^A T_A] = \prod_{A=1}^5 \exp[Y^A T_A] = \begin{pmatrix} e^{Y^1} & e^{Y^1 Y^3} & e^{Y^1 (Y^3 Y^4 + Y^5)} \\ 0 & e^{Y^2} & e^{Y^2 Y^4} \\ 0 & 0 & e^{-Y^1 - Y^2} \end{pmatrix} \tag{153}$$

The Left Invariant Cartan 1-form Matrix on S_5

Given the generic solvable Lie group element $\mathbb{L}(\mathbf{Y})$, we easily compute the solvable Lie algebra valued left-invariant 1-form Θ :

$$\begin{aligned}
\Theta &\equiv \mathbb{L}^{-1}(\mathbf{Y}) \cdot d\mathbb{L}(\mathbf{Y}) \\
&= \begin{pmatrix} dY^1 & dY^3 + Y^3(dY^1 - dY^2) & dY^5 + Y^4 dY^3 + Y^3 Y^4 dY^1 - Y^3 Y^4 dY^2 + 2Y^5 dY^1 + Y^5 dY^2 \\ 0 & dY^2 & dY^4 + Y^4(dY^1 + 2dY^2) \\ 0 & 0 & -dY^1 - dY^2 \end{pmatrix}
\end{aligned} \tag{154}$$

The Maurer–Cartan Forms 1-forms on S_5

The left-invariant Maurer–Cartan forms e^A are obtained by decomposing the upper triangular traceless matrix Θ , whose matrix elements are 1-forms on the solvable Lie group manifold S_5 , along the generator basis (A2):

$$\begin{aligned}
\Theta &= e^A T_A \\
e^1 &= dY^1 \\
e^2 &= dY^2 \\
e^3 &= dY^3 + Y^3(dY^1 - dY^2) \\
e^4 &= dY^4 + Y^4(dY^1 + 2dY^2) \\
e^5 &= dY^5 + Y^4 dY^3 + Y^3 Y^4 dY^1 - Y^3 Y^4 dY^2 + 2Y^5 dY^1 + Y^5 dY^2
\end{aligned} \tag{155}$$

By construction, the Maurer–Cartan forms e^A satisfy the Maurer–Cartan equations:

$$de^A + \frac{1}{2} f_{BC}^A e^B \wedge e^C = 0 \tag{156}$$

where $f_{BC}{}^A$ are the structure constants of the solvable Lie algebra $Solv_5$:

$$[T_B, T_C] = f_{BC}{}^A T_A \quad (157)$$

The vielbein of the unique $SL(3, \mathbb{R})$ invariant Einstein metric on the symmetric space $\mathcal{M}_5$

As explained above and in [1], the unique Einstein U-invariant metric on the symmetric manifold is algebraically obtained from an orthonormal basis of $\mathbb{K}$ generators in the Cartan decomposition of the U-Lie algebra: $\mathbb{U} = \mathbb{H} \oplus \mathbb{K}$, where $\mathbb{K}$ constitutes an irreducible representation of the maximal compact subalgebra $\mathbb{H}$ under its adjoint action. In our case $\mathbb{H} = \mathfrak{so}(3)$ and the 5-dimensional vector space $\mathbb{K}$ corresponds to the $J = 2$ irrep of the three-dimensional rotation group. An orthonormal basis is provided by the following symmetric matrices:

$$\begin{aligned} K^1 &= \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}; \quad K^2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} & 0 & 0 \\ 0 & \sqrt{\frac{2}{3}} & 0 \\ 0 & 0 & -\frac{1}{\sqrt{6}} \end{pmatrix}; \quad K^3 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; \quad K^4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \\ K^5 &= \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix}; \end{aligned} \quad (158)$$

that satisfy the following relation:

$$\text{Tr}(K^I \cdot K^J) = \delta^{IJ}; \quad I, J = 1, 2, \dots, 5 \quad (159)$$

The vielbein determining the metric is obtained from the Cartan left invariant matrix 1-form (154) by setting

$$V^I \equiv \text{Tr}(K^I \cdot \Theta) \quad (160)$$

and the metric is

$$\begin{aligned} ds^2 &\equiv \delta_{IJ} V^I \times V^J = g_{\alpha\beta}(\mathbf{Y}) dY_\alpha dY_\beta \\ &= \frac{1}{2} \left\{ 3dY^2 + (2dY^1 + dY^2)^2 + (dY^3 + Y^3(dY^1 - dY^2))^2 + (dY^4 + Y^4(dY^1 + 2dY^2))^2 \right. \\ &\quad \left. + (dY^5 + Y^4 dY^3 + Y^3 Y^4 dY^1 - Y^3 Y^4 dY^2 + 2Y^5 dY^1 + Y^5 dY^2)^2 \right\} \end{aligned} \quad (161)$$

The vielbein 1-forms V^I are linear combinations of the Maurer–Cartan 1-forms e^A :

$$V^I = v_A^I e^A; \quad v = \begin{pmatrix} \sqrt{2} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & \sqrt{\frac{3}{2}} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \quad (162)$$

So that the metric can also be written as:

$$ds^2 = \kappa_{AB} e^A \times e^B$$

$$\kappa \equiv \nu^T \cdot \nu = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \quad (163)$$

5.1. Hamiltonians in Involution and Generalized Thermodynamics

As we stressed in Section 4.2.5, the general definition (126) of the Poisson bracket on the total space of the tangent bundle can be reduced to functions $f(\mathbf{p})$ that depend only on the momenta, namely on the vertical fibre coordinates, and one obtains the reduced formula (147) which still satisfies all the properties mentioned in Definition A10 of Appendix A.7 in order to define a Poisson bracket. In the case of the here considered master model, Equation (147) takes the following explicit form for any pair of functions $F(\mathbf{p}), G(\mathbf{p})$:

$$\begin{aligned} \{F(\mathbf{p}), G(\mathbf{p})\}_{red} = & -2\mathbf{p}_3(-\partial_3 F \partial_1 G + \partial_3 F \partial_2 G + \partial_1 F \partial_3 G - \partial_2 F \partial_3 G) \\ & -2\mathbf{p}_4(-\partial_4 F \partial_1 G - 2\partial_4 F \partial_2 G + \partial_1 F \partial_4 G + 2\partial_2 F \partial_4 G) \\ & -2\mathbf{p}_5(-2\partial_5 F \partial_1 G - \partial_5 F \partial_2 G - \partial_4 F \partial_3 G + \partial_3 F \partial_4 G + 2\partial_1 F \partial_5 G + \partial_2 F \partial_5 G) \end{aligned} \quad (164)$$

where we have used the shorthand notation:

$$\partial_a f \equiv \frac{\partial f(\mathbf{p})}{\partial \mathbf{p}_a} \quad (165)$$

Utilizing the constructive recipe introduced by Arkhangel'skii in [57] and recalled in Equation (2.49) of [18]), we obtain the following three Hamiltonians in involution:

$$\mathfrak{H}_1 = \frac{1}{3}(\mathbf{p}_1^2 - \mathbf{p}_2 \mathbf{p}_1 + \mathbf{p}_2^2 + 3(\mathbf{p}_3^2 + \mathbf{p}_4^2 + \mathbf{p}_5^2)) \quad (166)$$

$$\mathfrak{H}_2 = \frac{1}{27}(-2\mathbf{p}_1^3 + 3\mathbf{p}_2 \mathbf{p}_1^2 + 3(\mathbf{p}_2^2 - 3(\mathbf{p}_3^2 - 2\mathbf{p}_4^2 + \mathbf{p}_5^2)))\mathbf{p}_1 - 2\mathbf{p}_2^3 - 54\mathbf{p}_3 \mathbf{p}_4 \mathbf{p}_5 - 9\mathbf{p}_2(\mathbf{p}_3^2 + \mathbf{p}_4^2 - 2\mathbf{p}_5^2) \quad (167)$$

$$\mathfrak{H}_3 = \frac{1}{3}\left(\mathbf{p}_1 - 2\mathbf{p}_2 + \frac{3\mathbf{p}_3 \mathbf{p}_4}{\mathbf{p}_5}\right) \quad (168)$$

$$0 = \{\mathfrak{H}_i, \mathfrak{H}_j\}_{red} \quad i, j = 1, 2, 3 \quad (169)$$

The first crucial observation is that the quadratic first Hamiltonian $\mathfrak{H}_1 = \mathcal{H}_{geoUH}$ coincides with the quadratic Hamiltonian (112) obtained from the Legendre transform of the geodesic dynamical system Lagrangian. Indeed, in our case, we have that κ is given by the second of Equation (163), and one immediately verifies that

$$\mathfrak{H}_1 = \frac{1}{2}(\kappa^{-1})^{AB} \mathbf{p}_A \mathbf{p}_B \equiv \frac{1}{2}\kappa^{AB} \mathbf{p}_A \mathbf{p}_B = \mathcal{H}_{geoUH}$$

$$\kappa^{-1} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & 0 & 0 & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \quad (170)$$

This implies that the method of Hamiltonians in involution selects the unique invariant norm on the solvable Lie algebra that corresponds to the unique U-symmetric Einstein

metric on the equivalent symmetric space $SL(N, \mathbb{R})/SO(N)$. This fact was already pointed out in [18].

The second important observation is that the functions depending only on the momenta $\mathbf{p}_A$ that are in involution with respect to the reduced Poisson bracket are in involution with respect to the complete Poisson bracket, so that they are conserved along all trajectories, namely all geodesic lines.

The case of the Casimir $\mathfrak{H}_3$ is different. It commutes with all the momenta, yet it does not commute with all the coordinates Y^α . Indeed, we have:

$$\{\mathfrak{H}_3, \mathbf{p}_A\} = 0; \quad \{\mathfrak{H}_3, Y^\alpha\} = \frac{\partial \mathfrak{H}_3}{\partial \mathbf{p}_A} e_A^\alpha(\mathbf{Y}) \quad (171)$$

5.2. Generalized Thermodynamics for a Geodesic Dynamical System on U/H

As we see from Equation (171), there is no way of fixing consistently the Hamiltonian $\mathfrak{H}_3$ to a constant value on the whole ten-dimensional space with canonical coordinates $\{\mathbf{p}_A, Y^\alpha\}$, obtaining in this way a reduced symplectic manifold of dimension eight. This conclusion is illustrated explicitly in the present master example, yet it is true for all manifolds $SL(N, \mathbb{R})/SO(N)$ since in all cases the Casimirs $\mathcal{C}_i(\mathbf{p})$ that have a vanishing reduced Poisson bracket with the momenta $\mathbf{p}$, have, with the manifold coordinates Y^α , a nonvanishing Poisson bracket:

$$\{\mathcal{C}_i(\mathbf{p}), Y^\alpha\} = \frac{\partial \mathcal{C}_i(\mathbf{p})}{\partial \mathbf{p}_A} e_A^\alpha(\mathbf{Y}) \quad (172)$$

Nevertheless, one can introduce, as in the classical statistical mechanics of free gases, a Gibbs state probability distribution that minimizes the Shannon functional as we discussed in Section 3 and we announced in Section 1.5. Indeed, given the set of conserved Hamiltonians in involution that depend only on the momenta $\mathbf{p}$ and that we denote $\mathfrak{H}_i(\mathbf{p})$, we can construct the Gibbs state probability distribution defined by Equations (26) and (27). For any symmetric space U/H of non-compact type, we have:

$$G(\lambda, V) = \frac{\exp[-\lambda \cdot \mathfrak{H}(\mathbf{p})]}{Z(\lambda, V)} \quad (173)$$

$$Z(\lambda, V) = \int_{\mathcal{M}_{2d}} \exp[-\lambda \cdot \mathfrak{H}(\mathbf{p})] d\lambda(\mathbf{p}, \mathbf{Y}) \quad (174)$$

$$d\lambda(\mathbf{p}, \mathbf{Y}) = \underbrace{e^1 \wedge e^2 \wedge \dots \wedge e^d \wedge d\mathbf{p}_1 \wedge \dots \wedge d\mathbf{p}_d}_{\text{Vol}(\mathcal{M}_{2d})} \quad (175)$$

where $\mathcal{M}_{2d}$ is the total space of the tangent bundle as specified in Equation (120) and $d\lambda(\mathbf{p}, \mathbf{Y})$ is the Liouville integration measure on such a space presented in Equation (119). Since all the Hamiltonians depend only on the momenta $\mathbf{p}_A$ and not on the coordinates Y^α , we are in a situation similar to that of Ideal Gases (see Appendix C.4), and the partition function (174) factorizes into the product of two integrals (or summations):

$$Z(\lambda, V) = \underbrace{\int_{\mathcal{T}} \exp[-\lambda \cdot \mathfrak{H}(\mathbf{p})] d\mathbf{p}_1 \wedge \dots \wedge d\mathbf{p}_d}_{\zeta(\lambda)} \times \underbrace{\int_{\text{Box} \subset \mathcal{S}} e^1 \wedge e^2 \wedge \dots \wedge e^d}_{V=\text{volume}} \quad (176)$$

Discussion on the Box Conception

In relation to the factorization in Equation (176), we have to pause for a moment and analyze the comparison with Ideal Gas Statistical Mechanics reviewed in Appendix C.4. In both cases, the basic Hamiltonian is a quadratic form in the momenta, and the integral over $\mathbf{p}$ is a multiple-Gaussian integral, the multiplicity being the number of degrees of freedom

n_f . In the Ideal Gas case, this number is $n_f = 3 \times N$ where N is the number of molecules composing the gas, namely of the order of the Avogadro number. In the geodesic dynamical system case, the number of degrees of freedom is $n_f = d$, namely a figure of the dimension of the symmetric space U/H or, if you prefer, of the corresponding solvable Lie group $S_{U/H}$. In the Ideal Gas case, the integral V is the volume of the portion of physical space $\mathbb{R}^3$ in which the sample of gas is confined, and the integration on it occurs N times, one for each molecule composing the gas. Hence, the factor that multiplies $\zeta(\lambda)$ is V^N . For the gases, the Gibbs state distribution expresses the probability that the N -particles are in a state where each of them has a given momentum, and it is at a given point. The independence of the Hamiltonian from the coordinates implies that such probability is insensitive to the position, namely, all positions have the same probability, while the momenta follow a Gaussian distribution. In the case of the geodesic dynamical system, the Gibbs state describes the probability that a geodesic starts at a given point with a given initial tangent vector (the momentum). The coordinate independence of the Hamiltonians implies that all points of the base manifold have the same probability, while the initial tangent vectors follow, as far as we consider only the quadratic Hamiltonian, a Gaussian distribution. As more Hamiltonians are introduced, the distribution becomes more complicated and more structured in momentum space, yet remains flat in coordinate space. In both cases, this flatness is the consequence of translation invariance, with respect to $\mathbb{R}^3$ in the Ideal Gas case, with respect to the solvable Lie group $S_{U/H}$ in the case of the geodesic dynamical system on U/H . Just as in the Ideal Gas case the *Box* is the portion of physical space in which the gas is confined, in the same way for the Geodesic Dynamical System case the *Box* is the portion of base manifold to which we confine the possible initial points of the geodesics of our interest: indeed the statistical distribution described by our thermodynamics is a statistical distribution in the space of geodesics.

5.2.1. Generalized Thermodynamics for the Chosen Master Example

In order to study the generalized thermodynamics of our master example, where the conserved Hamiltonians are given by Equations (166)–(168), we have to calculate the nontrivial part of the partition function:

$$\zeta(\lambda) = \int \exp \left[- \sum_{i=1}^3 \lambda_i \mathfrak{H}_i(\mathfrak{p}) \right] d^5 \mathfrak{p} \quad (177)$$

To this effect, it is convenient to diagonalize the quadratic form in $\mathfrak{p}_A$ that corresponds to the first Hamiltonian. Since we have

$$- \lambda_1 \mathfrak{H}_1 = \lambda_1 \underbrace{\left(-\frac{1}{2} \kappa^{-1} \right)^{AB}}_{= k^{AB}} \mathfrak{p}_A \mathfrak{p}_B \quad (178)$$

we need to diagonalize the above-defined matrix k . An easy calculation shows that the following matrix:

$$\mathcal{U} \equiv \begin{pmatrix} 0 & 0 & 0 & -1 & \sqrt{3} \\ 0 & 0 & 0 & 1 & \sqrt{3} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (179)$$

satisfies the condition:

$$\mathcal{U}^T \cdot \mathbf{k} \cdot \mathcal{U} = -\mathbf{1}_{5 \times 5} \quad (180)$$

Hence, defining new variables

$$\mathbf{w} \equiv \mathcal{U} \mathbf{p} \quad (181)$$

we obtain

$$-\mathbf{w}^T \mathbf{w} = -\frac{1}{2} \mathfrak{H}_1(\mathbf{p}) \quad (182)$$

Explicitly, the change of variables is as follows

$$\begin{aligned} \mathfrak{w}_1 &= \mathfrak{p}_5 & \mathfrak{p}_1 &= \sqrt{3}\mathfrak{w}_5 - \mathfrak{w}_4 \\ \mathfrak{w}_2 &= \mathfrak{p}_4 & \mathfrak{p}_2 &= \mathfrak{w}_4 + \sqrt{3}\mathfrak{w}_5 \\ \mathfrak{w}_3 &= \mathfrak{p}_3 & \mathfrak{p}_3 &= \mathfrak{w}_3 \\ \mathfrak{w}_4 &= \frac{1}{2}(\mathfrak{p}_2 - \mathfrak{p}_1) & \mathfrak{p}_4 &= \mathfrak{w}_2 \\ \mathfrak{w}_5 &= \frac{\mathfrak{p}_1 + \mathfrak{p}_2}{2\sqrt{3}} & \mathfrak{p}_5 &= \mathfrak{w}_1 \end{aligned} \quad (183)$$

In terms of new variables $\mathbf{w}$, the three Hamiltonians have the following expression:

$$\begin{aligned} \mathfrak{H}_1(\mathbf{w}) &= \mathfrak{w}_1^2 + \mathfrak{w}_2^2 + \mathfrak{w}_3^2 + \mathfrak{w}_4^2 + \mathfrak{w}_5^2 \\ \mathfrak{H}_2(\mathbf{w}) &= \left(\mathfrak{w}_4 + \frac{\mathfrak{w}_5}{\sqrt{3}} \right) \mathfrak{w}_1^2 - 2\mathfrak{w}_2\mathfrak{w}_3\mathfrak{w}_1 + \frac{1}{9} \left(\left(3\sqrt{3}\mathfrak{w}_5 - 9\mathfrak{w}_4 \right) \mathfrak{w}_2^2 + 2\sqrt{3}\mathfrak{w}_5 \left(\mathfrak{w}_5^2 - 3 \left(\mathfrak{w}_3^2 + \mathfrak{w}_4^2 \right) \right) \right) \\ \mathfrak{H}_3(\mathbf{w}) &= \frac{\mathfrak{w}_2\mathfrak{w}_3}{\mathfrak{w}_1} - \mathfrak{w}_4 - \frac{\mathfrak{w}_5}{\sqrt{3}} \end{aligned} \quad (184)$$

Hence, the partition function reduces to the following multiple integral

$$\zeta(\lambda_1, \lambda_2, \lambda_3) = \int_{-\infty}^{\infty} d\mathfrak{w}_1 \int_{-\infty}^{\infty} d\mathfrak{w}_2 \int_{-\infty}^{\infty} d\mathfrak{w}_3 \int_{-\infty}^{\infty} d\mathfrak{w}_4 \int_{-\infty}^{\infty} d\mathfrak{w}_5 \exp \left[- \sum_{i=1}^3 \lambda_i \mathfrak{H}_i(\mathbf{w}) \right] \quad (185)$$

Although we have not explored all the possible strategies to calculate the integral in the case that the three generalized temperatures λ_i are all different from zero, so such a calculation appears. On the other hand, the calculation is rather easy if we put $\lambda_2 = 0$, while keeping $\lambda_1 \neq 0$ and $\lambda_3 \neq 0$. We present such an explicit result that provides an illustrative example of what the geodesic thermodynamics on a non-compact symmetric space can be.

With an iterative Gaussian integration, we obtain:

$$\zeta(\lambda_1, 0, \lambda_3) = \frac{2\pi^{5/2} e^{\frac{\lambda_3^2}{12\lambda_1}}}{\lambda_1^{5/2}} \Rightarrow Z(\lambda_1, 0, \lambda_3, V) = \frac{2\pi^{5/2} e^{\frac{\lambda_3^2}{12\lambda_1}}}{\lambda_1^{5/2}} V \quad (186)$$

and consequently the stochastic Hamiltonian (see Equation (45)) reads as follows:

$$\mathcal{H} |_{\lambda_2=0} = -\log \left(\frac{2\pi^{5/2} e^{\frac{\lambda_3^2}{12\lambda_1}}}{\lambda_1^{5/2}} \right) - \log(V) = -\frac{\lambda_3^2}{12\lambda_1} - \frac{5}{2} \log \left(\frac{\pi}{\lambda_1} \right) - \log(V) - \log(2) \quad (187)$$

Recalling Equation (47), the Shannon Information Entropy (which coincides with minus the thermodynamical entropy S) is the following:

$$\mathcal{I} = \mathcal{H} - \lambda_1 \frac{\partial \mathcal{H}}{\partial \lambda_1} - \lambda_3 \frac{\partial \mathcal{H}}{\partial \lambda_3} = \frac{5 \log(\lambda_1)}{2} - \log(2V) - \frac{5}{2} - \frac{5 \log(\pi)}{2} \quad (188)$$

Finally, the Gibbs state probability distribution is

$$G(\lambda_1, \lambda_3, V, \mathbf{p}, \mathbf{Y}) = \frac{\lambda_1^{5/2} \exp\left(\frac{1}{12}\left(-\frac{\lambda_3^2}{\lambda_1} - 4\lambda_3\left(\mathbf{p}_1 - 2\mathbf{p}_2 + \frac{3\mathbf{p}_3\mathbf{p}_4}{\mathbf{p}_5}\right) - 4\lambda_1(\mathbf{p}_1^2 - \mathbf{p}_2\mathbf{p}_1 + \mathbf{p}_2^2 + 3(\mathbf{p}_3^2 + \mathbf{p}_4^2 + \mathbf{p}_5^2))\right)\right)}{2\pi^{5/2}V} \quad (189)$$

5.2.2. Final Remarks on the GDS Generalized Thermodynamics of the Master Model $SL(3, \mathbb{R})/SO(3)$

Let us observe that if we put $\lambda_3 = 0$ and we interpret $\lambda_1 = 1/(k_B T)$, the partition function (186) takes apart from constant factors the form of that of an ideal gas (compare the following with Equation (A110)):

$$Z(T, 0, V) = 2(\pi k_B T)^{5/2} V \quad (190)$$

where $3N$ is replaced by 5 and N by 1. The meaning of this is quite transparent. It is like we were dealing with a gas made of just one particle (hence $N = 1$) that moves in a space with five dimensions instead of three. Correspondingly, the thermodynamical metric at $\lambda_3 = 0$ is flat as for ideal gases. On the other hand, if we switch on the generalized temperature λ_3 , we have a surprise. Considering the three thermodynamical coordinates:

$$\mathbf{t} = \{\lambda_1, \lambda_3, V\} \quad (191)$$

and working out the thermodynamical metric from the stochastic Hamiltonian in Equation (187), we obtain:

$$\begin{aligned} ds_{therm}^2 &\equiv \frac{\partial^2 \mathcal{H}}{\partial t^i \partial t^j} dt^i dt^j \\ &= \frac{dV^2}{V^2} - \frac{d\lambda_1^2 (15\lambda_1 + \lambda_3^2) - 2d\lambda_1 d\lambda_3 \lambda_1 \lambda_3 + d\lambda_3^2 \lambda_1^2}{6\lambda_1^3} \end{aligned} \quad (192)$$

Clearly, the metric (192) is a metric on a direct product manifold, which is spanned by the two temperatures $\lambda_{1,3}$ and that spanned by the volume. Furthermore, the metric has a Lorentzian signature $\{-1, -1, 1\}$. Finally, the two-dimensional space associated with the two temperatures is not flat, as in the ideal gas case, it is a constant negative curvature manifold. Indeed, introducing the following dreibein:

$$\begin{aligned} \mathbf{E} &= \{E^1, E^2, E^3\} \\ E^1 &= \frac{\sqrt{15\lambda_1 + \lambda_3^2}}{\sqrt{6}\lambda_1^{3/2}} d\lambda_1 - \frac{\lambda_3}{\sqrt{6}\sqrt{\lambda_1}\sqrt{15\lambda_1 + \lambda_3^2}} d\lambda_3 \\ E^2 &= \frac{\sqrt{\frac{5}{2}}}{\sqrt{15\lambda_1 + \lambda_3^2}} d\lambda_3 \\ E^3 &= \frac{1}{V} dV \end{aligned} \quad (193)$$

we can write

$$ds_{therm}^2 = -(E^1)^2 - (E^2)^2 + (E^3)^2 \quad (194)$$

and we derive the following spin-connection

$$\omega^{ij} = \begin{pmatrix} 0 & -\frac{15\sqrt{\frac{3}{2}}\lambda_1^{3/2}}{(\lambda_3^2+15\lambda_1)^{3/2}} E^2 - \frac{\lambda_3^3}{\sqrt{10}(\lambda_3^2+15\lambda_1)^{3/2}} E^1 & 0 \\ \frac{15\sqrt{\frac{3}{2}}\lambda_1^{3/2}}{(\lambda_3^2+15\lambda_1)^{3/2}} E^2 + \frac{\lambda_3^3}{\sqrt{10}(\lambda_3^2+15\lambda_1)^{3/2}} E^1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (195)$$

which yields the following curvature 2-form:

$$R^{ij} = d\omega^{ij} + \omega^{ik} \wedge \omega^{pj} \eta_{kp} = \frac{1}{10} \begin{pmatrix} 0 & E^1 \wedge E^2 & 0 \\ -E^1 \wedge E^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (196)$$

which clearly demonstrates what we just stated. The two-dimensional thermodynamical subspace spanned by the generalized temperatures λ_1, λ_3 is a portion of a constant curvature manifold, namely a portion of a hyperbolic plane, since the negative signature of the metric implies a negative value of the constant curvature. In any case, the constancy of the thermodynamical curvature signals the absence of any critical point and phase transition.

6. Generalized Thermodynamics à la Souriau on Kähler Non-Compact U/H.s

Having clarified the status of generalized thermodynamics associated with the Geodesic Dynamical System and, in general, with Integrable Dynamical Systems, it clearly appears that the corresponding Gibbs Probability Distributions are of little use for Machine Learning algorithms, since they are non-trivial distributions only in momentum space and not on the very manifolds U/H that constitute the hidden layers of a neural network. Although nothing can be excluded a priori, what one needs in the ML context are non-trivial probability distributions on the very manifolds U/H that constitute the layers of the network. Such distributions are provided by generalized thermodynamics à la Souriau as schematically defined in Equations (32)–(35). Such a generalized thermodynamics replaces the Hamiltonians in involution of integrable systems with the moment-maps $\mathfrak{P}_A(\mathbf{Y})$ of a typically non-abelian, actually semisimple or simple, Lie algebra $\mathbb{U}$ that has a Poissonian realization:

$$\{\mathfrak{P}_A, \mathfrak{P}_B\} = f_{AB}^C \mathfrak{P}_C \quad (197)$$

and exists only on Kähler manifolds, as we extensively discussed in the introduction, analyzing the original conception of Barberesco et al. based on the notion of coadjoint orbits. As one sees, generalized thermodynamics à la Souriau is almost the opposite of the generalized thermodynamics related to integrable systems and with the corresponding Liouville Hamiltonians in involution. Its relevance precisely resides in the non-abelian character of the algebra satisfied by the Hamiltonians, that is, the algebra of Killing vector fields of the corresponding Riemannian manifold. The principle underlying the construction of generalized thermodynamics à la Souriau is another one, totally different from Liouville integrability: it is the convergence of the partition function integral (32), namely

$$Z_K(\beta) \equiv \int_{U/H} \exp[-\beta \cdot \mathfrak{P}(\mathbf{Y})] \underbrace{\mathcal{K} \wedge \mathcal{K} \wedge \dots \wedge \mathcal{K}}_{n\text{-times}} < \infty \quad (198)$$

that poses constraints on the vector β of generalized temperatures which, generically, identifies an element of the Lie algebra $\mathbb{U}$ in the chosen basis $\mathfrak{t}_A$ of its generators:

$$\beta^A \mathfrak{t}_A \in \mathbb{U} \quad (199)$$

The determination of the subset $\Omega \subset \mathbb{U}$ (typically not a subalgebra), for which the convergence constraint (198) is satisfied, encodes the new quality of this peculiar Kählerian generalized thermodynamics whose use in Machine Learning appears to be much more promising than generalized thermodynamics based on integrable dynamical systems.

We will demonstrate how the conditions defining the subspace Ω of generalized temperatures is distinctively handy when one uses solvable coordinates, namely while relying on the metric equivalence of non-compact U/H symmetric spaces with their corresponding solvable Lie Group Manifolds $\mathcal{S}_{U/H}$. Indeed, the inequalities defining the range of β are successively extracted from the negativity requirement of the quadratic term in Gaussian integrals over the real line:

$$\int_{-\infty}^{\infty} e^{-\alpha_2 x^2 - \alpha_1 x} dx = \frac{\sqrt{\pi} e^{\frac{\alpha_1^2}{4\alpha_2}}}{\sqrt{\alpha_2}} \quad \text{iff } \alpha_2 > 0 \quad (200)$$

To arrive at this, we first have to single out, among the non-compact symmetric manifolds U/H , the series of Kählerian ones. There are essentially two series, since, as we explained in the introduction, the Kähler character of a coset manifold U/H is uniquely signaled by the presence of a $\mathfrak{u}(1) \simeq \mathfrak{so}(2)$ addend in the compact subalgebra $\mathbb{H}$. Recalling the results and the discussion of the foundational paper [1] we see that the two series are:

$$\mathcal{M}^{[2,q]} \equiv \frac{\text{SO}(2, 2+q)}{\text{SO}(2) \times \text{SO}(2+q)}; \quad \mathbb{SH}_n \equiv \frac{\text{Sp}(2n, \mathbb{R})}{\text{U}(1) \times \text{SU}(n)} \underset{\text{by Cayley map}}{\equiv} \frac{\text{USp}(n, n)}{\text{U}(1) \times \text{SU}(n)} \quad (201)$$

For the generalized Cayley map, see for instance [36] (page 356, formulae (7.2.12)–(7.2.13)) and more generally [58]. The first series was already mentioned in Equation (36) and corresponds to an entire Tits Satake Universality class, the common Tits Satake submanifold of the entire class being:

$$\mathcal{M}_{TS}^{[2,q]} = \mathcal{M}^{[2,1]} = \frac{\text{SO}(2, 3)}{\text{SO}(2) \times \text{SO}(3)} \underset{\substack{\text{spinor double} \\ \text{covering}}}{\simeq} \frac{\text{Sp}(4, \mathbb{R})}{\text{U}(1) \times \text{SU}(2)} \quad (202)$$

As we see from Equation (204), the Tits Satake submanifold of the first series is equivalent, due to the low-dimensional Lie algebra isomorphisms, to the second manifold in the second series. Indeed, the second series is the series of Siegel half spaces of genus n and $\frac{\text{Sp}(4, \mathbb{R})}{\text{U}(1) \times \text{SU}(2)}$ is a double covering of $\frac{\text{SO}(2, 3)}{\text{SO}(2) \times \text{SO}(3)}$ obtained through the use of the four-dimensional spinor representation of $\text{SO}(2, 3)$ rather than its five-dimensional vector one (see [1,9]). The first manifold $n = 1$ of the second series is also the Tits Satake submanifold of the Hyperbolic Space universality class:

$$\mathcal{M}^{[1,1+q]} \equiv \frac{\text{SO}(1, 1+q)}{\text{SO}(1+q)} \quad (203)$$

In the above series, the manifolds are not Kählerian except for the case $q = 1$, which also corresponds to the Tits Satake submanifold of the entire class.

$$\mathcal{M}_{TS}^{[1,q]} = \mathcal{M}^{[1,1]} = \frac{\text{SO}(1, 2)}{\text{SO}(2)} \underset{\substack{\text{spinor double} \\ \text{covering}}}{\simeq} \frac{\text{Sp}(2, \mathbb{R})}{\text{U}(1)} = \frac{\text{SL}(2, \mathbb{R})}{\text{SO}(2)} \underset{\text{by Cayley map}}{\equiv} \frac{\text{SU}(1, 1)}{\text{U}(1)} \quad (204)$$

Finally, in view of the visions of Paint Group invariance and of its relevance in ML algorithms discussed in [1,4] one realizes that the most interesting Kähler manifolds are those of the series $\mathcal{M}^{[2,q]}$ in (201) that are also named Calabi–Vesentini manifolds. An interesting point, whose implications for ML are all to be studied, is that the Calabi–Vesentini manifolds times a hyperbolic plane constitute Special Kähler manifolds (see [1,58]) and can be described by a suitable section of the corresponding flat holomorphic symplectic bundle, as recalled in [1].

6.1. The General Setup

Having singled out the series of non-compact symmetric spaces U/H that are Kählerian and correspondingly apt to support generalized thermodynamics à la Souriau let us develop the general setup for the construction of this latter.

For all U/H , we have the two sinergic Lie algebra decompositions:

$$\mathbb{U} = \mathbb{H} \oplus \mathbb{K} \quad (205)$$

$$\mathbb{U} = \mathbb{H} + \text{Solv}_{U/H} \quad (206)$$

where $\mathbb{H}$ is the maximal compact subalgebra of $\mathbb{U}$ and $\mathbb{K}$ constitutes a linear representation of $\mathbb{H}$ under its adjoint action, but it is not a closed subalgebra:

$$[\mathbb{H}, \mathbb{K}] \subset \mathbb{K}; \quad [\mathbb{K}, \mathbb{K}] \not\subset \mathbb{K} \quad \text{rather} \quad [\mathbb{K}, \mathbb{K}] \subset \mathbb{H} \quad (207)$$

while $\text{Solv}_{U/H}$, that has the same dimension as $\mathbb{K}$, is a closed Lie subalgebra (a solvable one), but it is not a linear representation of $\mathbb{H}$ under its adjoint action:

$$[\mathbb{H}, \text{Solv}_{U/H}] \not\subset \text{Solv}_{U/H}; \quad [\text{Solv}_{U/H}, \text{Solv}_{U/H}] \subset \text{Solv}_{U/H} \quad (208)$$

In the case when U/H is Kählerian, we have the additional essential property:

$$\mathbb{H} = \mathbb{H}_0 \oplus \mathfrak{so}_c(2); \quad [\mathbb{H}_0, \mathbb{H}_0] \subset \mathbb{H}_0; \quad [\mathbb{H}_0, \mathfrak{so}_c(2)] = \mathbf{0} \quad (209)$$

In all cases (compare with Section 4.2), we construct the U -invariant metric on U/H utilizing the vielbein extracted from the left-invariant matrix 1-form Θ of the metric equivalent solvable Lie group $\mathcal{S}_{U/H} \subset U$:

$$\Theta \equiv \mathbb{L}^{-1}(\mathbf{Y}) d\mathbb{L}(\mathbf{Y}) \quad (210)$$

by projecting it onto an orthonormal basis of $\mathbb{K}$ generators:

$$\begin{aligned} \mathbf{V}^A &\equiv \text{Tr}(\Theta \cdot K^A); \quad \text{Tr}(K^A \cdot K^B) = \delta^{AB} \\ ds_{U/H}^2 &= \mathbf{V}^A \times \mathbf{V}^B \delta_{AB} \end{aligned} \quad (211)$$

and we find that we can equivalently write (compare with Equation (101)):

$$ds_{U/H}^2 = \underbrace{\kappa_{AB}}_{\text{const. symm.}} \mathbf{e}^A \times \mathbf{e}^B \quad (212)$$

where the 1-forms $\mathbf{e}^A$ are defined by the expansion of Θ along a basis of generators T_A of the solvable Lie algebra $\text{Solv}_{U/H}$:

$$\Theta = \mathbf{e}^A T_A \quad (213)$$

since the relation between the two synergic decompositions (205) and (206) of the same Lie algebra $\mathbb{U}$ implies that there always exists a constant matrix ν such that (compare with Equation (163)):

$$V^A = \nu^A_B e^B \Rightarrow \kappa = \nu^T \cdot \nu \quad (214)$$

In the Kählerian case, we have the additional item of the Kähler 2-form whose form is general in terms of the matrix representing the adjoint action of the $\mathfrak{so}_c(2)$ generator on the space $\mathbb{K}$. Let us name $\mathbf{X}^c$ such a generator and construct its d -dimensional matrix representation on $\mathbb{K}$:

$$\mathbf{K}_{AB}^c = \delta_{AC} \delta_{BD} \text{Tr} \left(\left[\mathbf{X}^c, K^C \right] \cdot K^D \right) \quad (215)$$

Since $\mathfrak{so}_c(2)$ is a compact subalgebra, the matrix $\mathbf{K}^c$ is necessarily antisymmetric, and the H invariance of the metric guarantees that the Kähler 2-form defined by:

$$\mathcal{K} \equiv \mathbf{K}_{AB}^c \mathbf{V}^A \wedge \mathbf{V}^B \quad (216)$$

is necessarily closed. The proof is simple. By definition of the Levi-Civita connection in vielbein/spin-connection formalism we have:

$$d\mathbf{V}^A = \omega^{AC} \wedge \mathbf{V}^C \quad (217)$$

where $\omega^{IJ} = -\omega^{JI}$ is valued in the d -dimensional representation $\mathbb{K}$ of the $\mathbb{H}$ Lie algebra. Using (217), we obtain

$$d\mathcal{K} = - \underbrace{[\mathbf{K}^c, \omega]_{IJ}}_{=0} V^I \wedge V^J \quad (218)$$

Indeed, $\mathbf{K}^c$ is the $\mathfrak{so}_c(2)$ subalgebra, and it commutes with the whole $\mathbb{H}$ -algebra.

The next question concerns the Killing vector fields and their Hamiltonian representation. It is important to stress another clear-cut distinction. The symmetric space U/H is diffeomorphic to the solvable Lie group $\mathcal{S}_{U/H}$ and the latter, as any Lie group manifold G , possesses two commuting sets of vector fields $\mathbf{t}_A^{[L/R]}$ satisfying the Lie algebra $\mathbb{G}$ of the group:

$$\left[\mathbf{t}_A^{[L/R]}, \mathbf{t}_B^{[L/R]} \right] = f_{AB}^C \mathbf{t}_C^{[L/R]}; \quad \left[\mathbf{t}_A^{[L]}, \mathbf{t}_B^{[R]} \right] = 0 \quad (219)$$

the left-invariant ones $\mathbf{t}_A \equiv \mathbf{t}_A^{[L]}$, dual to the left-invariant 1-forms $\mathbf{e}^B$, according to Equations (104) and (105), generate right-translations, while the right-invariant ones $\mathbf{t}_A^{[R]}$, dual to the right-invariant 1-forms $\mathbf{e}_{[R]}^A$ defined by the expansion along a generator basis T_A of $Solv_{U/H}$ of the right-invariant matrix 1-form:

$$\Theta_{[R]} \equiv d\mathbb{L} \cdot \mathbb{L}^{-1} = \mathbf{e}_{[R]}^A T_A; \quad \mathbf{e}_{[R]}^A \left(\mathbf{t}_A^{[R]} \right) = \delta_B^A \quad (220)$$

generate the left-translations. For this reason, only the right-invariant vector fields $\mathbf{t}_A^{[R]}$ are **Killing vector fields** of the symmetric space metric (211) and, as such, they are also **symplectic Killing vector fields** for the symplectic structure provided by the Kähler 2-form in Equation (216):

$$\ell_{\mathbf{t}_A^{[R]}} \mathcal{K} = 0 \quad (221)$$

where $\ell_{\mathbf{t}}$ denotes the Lie derivative along the vector field $\mathbf{t}$. The right-invariant vector fields $\mathbf{t}_A^{[R]}$ are not the only Killing vectors and hence symplectic Killing vectors: Indeed, the symmetric space metric is invariant with respect to the whole group U and we need a set of Killing vector fields satisfying the whole $\mathbb{U}$ Lie algebra. According to the decomposi-

tion (206), we just have to add the Killing vector fields spanning the compact subalgebra $\mathbb{H}$. The question is how to obtain their explicit expression.

6.1.1. General Construction Method of the Killing Vector Fields

In order to construct the expression in solvable coordinate $\mathbf{Y}$ of the Killing vector fields associated with the compact generators, we utilize the following procedure, which is general and applies to any Killing vector field.

From the general theory of coset manifolds (see, for instance, [55], second volume, section 5.2.3, page 114 and the following ones), we know that when we act on the left on a coset representative $\mathbb{L}(y)$ with any element $g \in U$ we have:

$$g \mathbb{L}(y) = \mathbb{L}(g(y)) \cdot h(y, g); \quad h(y, g) \in H \subset U \quad (222)$$

where $y' = g(y)$ is the coordinate of the new point in the coset reached by the g transformation and $h(y, g)$, which lies in the subgroup, is named the compensator. Typically, the determination of the compensator is a cumbersome task; yet, in the solvable parameterization, there is a well-defined universal algorithm for the determination of $g(y)$ that bypasses the determination of the compensator. Our coset representative is an element of the solvable Lie group and as demonstrated in [1,2] we can always use for any U/H the so-called triangular basis, where the solvable coset representative is an upper triangular matrix. Hence, the solvable coset representative $\mathbb{L}(\mathbf{Y})$ is upper triangular, and the matrices $h \in H$ of the compact subgroup, in particular the compensators, are all orthogonal $h \cdot h^T = \mathbf{1}$. It follows that defining the symmetric matrix:

$$\mathcal{M}(\mathbf{Y}) \equiv \mathbb{L}(\mathbf{Y}) \cdot \mathbb{L}^T(\mathbf{Y}) \quad (223)$$

We have

$$\forall g \in U \quad : \quad g \cdot \mathcal{M}(\mathbf{Y}) \cdot g^T = \mathcal{M}(g(\mathbf{Y})) \quad (224)$$

In order to obtain $g(\mathbf{Y})$, which is our goal, it suffices to utilize the finite recursive Cholewsky–Crout algebraic algorithm (see [1,2]) that, given $\mathcal{M}(g(\mathbf{Y}))$, uniquely determines the upper triangular matrix $\mathbb{L}(g(y))$ such that:

$$\mathcal{M}(g(\mathbf{Y})) = \mathbb{L}(g(\mathbf{Y})) \cdot \mathbb{L}^T(g(\mathbf{Y})) \quad (225)$$

Then applying the inverse of the Σ exponential map according to the definitions and conventions of [1,2] we get:

$$\Sigma^{-1}[\mathbb{L}(g(\mathbf{Y}))] = g(\mathbf{Y})^\alpha T_\alpha \quad (226)$$

In Equation (226), consider the compact group subgroup elements of the form:

$$g_i[\theta] = \exp[\theta J_i] \quad (227)$$

where J_i ($i = 1, \dots, m = \dim \mathbb{H}$) is a basis of generators of the compact subalgebra $\mathbb{H}$. The $g_i[\theta]$ are elements of the m one-parameter subgroups of H . Expanding in a power series of θ , we get:

$$g_i[\theta](\mathbf{Y})^\alpha = \mathbf{Y}^\alpha + f_i^\alpha(\mathbf{Y}) + \mathcal{O}(\theta^2) \quad (228)$$

The searched for Killing vector fields are:

$$\mathfrak{k}_i \equiv f_i^\alpha(\mathbf{Y}) \frac{\partial}{\partial Y^\alpha} \quad (229)$$

and note that by construction they satisfy the Lie algebra $\mathbb{H}$ and are symplectic Killing vector fields, according to the definition (221). Note also that the above construction of the Killing vector fields that was indispensable for the compact subalgebra $\mathbb{H}$ could also be applied to the determination of the Killing vector fields associated with the solvable Lie subalgebra, if we did not know them independently, or to those associated with the $\mathbb{K}$ -generators, corresponding to the decomposition (205). Indeed, the above procedure is completely general for any set of generators J_Λ of the entire Lie algebra $\mathbb{U}$.

6.1.2. The General Form of the Moment-Maps

Given any set $\mathfrak{k}_\Lambda$ of Killing vector fields, each uniquely associated with a generator J_Λ of the Lie algebra $\mathbb{U}$ via the construction mentioned above in Equations (227) and (229), one obtains the corresponding moment map via another fully general formula which is extremely simple:

$$\begin{aligned} \mathfrak{P} &: \mathbb{U} \longrightarrow \mathbb{C}^\infty\left(\frac{\mathbb{U}}{\mathbb{H}}\right) \\ \mathfrak{P}_\Lambda(\mathbf{Y}) &= \frac{1}{2} \text{Tr}\left[X_c \cdot \mathbb{L}^{-1}(\mathbf{Y}) \cdot J_\Lambda \cdot \mathbb{L}(\mathbf{Y})\right] \end{aligned} \quad (230)$$

The functions $\mathfrak{P}_\Lambda(\mathbf{Y})$ satisfy the necessary condition with respect to the Killing vector fields $\mathfrak{k}_\Lambda$:

$$i_{\mathfrak{k}_\Lambda} \mathcal{K} = d\mathfrak{P}_\Lambda(\mathbf{Y}) \quad (231)$$

and their definition (230) has a very simple interpretation; they are the projection onto the $\mathfrak{so}(2)_c$ central subalgebra defining the Kähler structure of the adjoint transformation of the generator J_Λ . An important comment is the following. Since all the non-compact symmetric spaces $\mathbb{U}/\mathbb{H}$ are Hadamard–Cartan manifolds diffeomorphic to $\mathbb{R}^n$ (see [2]), they can be covered by just one open chart, and the solvable coordinates $\mathbf{Y}$ provide such a chart. For this reason, the moment maps in solvable coordinates $\mathfrak{P}_\Lambda(\mathbf{Y})$ are globally defined functions over the whole manifold and Equation (231) holds true globally.

6.1.3. The Partition Function and the Gibbs Probability Distribution

Equipped with the above general weapons, one can address the conditions on the temperature vector β for the convergence of the integral (198) defining the partition function $Z(\beta)$. The first observation is that in the privileged solvable coordinate chart, the volume form reduces to:

$$\underbrace{\mathcal{K} \wedge \mathcal{K} \wedge \dots \wedge \mathcal{K}}_{n \text{ times}} = \text{const} \times \underbrace{dY^1 \wedge \dots \wedge dY^{r_{n.c.}}}_{\text{Cartan coordinates}} \underbrace{dY^{r_{n.c.}+1} \wedge \dots \wedge dY^{2n}}_{\text{nilpotent coordinates}} \quad (232)$$

where $r_{n.c.}$ is the non-compact rank of the considered Kählerian symmetric space $\mathbb{U}/\mathbb{H}$. If we choose the first series in Equation (201), then the non-compact rank is always $r_{n.c.}$ and we just have two solvable coordinates associated with the unique two Cartan generators, while all the other coordinates $Y^{r_{n.c.}+1}, \dots, Y^{2n}$ are associated with nilpotent generators corresponding to long and short roots, the latter organized in two multiplets assigned to the fundamental representation of the Paint Group $G_{\text{Paint}} = \text{SO}(q)$. On the contrary, all the Kähler manifolds in the second series are maximally split, and the non-compact rank increases linearly in n . In any case, the strategy to calculate the partition function is that of starting with the integration one-by-one on the nilpotent coordinates, which leads each time to a Gaussian integral of the form (200) and to a corresponding positivity constraint on the β vector. As we show in the two explicitly constructed examples, the integration of the nilpotents for $r_{n.c.} = 2$ can be explicitly performed analytically until we reach a function to be integrated on the Cartan coordinates, which can be verified to

be explicitly positive definite and exponentially decreasing at infinity so as to guarantee partition function convergence. In the simplest case of the hyperbolic plane, we can also perform the last integration analytically, obtaining the partition function in closed form and the thermodynamical metric. For the Siegel $n = 2$ plane, the last integration over the Cartan fields has to be done numerically, yet all the thermodynamical quantities are accessible as compiled functions. The extension of the results for the Siegel plane to all the manifolds of the first series in (201) via a convenient use of Paint Group invariance is an issue of further research. In Appendix D, we present the preliminary setup calculation for the case $q = 2$.

6.2. Generalized Thermodynamics à la Souriau of the Poincaré–Lobachevsky Hyperbolic Plane $\mathbb{H}_2$

Let us begin with the simplest example, namely with the hyperbolic plane $\mathbb{H}_2$. Utilizing the representation:

$$\mathbb{H}_2 = \frac{\mathrm{SL}(2, \mathbb{R})}{\mathrm{SO}(2)} \quad (233)$$

and following all the conventions of [1], we write the generators of the full $\mathrm{SL}(2, \mathbb{R})$ Lie algebra in two ways, namely in the orthogonal decomposition:

$$\mathfrak{sl}(2, \mathbb{R}) = \mathfrak{so}(2) \oplus \mathbb{K} \quad (234)$$

and in the solvable subalgebra decomposition:

$$\mathfrak{sl}(2, \mathbb{R}) = \mathfrak{so}(2) + \mathrm{Solv}_2 \quad (235)$$

For the 1-dimensional compact subalgebra $\mathbb{H} = \mathfrak{so}(2)$, both in Equations (234) and (235), we choose the same generator:

$$X_c = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (236)$$

The two generators of the orthogonal non-compact space $\mathbb{K}$ are given by:

$$K_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}, \quad K_2 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \quad (237)$$

while our chosen basis of generators for the solvable Lie algebra is:

$$\mathrm{Solv}_2 = \mathrm{span}\{T_1, T_2\}; \quad T_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad T_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (238)$$

As explained in [1], the general form of a solvable Lie group element is:

$$\mathbb{L}(\mathbf{Y}) = \begin{pmatrix} e^{Y_1} & e^{Y_1} Y_2 \\ 0 & e^{-Y_1} \end{pmatrix} \quad (239)$$

and the left-invariant matrix 1-form decomposes as follows in the solvable Lie algebra basis:

$$\Theta \equiv \mathbb{L}^{-1} d\mathbb{L} = \mathbf{e}^1 T_1 + \mathbf{e}^2 T_2 \quad (240)$$

where the two left-invariant 1-forms $\mathbf{e}^{1,2}$ have the following appearance:

$$\begin{aligned}\mathbf{e}^1 &= dY_1 \\ \mathbf{e}^2 &= 2Y_2 dY_1 + dY_2\end{aligned}\quad (241)$$

The zweibein of the 2-dimensional space is instead defined by the projection of the left-invariant 1-form along the $\mathbb{K}$ generators:

$$\mathbf{V}^i = \text{Tr}[\Theta \cdot K_i] \quad (242)$$

and one obtains the relation:

$$\mathbf{V} = \nu \mathbf{e}; \quad \nu = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \quad (243)$$

so that the norm form on the solvable Lie algebra mentioned in Equation (101) is as follows:

$$\kappa = \nu^T \cdot \nu = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \quad (244)$$

Defining the adjoint action of the unique $\mathfrak{so}(2)$ generator on the K_i generators and hence on the zweibein $\mathbf{V}$:

$$K^c = \text{adj}_{X_c}[\mathbb{K}] = \text{Tr}([X_c, K_i] \cdot K_j) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (245)$$

we obtain the Kähler 2-form:

$$\mathcal{K} = K_{ij}^c \mathbf{V}^i \wedge \mathbf{V}^j = -2 \mathbf{V}^1 \wedge \mathbf{V}^2 = -2 dY_1 \wedge dY_2 \quad (246)$$

whose closure $d\mathcal{K} = 0$ is immediately verified.

The Kähler metric and its inverse are immediately calculated from the above data:

$$\mathbf{g}_{ij} = \begin{pmatrix} 2(Y_2^2 + 1) & Y_2 \\ Y_2 & \frac{1}{2} \end{pmatrix}; \quad \mathbf{g}^{-1|ij} = \begin{pmatrix} \frac{1}{2} & -Y_2 \\ -Y_2 & 2(Y_2^2 + 1) \end{pmatrix} \quad (247)$$

and the complex structure tensor in the solvable coordinate basis is obtained from the explicit form of the Kähler 2-form provided in Equation (246):

$$J_c = K^c \cdot g^{-1} = \begin{pmatrix} Y_2 & -2(Y_2^2 + 1) \\ \frac{1}{2} & -Y_2 \end{pmatrix}; \quad J_c^2 = -\mathbf{1} \quad (248)$$

The metric (247) is explicitly Hermitian with respect to the complex structure (248):

$$J_c \cdot \mathbf{g} \cdot J_c^T = \mathbf{g} \quad (249)$$

The left-invariant 1-forms $\mathbf{e}$ on the solvable Lie group manifold $\mathcal{S}_2$, in terms of which we have constructed the Kähler metric, satisfy the Maurer–Cartan equation written below that defines the structure constant of the solvable Lie algebra $\text{Sol}v_2$

$$d\mathbf{e}^1 = 0; \quad d\mathbf{e}^2 + 2\mathbf{e}^1 \wedge \mathbf{e}^2 = 0 \quad (250)$$

The same Maurer–Cartan Equation (250) is also satisfied by the right-invariant 1-forms obtained by decomposing the right-invariant matrix 1-form $d\mathbb{L} \cdot \mathbb{L}$.

Following all the procedures detailed in the previous Section 6.1 we obtain the explicit form of the three Killing vector fields in solvable coordinates:

$$\begin{aligned}\mathfrak{k}_0 &= e^{2Y_1} Y_2 \frac{\partial}{\partial Y^1} + \left(e^{-2Y_1} - e^{2Y_1} (Y_2^2 + 1) \right) \frac{\partial}{\partial Y^2} \\ \mathfrak{k}_1 &= \frac{\partial}{\partial Y^1} \\ \mathfrak{k}_2 &= \exp[-2Y^1] \frac{\partial}{\partial Y^2}\end{aligned}\quad (251)$$

where $\mathfrak{k}_0$ is the Killing vector field associated with compact generator X_c , while $\mathfrak{k}_{1,2}$ are the Killing vector fields associated with the solvable Lie algebra generators $T_{1,2}$. The corresponding moment maps calculated by means of Equation (230) are the following ones:

$$\begin{aligned}\mathfrak{P}_0(\mathbf{Y}) &= \frac{1}{2} e^{-2Y_1} \left(-e^{4Y_1} (Y_2^2 + 1) - 1 \right) \\ \mathfrak{P}_1(\mathbf{Y}) &= -Y_2 \\ \mathfrak{P}_2(\mathbf{Y}) &= -\frac{1}{2} e^{-2Y_1}\end{aligned}\quad (252)$$

6.2.1. Calculation of the Partition Function

For simplicity of writing, naming α, β, γ the three components $\beta^0, \beta^1, \beta^2$ of the temperature vector β , and $x = Y_1, y = Y_2$ the two solvable coordinates, the partition function to be computed is the following:

$$Z(\alpha, \beta, \gamma) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \exp \left[-\frac{1}{2} e^{-2x} \left(\alpha + \gamma + \alpha e^{4x} (y^2 + 1) + 2\beta e^{2x} y \right) \right] \quad (253)$$

and using the convergence condition of the Gaussian integrals recalled in Equation (200), we get the following constraints:

$$\alpha > 0; \quad \alpha(\alpha + \gamma) - \beta^2 > 0 \quad (254)$$

When the above conditions are satisfied, the integrals are easily calculated, and we obtain:

$$Z(\alpha, \beta, \gamma) = \frac{\pi e^{-\sqrt{\alpha(\alpha+\gamma)-\beta^2}}}{\sqrt{\alpha(\alpha+\gamma)-\beta^2}} \quad (255)$$

The corresponding **Gibbs probability distribution** takes the following appearance:

$$G(\alpha, \beta, \gamma, Y^1, Y^2) = \frac{\sqrt{\alpha(\alpha+\gamma)-\beta^2} \exp \left[\sqrt{\alpha(\alpha+\gamma)-\beta^2} - \frac{1}{2} e^{-2Y_1} (\alpha e^{4Y_1} (Y_2^2 + 1) + \alpha + 2\beta e^{2Y_1} Y_2 + \gamma) \right]}{\pi} \quad (256)$$

Equations (254)–(256) take a nicer form by a linear redefinition of the temperatures:

$$\alpha = \delta + \zeta; \quad \gamma = -2\zeta; \quad \beta = \beta \quad (257)$$

where δ, ζ are the new temperature parameters. We get:

$$\delta > 0; \quad \delta^2 - \beta^2 - \zeta^2 > 0 \quad (258)$$

$$Z(\delta, \beta, \zeta) = \frac{\pi e^{-\sqrt{\delta^2 - \beta^2 - \zeta^2}}}{\sqrt{\delta^2 - \beta^2 - \zeta^2}} \quad (259)$$

$$G(\delta, \beta, \zeta, Y^1, Y^2) = \frac{\sqrt{-\beta^2 + \delta^2 - \zeta^2} \exp\left(\sqrt{-\beta^2 + \delta^2 - \zeta^2} - \frac{1}{2}e^{-2Y_1}(2\beta e^{2Y_1}Y_2 + e^{4Y_1}(Y_2^2 + 1)(\delta + \zeta) + \delta - \zeta)\right)}{\pi} \quad (260)$$

It appears from Equation (258) that the Souriau $\Omega \subset \mathfrak{sl}(2, \mathbb{R})$ subspace of allowed temperatures is a cone in the three-dimensional Lie algebra space:

$$= \left\{ \begin{pmatrix} \beta & \delta - \zeta \\ -\delta - \zeta & -\beta \end{pmatrix} \in \mathfrak{sl}(2, \mathbb{R}) \mid \delta > 0, \delta^2 - \beta^2 - \zeta^2 > 0 \right\} \quad (261)$$

as it is shown in Figure 1.

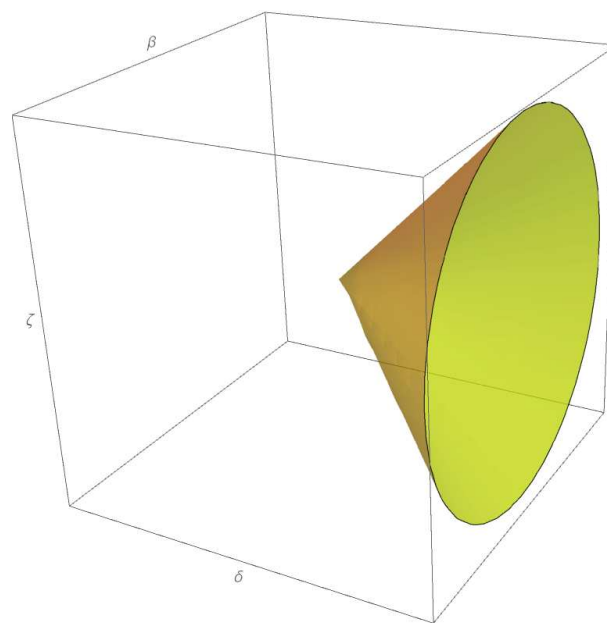


Figure 1. The cone Ω of Souriau allowed temperature vectors in the $\mathfrak{sl}(2, \mathbb{R})$ Lie algebra space as defined in Equation (261).

6.2.2. Visualization of the Gibbs Probability Distributions

In the perspective of Data Science applications, the temperature vectors $\beta = \{\delta, \beta, \zeta\} \in \Omega$ define, as the Gibbs states (260), probability distributions over the symmetric space $U/H = \text{SL}(2, \mathbb{R})/\text{O}(2)$, namely the Poincaré hyperbolic plane. Such Gibbs states are the appropriate generalizations to not-flat Cartan–Hadamard spaces of the familiar Gaussian distributions pertaining to flat space. The temperature vector models such distributions. It is very helpful to visualize the Gibbs states (260) utilizing the disk model of the Poincaré plane:

$$\text{Disk} = \left\{ \{x, y\} \in \mathbb{R}^2 \mid x^2 + y^2 < 1 \right\} \quad (262)$$

The relation between the solvable coordinates Y_1, Y_2 , and the coordinates x, y is provided by the following formula (see [1,4]):

$$Y_2 = \frac{4y}{x^2 + y^2 - 1} \quad , \quad Y_1 = \log\left(-\frac{x^2 + y^2 - 1}{x^2 - 2x + y^2 + 1}\right) \quad (263)$$

Substituting Equation (263) in Equation (260) and furthermore utilizing, polar coordinates in the β, ζ plane of the temperature space:

$$\beta = \mu \cos(\theta); \quad \zeta = \mu \sin(\theta); \quad 0 < \mu < \delta \quad (264)$$

we obtain the following three-parameter family of probability distributions over the disk (262):

$$G(\delta, \mu, \theta, x, y) = \frac{\sqrt{\delta^2 - \mu^2}}{\pi} \times \exp \left[\sqrt{\delta^2 - \mu^2} - \frac{(x^2 - 2x + y^2 + 1)^2 \left(\delta - \mu \sin(\theta) + \frac{\left(\frac{16y^2}{(x^2 + y^2 - 1)^2} + 1 \right) (x^2 + y^2 - 1)^4 (\delta + \mu \sin(\theta))}{(x^2 - 2x + y^2 + 1)^4} + \frac{8\mu y \cos(\theta) (x^2 + y^2 - 1)}{(x^2 - 2x + y^2 + 1)^2} \right)}{2(x^2 + y^2 - 1)^2} \right] \quad (265)$$

We present in Figure 2 a few examples of plots of such probability distributions.

6.2.3. The Kähler Geothermodynamic Metric and Curvature

The convex cone conditions (258) defining the Souriau temperature space in the case $\mathfrak{so}(2, 1) \simeq \mathfrak{su}(1, 1) \simeq \mathfrak{sl}(2, \mathbb{R})$ were found, in different notations and setups, also by the authors of [30–34]. These authors interpreted the cone as the future-directed light cone of three-dimensional Minkowski space, yet this is just a special feature of the low-dimensional case of U/H Kähler manifolds under consideration. What is general is that the temperature associated with the $\mathfrak{so}(2)_c$ subalgebra defining the Kähler structure must be strictly positive, and then the temperatures associated with the other generators receive constraints in terms of the latter in order to maintain convergence of the other Gaussian integrals.

Independent of such observations, Gibbs states are parameterized probability distributions that can be used to interpolate data by fitting their parameters, namely their temperatures. Generalized thermodynamics provides a metric on the space of temperatures and therefore yields a distance between two distributions, each corresponding to an equilibrium state.

According to the general theory discussed in previous sections, we have the stochastic Hamiltonian:

$$\mathcal{H}^{stoch} \equiv -\log[Z(\delta, \beta, \zeta)] = \sqrt{-\beta^2 + \delta^2 - \zeta^2} + \frac{1}{2} \log[-\beta^2 + \delta^2 - \zeta^2] - \log(\pi) \quad (266)$$

and according to Equation (47) we calculate Shannon Information Functional:

$$\mathcal{I} = \mathcal{H}^{stoch} - \left(\delta \frac{\partial}{\partial \delta} + \beta \frac{\partial}{\partial \beta} + \zeta \frac{\partial}{\partial \zeta} \right) \mathcal{H}^{stoch} = \frac{1}{2} \left(\log \left[\frac{\delta^2 - \beta^2 - \zeta^2}{\pi^2} \right] - 2 \right) \quad (267)$$

As we see the Information Functional tends to $-\infty$ on the boundary of the cone of Figure 1, confirming that the norm $\mathbb{N}(\beta)$ of the temperature vector is the analogue of the inverse thermodynamical temperature

$$\mathbb{N}(\beta) = \sqrt{\delta^2 - \beta^2 - \zeta^2} \sim \frac{1}{T} \quad (268)$$

and that Shannon Information Functional $\mathcal{I}$ is just minus the Thermodynamical Entropy S :

$$\mathcal{I} \sim -S \quad (269)$$

When temperature T goes to ∞ , the Thermodynamical Entropy S goes to infinity, and the information content of the Gibbs probability distribution is largely negative since we have maximal disorder. On the contrary, when $T \rightarrow 0$, which means that $N(\beta) \rightarrow \infty$ the Information Entropy tends to ∞ as well, logarithmically slowly, since we have a lot of information. As an illustration of this basic feature, the reader is referred to Figure 3.

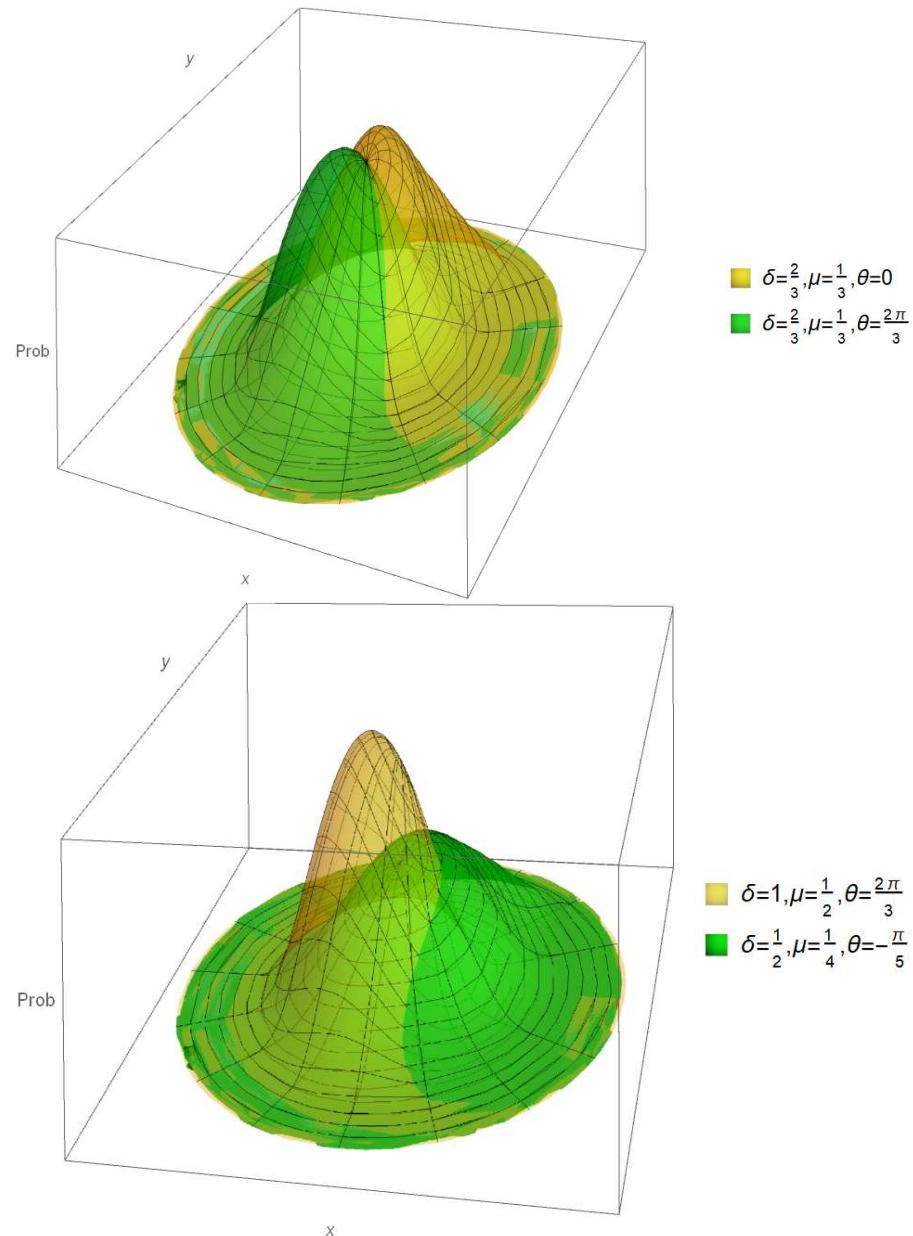


Figure 2. Examples of plots of the Gibbs probability distributions (265) over the Poincaré disk, labeled by different set of temperatures. The exponential Gaussian decay toward infinity is visually evident, as much as the deformed bell shape. In the **first image**, we compare two distributions with the same values of $\delta\mu$ but with a different angle θ . In the second image, we compare two distributions that differ in all parameters.

Recalling next the general form of the thermodynamical metric expressed in terms of the Hessian of the stochastic Hamiltonian (see Equation (69)), we write

$$\begin{aligned}
ds_{\text{geotherm}}^2 &= \frac{\partial^2 \mathcal{H}^{\text{sto}}}{\partial \beta^i \partial \beta^j} d\beta^i \times d\beta^j \\
&= \frac{1}{\mathbb{N}^4} \times \left\{ d\beta^2 \left(-\left(\beta^2 + (\mathbb{N}+1)(\delta - \zeta)(\delta + \zeta) \right) \right) + 2(\mathbb{N}+2) \beta d\beta \times (\delta d\delta - \zeta d\zeta) \right. \\
&\quad \left. - d\delta^2 \left(\delta^2 + \beta^2(\mathbb{N}+1) + \zeta^2(\mathbb{N}+1) \right) + 2(\mathbb{N}+2) \delta \times \zeta \times d\delta \times d\zeta + d\zeta^2 \left((\mathbb{N}+1)(\beta - \delta)(\beta + \delta) - \zeta^2 \right) \right\} \quad (270)
\end{aligned}$$

where $\mathbb{N}$ is a shorthand for $\mathbb{N}(\beta)$ as defined in Equation (268) and $\beta \equiv \{\delta, \beta, \zeta\}$.

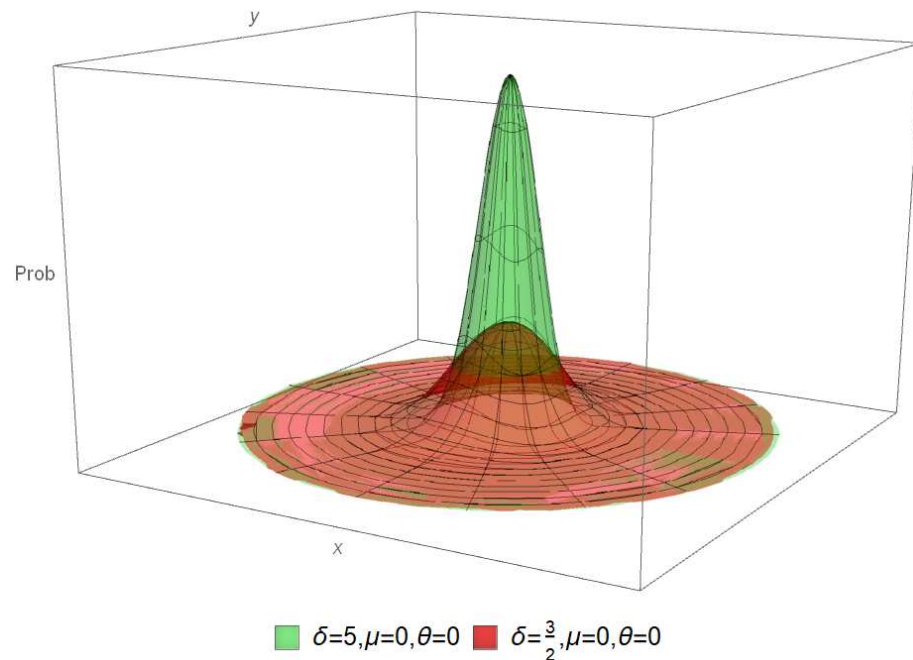


Figure 3. In this figure, we compare two Gibbs distributions on the Poincaré plane corresponding to a lower and higher value of the norm (268). As one sees for a high value of the norm, the distribution is very sharply shaped around its maximal value, while for a lower norm, it is much broader. For high norm, we know with much more precision the actual location of the stochastic variable in the plane.

We are interested in calculating the Riemannian curvature of the metric (270), but in order to better understand the intrinsic properties of the curvature 2-form, we prefer to study it in the anholonomic basis provided by the vielbein formalism. For this reason, we need to construct a suitable **dreibein** that reproduces (270) as the sum of squares of appropriate 1-forms:

$$ds_{\text{geotherm}}^2 = - \sum_{i=1}^3 V^i \times V^i \quad (271)$$

With some work, we have found that the following dreibein fulfils its own job:

$$\begin{aligned}
V^1 &= \frac{1}{\mathbb{N}^2 \sqrt{\mathbb{N}+2} (\delta - \zeta)} \times \left\{ -\beta d\beta (\mathbb{N}+2)(\delta - \zeta) + d\delta \left(\beta^2(\mathbb{N}+1) + (\delta - \zeta)(\delta - \zeta(\mathbb{N}+1)) \right) \right. \\
&\quad \left. + d\zeta \left((\delta - \zeta)(\delta - \zeta + \delta\mathbb{N}) - \beta^2(\mathbb{N}+1) \right) \right\} \\
V^2 &= \frac{\sqrt{\mathbb{N}+1} (d\beta(\delta - \zeta) + \beta(d\zeta - d\delta))}{\mathbb{N}(\delta - \zeta)} \\
V^3 &= (d\delta - d\zeta) \sqrt{\frac{-\mathbb{N}^2 + \mathbb{N} + 2}{(4 - \mathbb{N}^2)(\delta - \zeta)^2}} \quad (272)
\end{aligned}$$

and using the MATHEMATICA code **Vielbgrav23** developed by one of us (see [9]), we have found the explicit form of the spin connection and of the curvature 2-form:

$$\mathbf{d}V^i + \omega^{ij} \wedge V^k \eta_{jk} = 0; \quad \mathfrak{R}^{ij} \equiv \mathbf{d}\omega^{ij} + \omega^{ik} \wedge \omega^{\ell j} \eta_{k\ell}; \quad \eta = \text{diag}(-1, -1, -1) \quad (273)$$

We got:

$$\mathfrak{R}^{ij} = \begin{pmatrix} 0 & \mathcal{G}(\beta) V^2 \wedge V^3 + \mathcal{F}(\beta) V^1 \wedge V^2 & \mathcal{Q}(\beta) V^1 \wedge V^3 \\ -\mathcal{G}(\beta) V^2 \wedge V^3 - \mathcal{F}(\beta) V^1 \wedge V^2 & 0 & \mathcal{P}(\beta) V^2 \wedge V^3 + \mathcal{G}(\beta) V^1 \wedge V^2 \\ -\mathcal{Q}(\beta) V^2 \wedge V^3 & -\mathcal{P}(\beta) V^2 \wedge V^3 - \mathcal{G}(\beta) V^1 \wedge V^2 & 0 \end{pmatrix} \quad (274)$$

where the four coefficients $\mathcal{F}(\beta), \mathcal{G}(\beta), \mathcal{Q}(\beta), \mathcal{P}(\beta)$ are not constants; rather, they have a non-trivial dependence on the temperature vector components. However, they display the $\text{SO}(2)_c$ invariance of the geothermic metric. Indeed using the polar parameterization (264) of the two non-compact temperatures β, ζ , we find that the intrinsic components of the curvature 2-form, $\mathcal{F}(\beta), \mathcal{G}(\beta), \mathcal{Q}(\beta), \mathcal{P}(\beta)$, depend only on δ, μ , and do not depend on the angle θ . Explicitly, we found:

$$\begin{aligned} \mathcal{F} &= \frac{N_F}{D_F} \\ N_F &= -(\delta^4 - 2\delta^2\mu^2 - 4\delta^2 + \mu^4 + 4\mu^2) \left(\delta^8 - 4\delta^6\mu^2 + 71\delta^6 + 6\delta^4\mu^4 - 213\delta^4\mu^2 + 384\delta^4 - 4\delta^2\mu^6 \right. \\ &\quad + 213\delta^2\mu^4 - 768\delta^2\mu^2 - 426\delta^2\mu^2\sqrt{\delta^2 - \mu^2} + 426\delta^2\sqrt{\delta^2 - \mu^2} - 426\mu^2\sqrt{\delta^2 - \mu^2} + 104\sqrt{\delta^2 - \mu^2} \\ &\quad - 13\mu^6\sqrt{\delta^2 - \mu^2} + 39\delta^2\mu^4\sqrt{\delta^2 - \mu^2} + 213\mu^4\sqrt{\delta^2 - \mu^2} + 284\delta^2 + 13\delta^6\sqrt{\delta^2 - \mu^2} - 39\delta^4\mu^2\sqrt{\delta^2 - \mu^2} \\ &\quad \left. + 213\delta^4\sqrt{\delta^2 - \mu^2} + \mu^8 - 71\mu^6 + 384\mu^4 - 284\mu^2 + 16 \right) \\ D_F &= 4 \left(\sqrt{\delta^2 - \mu^2} + 1 \right) \left(\sqrt{\delta^2 - \mu^2} + 2 \right)^3 \left(\sqrt{\delta^2 - \mu^2} + \delta^2 - \mu^2 \right)^2 \times \\ &\quad \times \left(\sqrt{\delta^2 - \mu^2} - \delta^2 + \mu^2 + 2 \right) \left(3\sqrt{\delta^2 - \mu^2} + \delta^2 - \mu^2 + 2 \right) \end{aligned} \quad (275)$$

$$\begin{aligned} \mathcal{G} &= \frac{N_G}{D_G} \\ N_G &= (\delta^2 - \mu^2) \left(\frac{-\delta^2 + \mu^2 + 4}{\sqrt{\delta^2 - \mu^2} - \delta^2 + \mu^2 + 2} \right)^{3/2} \left(\delta^6 - 3\delta^4\mu^2 + 26\delta^4 + 3\delta^2\mu^4 \right. \\ &\quad - 52\delta^2\mu^2 - 16\delta^2\mu^2\sqrt{\delta^2 - \mu^2} + 44\delta^2\sqrt{\delta^2 - \mu^2} - 44\mu^2\sqrt{\delta^2 - \mu^2} + 20\sqrt{\delta^2 - \mu^2} + 8\mu^4\sqrt{\delta^2 - \mu^2} \\ &\quad \left. + 41\delta^2 + 8\delta^4\sqrt{\delta^2 - \mu^2} - \mu^6 + 26\mu^4 - 41\mu^2 + 4 \right) \\ D_G &= 2 \left(\sqrt{\delta^2 - \mu^2} + 1 \right) \left(\sqrt{\delta^2 - \mu^2} + 2 \right)^{5/2} \left(\sqrt{\delta^2 - \mu^2} + \delta^2 - \mu^2 \right)^2 \left(3\sqrt{\delta^2 - \mu^2} + \delta^2 - \mu^2 + 2 \right) \end{aligned} \quad (276)$$

$$\begin{aligned} \mathcal{Q} &= \frac{N_Q}{D_Q} \\ N_Q &= (\delta^2 - \mu^2 - 4) \left(\delta^6 - 3\delta^4\mu^2 + 25\delta^4 + 3\delta^2\mu^4 - 50\delta^2\mu^2 - 16\delta^2\mu^2\sqrt{\delta^2 - \mu^2} + 38\delta^2\sqrt{\delta^2 - \mu^2} \right. \\ &\quad - 38\mu^2\sqrt{\delta^2 - \mu^2} + 8\sqrt{\delta^2 - \mu^2} + 8\mu^4\sqrt{\delta^2 - \mu^2} + 28\delta^2 + 8\delta^4\sqrt{\delta^2 - \mu^2} - \mu^6 + 25\mu^4 - 28\mu^2 \left. \right) \\ D_Q &= 4 \left(\sqrt{\delta^2 - \mu^2} + 2 \right)^6 \left(\sqrt{\delta^2 - \mu^2} - \delta^2 + \mu^2 + 2 \right)^4 \end{aligned} \quad (277)$$

$$\begin{aligned}
\mathcal{P} &= \frac{N_P}{D_P} \\
N_P &= \left(\delta^2 - \mu^2 - 4 \right)^2 \left(\delta^2 - \mu^2 \right) \left(14\delta^{10} - 70\delta^8\mu^2 + 340\delta^8 + 140\delta^6\mu^4 - 1360\delta^6\mu^2 + 1562\delta^6 - 140\delta^4\mu^6 \right. \\
&\quad + 2040\delta^4\mu^4 - 4686\delta^4\mu^2 + 1864\delta^4 + 70\delta^2\mu^8 - 1360\delta^2\mu^6 + 4686\delta^2\mu^4 - 3728\delta^2\mu^2 \\
&\quad - 4030\delta^2\mu^2\sqrt{\delta^2 - \mu^2} + 1210\delta^2\sqrt{\delta^2 - \mu^2} - 1210\mu^2\sqrt{\delta^2 - \mu^2} + 136\sqrt{\delta^2 - \mu^2} - \mu^{10}\sqrt{\delta^2 - \mu^2} \\
&\quad + 5\delta^2\mu^8\sqrt{\delta^2 - \mu^2} + 89\mu^8\sqrt{\delta^2 - \mu^2} - 356\delta^2\mu^6\sqrt{\delta^2 - \mu^2} - 869\mu^6\sqrt{\delta^2 - \mu^2} + 2607\delta^2\mu^4\sqrt{\delta^2 - \mu^2} \\
&\quad + 2015\mu^4\sqrt{\delta^2 - \mu^2} + 524\delta^2 + \delta^{10}\sqrt{\delta^2 - \mu^2} \\
&\quad - 5\delta^8\mu^2\sqrt{\delta^2 - \mu^2} + 89\delta^8\sqrt{\delta^2 - \mu^2} - 356\delta^6\mu^2\sqrt{\delta^2 - \mu^2} + 869\delta^6\sqrt{\delta^2 - \mu^2} + 10\delta^6\mu^4\sqrt{\delta^2 - \mu^2} \\
&\quad - 2607\delta^4\mu^2\sqrt{\delta^2 - \mu^2} + 2015\delta^4\sqrt{\delta^2 - \mu^2} - 10\delta^4\mu^6\sqrt{\delta^2 - \mu^2} + 534\delta^4\mu^4\sqrt{\delta^2 - \mu^2} - 14\mu^{10} + 340\mu^8 \\
&\quad \left. - 1562\mu^6 + 1864\mu^4 - 524\mu^2 + 16 \right) \\
D_P &= 4 \left(\sqrt{\delta^2 - \mu^2} + 1 \right)^2 \left(\sqrt{\delta^2 - \mu^2} + 2 \right)^3 \left(\sqrt{\delta^2 - \mu^2} + \delta^2 - \mu^2 \right)^2 \times \\
&\quad \times \left(\sqrt{\delta^2 - \mu^2} - \delta^2 + \mu^2 + 2 \right)^2 \left(3\sqrt{\delta^2 - \mu^2} + \delta^2 - \mu^2 + 2 \right)^2
\end{aligned} \tag{278}$$

The behavior of these four functions is displayed in two pictures in Figure 4.

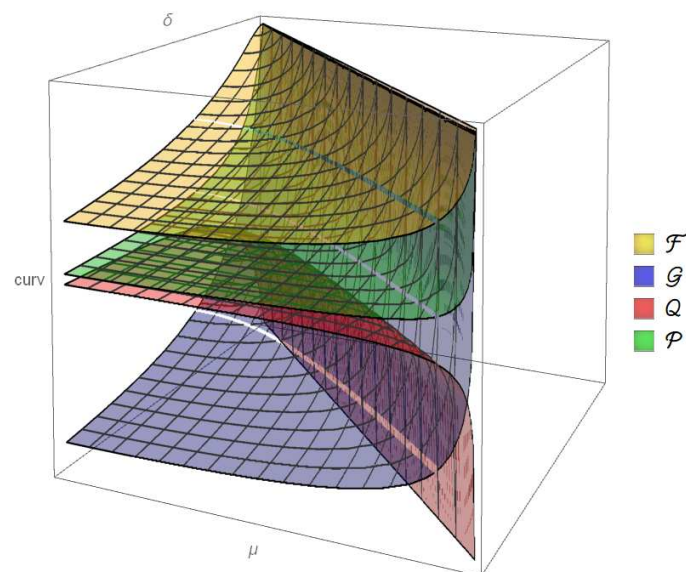


Figure 4. Cont.

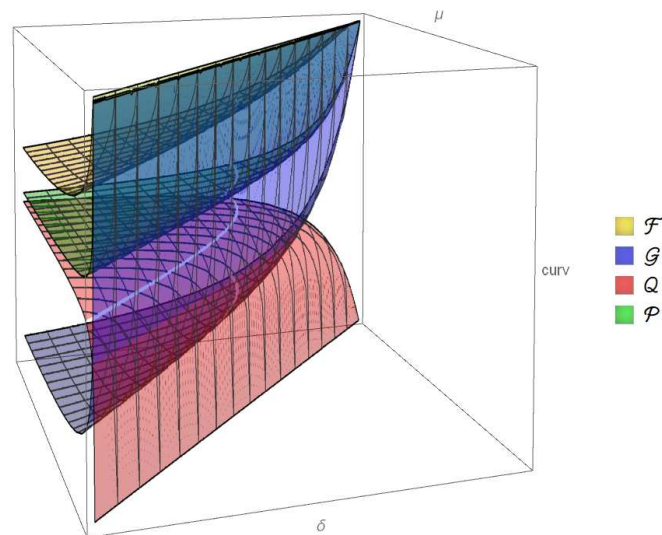


Figure 4. In the **first** and **second** pictures in this figure, we display the behavior of the four intrinsic curvature components $\mathcal{F}, \mathcal{G}, \mathcal{Q}, \mathcal{P}$ as respectively seen from the two sides of the vertical plane that has the line $\delta = \mu$ as a base. Such a line corresponds to the boundary of the cone in Figure 1. All the curvature components become singular on such a line.

As it becomes clear from the previous detailed discussion, the generalized thermodynamics à la Souriau yields a 3-dimensional temperature space equipped with a completely non-trivial metric which describes the distance between different Gibbs state probability distributions. The richness and non-triviality of this generalized thermodynamics is to be contrasted with the essentially trivial thermodynamics (Ideal-Gas-like) associated with the geodesic dynamical system and with any other conceivable integrable dynamical system. What one needs in Machine Learning algorithms are probability distributions on the very base manifold, constituting the mathematical model of the hidden layers. Probability distributions on the fibres of the tangent bundle are not that useful in this context.

6.3. Generalized Thermodynamics à la Souriau of the Siegel Half Plane $\mathbb{SH}_2$

The solvable coordinate description of the Siegel half-plane $\mathbb{SH}_2$ and its theory are presented in section 7.2 of the foundational paper [1] written by two of us together with Ugo Bruzzo. To that paper and to that section, we refer the reader for all the items we use here to derive the generalized thermodynamics à la Souriau of this manifold. First of all, we recall from the introduction to section 7.2 of [1] two conceptual points that are relevant from the perspective of Machine Learning applications.

In the four papers [1–4] by means of which the new paradigm of **Cartan Neural Networks** was elaborated and presented to the scientific community, we mainly focused on the series of U/H manifolds of the type:

$$\mathcal{M}^{[r,q]} \equiv \frac{\mathrm{SO}(r, r+q)}{\mathrm{SO}(r) \times \mathrm{SO}(r+q)} \quad (279)$$

In order to enlighten the role of non-maximal split manifolds having a non-trivial Tits Satake projection, it is conjectured to be a universal mechanism for clustering of data. In particular we analyzed the cases of non-compact rank $r = 1, 2$ and in [1], at the beginning of section 7.2, we wrote:

We focus on the case $r = 2$ of the series of symmetric manifolds (279) in a completely synoptic setup with respect to our previous treatment of the case $r = 1$. The motivation for this synopsis is twofold:

1. *On one hand we want to stress that the $r = 1$ case is completely aligned with all the subsequent $r > 1$ ones and that the Tits Satake projection, which went unnoticed and unexploited by the*

authors in [59–61], is actually the conceptual back-bone for all the members of the considered series of manifolds.

2. On the other hand we want to emphasize that the $r = 2$ and $r = 1$ cases are twins inside the entire series since their respective Tits Satake submanifolds $\mathcal{M}^{[1,1]}$, $\mathcal{M}^{[2,1]}$ are just the first and the second instance of a Siegel upper complex plane, which is the appropriate generalization of the Lobachevsky-Poincaré hyperbolic plane. Instead, for values $r > 2$, the Tits Satake submanifold, that, by definition, is always a maximally split symmetric space, is not a further instance of a Siegel upper complex plane. Indeed the appearance of the first two Siegel planes is strictly linked with the low rank sporadic isomorphisms of simple Lie algebras.

As we already noticed above from the point of view of Souriau generalized thermodynamics, the relevant U/H are the Kählerian ones, which essentially means the two infinite series mentioned in Equation (201). The second series (the Siegel planes) is made of maximally split symmetric manifolds, while the first series displays, for $q > 1$, the Tits Satake projection mechanism gives rise to non-trivial Paint Group invariances. Hence, if we want to join the Tits Satake projection with generalized thermodynamics à la Souriau, we have to choose the first series $\mathcal{M}^{[2,q]}$. Yet, due to low-dimensional Lie algebra isomorphism, the Tits Satake submanifold $\mathcal{M}^{[2,1]}$ for the TS universality class $\mathcal{M}^{[2,q]}$ is locally isomorphic (by double covering) to the second element of the second series, namely the Siegel plane $\mathbb{SH}_2$, as already emphasized in Equation (204). These preliminary observations are instrumental to understand the reason behind our following exposition which, summarizing and importing the results of section 7.2 of [1], emphasizes the double description of $\mathbb{SH}_2$ in terms of the 4-dimensional spinor representation of $\text{SO}(2,3)$, identified with the $\text{Sp}(4, \mathbb{R})$ fundamental representation, and in terms of the 5-dimensional vector, defining, representation of $\text{SO}(2,3)$. This is due to our interest in extending the results inherent to the Tits Satake submanifold to the full TS-universality class by relying on Paint Group covariance. Indeed, in Appendix D we provide a preliminary study of the generalized thermodynamics setup for the case $\mathcal{M}^{[2,2]}$ in order to emphasize the role of Paint Group in this context.

After these preliminary clarifications, we start the analysis of $\mathbb{SH}_2$ following section 7.2 of our foundational paper. The relation between the spinor and the vector representation that provides the local isomorphism of $\text{SO}(2,3)$ with $\text{Sp}(4, \mathbb{R})$ is given by the gamma-matrices $\Gamma_i, i = 1, \dots, 5$. Those well adapted to the triangular basis, which is that where the preserved metric with signature (2,3) is the η_t -matrix in Equation (7.34) of [1] are displayed in Equation (7.41) of the same paper. The charge conjugation matrix C_s that becomes the symplectic invariant form of $\text{Sp}(4, \mathbb{R})$ is displayed in Equation (7.37) of the same paper. The 10 generators J_{ij} of the $\mathfrak{so}(2,3) \simeq \mathfrak{sp}(4, \mathbb{R})$ Lie algebra are defined in Equation (7.39) of [1] as commutators of gamma matrices. The explicit expression of the double covering of $\text{SO}(2,3)$ group by means of $\text{Sp}(4, \mathbb{R})$ group is provided by Equation (7.44) of the same reference that we repeat here for the reader's convenience:

$$\forall S \in \text{Sp}(4, \mathbb{R}) \quad : \quad \mathcal{O}_j^i[S] \equiv \frac{1}{4} \text{Tr} \left(\Gamma_i^T S^{-1} \Gamma_j S \right) \in \text{SO}(2,3) \quad (280)$$

The fundamental next step is the construction of the solvable coset representative both in the vector and in the spinor representation. We utilize the notations and the results of [1] section 7.2. Hence, we use the letter **W** instead of **Y** for the vector of solvable coordinates, namely of parameters of the solvable Lie group metrically equivalent to U/H. Thinking of $\text{SO}(2,3)$ as the Tits Satake subgroup of $\text{SO}(2, 2 + 2s)$ the solvable coordinate vector is the one given in Equation (7.52) of [1], namely:

$$\mathbf{W} = \left\{ w_1, w_2, w_3, w_4, w_5, \underbrace{0, \dots, 0}_{(2s-1)}, w_6, \underbrace{0, \dots, 0}_{(2s-1)} \right\} \quad (281)$$

the zeros corresponding to the Tits Satake projection. The corresponding solvable group element constructed with the Σ -exponential map (see a discussion of its definition in [2]) in terms of the solvable coordinates is the following one:

$$\mathbb{L}(\mathbf{W}) = \begin{pmatrix} e^{w_1} & \frac{e^{w_1} w_3}{\sqrt{2}} & \frac{1}{2} e^{w_1} (w_3 w_6 + \sqrt{2} w_5) & \frac{1}{8} e^{w_1} (-\sqrt{2} w_3 w_6^2 + 4\sqrt{2} w_4 - 4w_5 w_6) & -\frac{1}{4} e^{w_1} (2w_3 w_4 + w_5^2) \\ 0 & e^{w_2} & \frac{e^{w_2} w_6}{\sqrt{2}} & -\frac{1}{4} e^{w_2} w_6^2 & -\frac{e^{w_2} w_4}{\sqrt{2}} \\ 0 & 0 & 1 & -\frac{w_6}{\sqrt{2}} & -\frac{w_5}{\sqrt{2}} \\ 0 & 0 & 0 & e^{-w_2} & -\frac{e^{-w_2} w_3}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & e^{-w_1} \end{pmatrix} \quad (282)$$

which is Equation (7.54) of [1]. Imposing the identification:

$$\mathcal{O}[\mathcal{S}_s(\mathbf{W})] = \mathbb{L}(\mathbf{W}) \quad (283)$$

One finds a unique solution for the $\mathcal{S}_s$ that is the following:

$$\mathcal{S}_s[\mathbf{W}] = \begin{pmatrix} e^{\frac{1}{2}(-w_1-w_2)} & \frac{1}{2} e^{\frac{1}{2}(w_2-w_1)} w_6 & \frac{1}{4} e^{\frac{1}{2}(w_1+w_2)} (w_5 w_6 - 2\sqrt{2} w_4) & \frac{1}{4} e^{\frac{1}{2}(w_1-w_2)} (2w_5 + \sqrt{2} w_3 w_6) \\ 0 & e^{\frac{1}{2}(w_2-w_1)} & \frac{1}{2} e^{\frac{1}{2}(w_1+w_2)} w_5 & \frac{e^{\frac{1}{2}(w_1-w_2)} w_3}{\sqrt{2}} \\ 0 & 0 & e^{\frac{1}{2}(w_1+w_2)} & 0 \\ 0 & 0 & -\frac{1}{2} e^{\frac{1}{2}(w_1+w_2)} w_6 & e^{\frac{1}{2}(w_1-w_2)} \end{pmatrix} \quad (284)$$

6.3.1. The Siegel Upper Plane

For completeness, we import from [1] the necessary concepts and formulae for the Siegel upper plane. This is particularly important in order to compare our results with those of Barbaresco et al. [30–34]. The *Siegel upper complex plane* of degree (or genus) g is the generalization to higher dimensions of the Lobachevsky–Poincaré hyperbolic plane.

Just as the standard hyperbolic plane with the Poincaré metric is a complex analytic realization of a maximally split symmetric space, namely $\mathrm{SL}(2, \mathbb{R})/\mathrm{SO}(2)$, in the same way, the upper Siegel plane of degree g is the complex analytic realization of the symmetric space:

$$\mathcal{M}_{\text{Siegel}} = \frac{\mathrm{Sp}(2g, \mathbb{R})}{\mathrm{S}[\mathrm{U}(1) \times \mathrm{U}(g)]} \quad (285)$$

The key observation is the following. Just in the same way as the *fractional linear transformation*:

$$z \rightarrow \tilde{z} \equiv \frac{az + b}{cz + d}; \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}(2, \mathbb{R}) \quad (286)$$

maps complex numbers z with strictly positive imaginary part into complex numbers $\tilde{z}$ with the same property, **the fractional linear matrix transformation**:

$$Z_{g \times g} \rightarrow \tilde{Z}_{g \times g} \equiv (A_{g \times g} Z_{g \times g} + B_{g \times g}) \cdot (C_{g \times g} Z_{g \times g} + D_{g \times g})^{-1}; \quad \left(\begin{array}{c|c} A_{g \times g} & B_{g \times g} \\ \hline C_{g \times g} & D_{g \times g} \end{array} \right) \in \mathrm{Sp}(2g, \mathbb{R}) \quad (287)$$

maps **complex symmetric matrices**:

$$Z_{g \times g} = X_{g \times g} + i Y_{g \times g}; \quad Z_{g \times g}^T = Z_{g \times g} \quad (288)$$

whose imaginary part $Y_{g \times g}$ is positive definite (namely has strictly positive eigenvalues) into **complex symmetric matrices** $\tilde{Z}_{g \times g}$ with the same property. The relations among the $g \times g$ blocks:

$$A^T C = C^T A; \quad B^T D = D^T B; \quad A^T D - C^T B = \mathbf{1} \quad (289)$$

following from the very definition of the $\mathrm{Sp}(2g, \mathbb{R})$ group, are instrumental in the lengthy yet straightforward proof of what was stated above.

The number of real components of $Z_{g \times g}$ exactly matches the dimension of the *maximally split symmetric space* defined in Equation (285) so that the upper Siegel plane constitutes its holomorphic realization. Furthermore, the choice of the Borel solvable subgroup inside $\mathrm{Sp}(2g, \mathbb{R})$ provides a convenient parameterization of the matrix $Z_{g \times g}$. Indeed, this latter is the orbit under the fractional linear action of the Borel subgroup of the special matrix $Z_0 = i \mathbf{1}_{g \times g}$.

Applying this idea to the case $g = r = 2$ which is ours, and utilizing the parameterization of the Borel solvable subgroup provided in Equation (284), we obtain:

$$\begin{aligned} Z &= X + iY \\ X &= \begin{pmatrix} \frac{1}{8} \left(w_6 (4w_5 + \sqrt{2}w_3w_6) - 4\sqrt{2}w_4 \right) & \frac{1}{4} (2w_5 + \sqrt{2}w_3w_6) \\ \frac{1}{4} (2w_5 + \sqrt{2}w_3w_6) & \frac{w_3}{\sqrt{2}} \end{pmatrix} \\ Y &= \begin{pmatrix} \frac{1}{4} e^{-w_1-w_2} (e^{2w_2}w_6^2 + 4) & \frac{1}{2} e^{w_2-w_1} w_6 \\ \frac{1}{2} e^{w_2-w_1} w_6 & e^{w_2-w_1} \end{pmatrix} \end{aligned} \quad (290)$$

According to Equation (287) the action of any element $g \in \mathrm{Sp}(4, \mathbb{R})$ on the coset manifold is the fractional linear transformation of the symmetric complex matrix:

$$Z = \begin{pmatrix} z & \omega \\ \omega & \zeta \end{pmatrix}; \quad z, \omega, \zeta \in \mathbb{C} \quad (291)$$

which represents the entire manifold. When Z is diagonal, namely when $\omega = 0$, the two remaining complex entries z, ζ represent the coordinates of two hyperbolic upper planes. The subgroup $\Gamma \subset \mathrm{Sp}(4, \mathbb{R})$ which respects diagonality, namely the condition $\omega = 0$ is $\Gamma = \mathrm{SL}(2, \mathbb{R}) \times \mathrm{SL}(2, \mathbb{R})$.

It is convenient to recall Equation (290), which provides the parameterization of the complex matrix (291) in terms of the solvable coordinates $w_1, w_2, w_3, w_4, w_5, w_6$ that can be summarized by stating:

$$\begin{aligned} z &= \frac{1}{8} \left(-4\sqrt{2}w_4 + w_6 (4w_5 + \sqrt{2}w_3w_6) + 2ie^{-w_1-w_2} (e^{2w_2}w_6^2 + 4) \right) \\ \zeta &= \frac{w_3}{\sqrt{2}} + ie^{w_2-w_1} \\ \omega &= \frac{1}{4} \left(2w_5 + 2ie^{w_2-w_1}w_6 + \sqrt{2}w_3w_6 \right) \end{aligned} \quad (292)$$

As one sees from Equation (292), the diagonalization condition of the matrix Z corresponds to setting $w_5 = w_6 = 0$, which implies that the other two solvable coordinates w_3, w_4 obtain the interpretation of real parts of the complex coordinates ζ and z , respectively.

The three complex numbers z, ζ, ω can be utilized as complex coordinates of the symmetric space and the Kähler metric can be derived from a suitable Kähler potential. Similarly, the Kählerian moment maps for all the Killing vector fields can be obtained from the Kähler potential, just as the Kähler 2-form. This, however, is not the best approach for our goals. In order to construct the generalized thermodynamics à la Souriau, it is much

more convenient to utilize real solvable coordinates and obtain the Kähler 2-form just as we did in the Poincaré case from the unique $U(1)$ generator. This is what we do in next subsection.

6.3.2. The Kähler 2-Form, the Killing Vector Fields, and the Moment Maps

In Order to construct all the items of generalized thermodynamics we need the vielbein, the Kähler 2-form and the moment maps of all Killing vectors. To this effect we need a well-normalized basis of generators of the full U Lie algebra. Such a basis is presented in two versions in the spinor representation in Table 1 and in the vector representation in Table 2.

Table 1. In this table, we display the complete list of the 10 generators of the $\mathfrak{sp}(4, R)$ Lie algebra.

$T_1^s = \begin{pmatrix} -\frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}$	$T_2^s = \begin{pmatrix} -\frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \end{pmatrix}$
$T_3^s = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix}$	$T_4^s = \begin{pmatrix} 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
$T_5^s = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix}$	$T_6^s = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \end{pmatrix}$
$T_7^s = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \end{pmatrix}$	$T_8^s = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \end{pmatrix}$
$T_9^s = \begin{pmatrix} 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \end{pmatrix}$	$T_{10}^s = \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \end{pmatrix}$

The generators in the two lists are in one-to-one correspondence and satisfy the Lie algebra commutation relations with the very same structure constants. The 10 generators, irrespectively of their label s or v are ordered in the following way:

$$T_{1,\dots,6}^{s/v} = K_i \quad (293)$$

are the 6 non-compact coset generators spanning the vector subspace $\mathbb{K}$ in the orthogonal decomposition

$$\mathbb{U} = \mathbb{H} \oplus \mathbb{K} \quad (294)$$

Furthermore, the first two generators $T_{1,2}^{s/v}$ are the two non-compact Cartan generators.

The generators:

$$T_{7,8,9}^{s/v} = H_{1,2,3} \quad (295)$$

are the generators of the $\mathfrak{su}(2) \simeq \mathfrak{so}(3)$ subalgebra of $\mathbb{H} = \mathfrak{su}(2) \oplus \mathfrak{u}(1)$.

Finally,

$$T_{10}^{s/v} = H_0 \quad (296)$$

is the $\mathfrak{u}(1) \simeq \mathfrak{so}(2)$ generator responsible for the Kähler structure.

Table 2. In this table, we display a complete list of the 10 generators of the $\mathfrak{so}(2,3)$ Lie algebra, which are in one-to-one correspondence with and in the same order as the generators of the $\mathfrak{sp}(4, R)$ Lie algebra listed in Table 1.

$$T_1^v = \sqrt{2} K_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \quad T_2^v = \sqrt{2} K_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$T_3^v = \sqrt{2} K_3 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix} \quad T_4^v = \sqrt{2} K_4 = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \end{pmatrix}$$

$$T_5^v = \sqrt{2} K_5 = \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix} \quad T_6^v = \sqrt{2} K_6 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$T_7^v = H_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad T_8^v = H_2 = \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix}$$

$$T_9^v = H_3 = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix} \quad T_{10}^v = H_0 = \begin{pmatrix} 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

Introducing the left-invariant 1-form in either the spinor or the vector form and projecting it onto the $\mathbb{K}$ -subspace, we find the sechsbein of the 6-dimensional space which turns out to be the same in the two cases, as it should. We choose the vector form and we write:

$$\begin{aligned}
\Theta^v &\equiv \mathbb{L}^{-1}(\mathbf{W}) \cdot d\mathbb{L}(\mathbf{W}) \\
\mathbf{e}^i &= \text{Tr}(\Theta^v \cdot K_i^\dagger) \\
\delta_{ij} &= \text{Tr}(K_i \cdot K_j^\dagger); \text{Tr}(H_i \cdot K_j^\dagger) = 0 \quad \text{normalization of the adjoints } K_j^\dagger \quad (297)
\end{aligned}$$

, and explicitly we have:

$$\begin{aligned}
\mathbf{e}^1 &= d\mathbf{w}_1 \\
\mathbf{e}^2 &= d\mathbf{w}_2 \\
\mathbf{e}^3 &= \frac{1}{2}(d\mathbf{w}_3 + d\mathbf{w}_1 w_3 - d\mathbf{w}_2 w_3) \\
\mathbf{e}^4 &= \frac{1}{8}\left(w_6^2(-(d\mathbf{w}_3 + d\mathbf{w}_1 w_3 - d\mathbf{w}_2 w_3)) - 2\sqrt{2}w_6(d\mathbf{w}_5 + d\mathbf{w}_1 w_5) + 4(d\mathbf{w}_4 + (d\mathbf{w}_1 + d\mathbf{w}_2)w_4)\right) \\
\mathbf{e}^5 &= \frac{1}{4}\left(2d\mathbf{w}_5 + 2d\mathbf{w}_1 w_5 + \sqrt{2}w_6(d\mathbf{w}_3 + d\mathbf{w}_1 w_3 - d\mathbf{w}_2 w_3)\right) \\
\mathbf{e}^6 &= \frac{1}{2}(d\mathbf{w}_6 + d\mathbf{w}_2 w_6) \quad (298)
\end{aligned}$$

In this case, the sechsbein coincide with the left-invariant 1-forms and satisfy the Maurer–Cartan equations of the solvable Lie group; however, this is not particularly relevant for our goals. The explicit form of the metric is given by:

$$ds_{\text{Siegel}}^2 = \sum_{i=1}^6 \mathbf{e}^i \times \mathbf{e}^i \quad (299)$$

The Kähler 2-form is obtained by constructing the adjoint representation of the generator H_0 on the subspace $\mathbb{K}$. Working in the vector representation and defining the normalized generators:

$$K_i = \frac{1}{\sqrt{2}} T_i^v \quad (i, 1, \dots, 6); \quad \text{Tr}(K_i \cdot K_j) = \delta_{ij} \quad (300)$$

we obtain

$$[H_0, K_i] = (\text{Adj}H_0)_{ij} K_j \quad (301)$$

where

$$\text{Adj}H_0 = \begin{pmatrix} 0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad (302)$$

and the Kähler 2-form takes the form:

$$\begin{aligned}
\mathcal{K} &= (\text{Adj}H_0)_{ij} \mathbf{e}^i \wedge \mathbf{e}^j \\
&= -\sqrt{2} \mathbf{e}^1 \wedge \mathbf{e}^3 + \sqrt{2} \mathbf{e}^1 \wedge \mathbf{e}^4 + \sqrt{2} \mathbf{e}^2 \wedge \mathbf{e}^3 + \sqrt{2} \mathbf{e}^2 \wedge \mathbf{e}^4 - 2 \mathbf{e}^5 \wedge \mathbf{e}^6 \quad (303)
\end{aligned}$$

The explicit expression of the Kähler 2 in the solvable coordinate basis is obtained from Equation (303) by substituting the explicit form of the sechsbein (298); so doing, one immediately verifies that the 2-form $\mathcal{K}$ is closed, namely $d\mathcal{K} = 0$.

Then, utilizing the vector representation and utilizing the method described in Section 6.1.1 we derive the explicit form of the 10 Killing vector fields expressed in terms of the solvable coordinate basis associated with the Lie algebra generators, as ordered and displayed in Table 2. We obtain the following result. The 6 Killing vector fields, generating the coset translations associated with the $\mathbb{K}$ -generators, have the following explicit form:

$$\begin{aligned}
 \mathfrak{k}_1 &= \partial_1 \\
 \mathfrak{k}_2 &= \partial_2 \\
 \mathfrak{k}_3 &= -\frac{1}{2}e^{w_1-w_2}w_3\partial_1 + \frac{1}{2}e^{w_1-w_2}w_3\partial_2 + \left(\frac{1}{2}e^{w_1-w_2}(w_3^2+2) + e^{w_2-w_1}\right)\partial_3 + \frac{1}{4}e^{w_1-w_2}(w_5-w_6)(w_5+w_6)\partial_4 \\
 &\quad - \frac{e^{w_1-w_2}w_6}{\sqrt{2}}\partial_5 + \frac{e^{w_1-w_2}w_5}{\sqrt{2}}\partial_6 \\
 \mathfrak{k}_4 &= -\frac{1}{2}e^{w_1+w_2}w_4\partial_1 + \frac{1}{4}e^{w_1+w_2}(\sqrt{2}w_5w_6-2w_4)\partial_2 + \frac{1}{4}e^{w_1+w_2}(w_5^2+\sqrt{2}w_3w_6w_5-w_6^2)\partial_3 \\
 &\quad + \left(\frac{1}{16}e^{w_1+w_2}(8w_4^2-4\sqrt{2}w_5w_6w_4+(w_6^2+4)^2) + e^{-w_1-w_2}\right)\partial_4 \\
 &\quad + \frac{e^{w_1+w_2}w_6(w_6^2+4)}{4\sqrt{2}}\partial_5 - \frac{e^{w_1+w_2}w_5(w_6^2+4)}{4\sqrt{2}}\partial_6 \\
 \mathfrak{k}_5 &= -\frac{1}{2}e^{w_1}w_5\partial_1 - \frac{e^{w_1}w_3w_6}{2\sqrt{2}}\partial_2 - \frac{e^{w_1}(w_3^2+2)w_6}{2\sqrt{2}}\partial_3 + \frac{e^{w_1}w_6(w_6^2+2w_3w_4+4)}{4\sqrt{2}}\partial_4 \\
 &\quad + \left(\frac{1}{4}e^{w_1}(w_5^2+2w_6^2+2w_3w_4+4) + e^{-w_1}\right)\partial_5 + \frac{e^{w_1}(w_3(w_6^2+4)-4w_4)}{4\sqrt{2}}\partial_6 \\
 \mathfrak{k}_6 &= -\frac{1}{2}e^{w_2}w_6\partial_2 - \frac{1}{2}e^{w_2}(\sqrt{2}w_5+w_3w_6)\partial_3 + \left(\frac{e^{-w_2}w_5}{\sqrt{2}} + \frac{1}{2}e^{w_2}w_4w_6\right)\partial_4 \\
 &\quad + \frac{e^{-w_2}(e^{2w_2}w_4-w_3)}{\sqrt{2}}\partial_5 + \left(\frac{1}{4}e^{w_2}(w_6^2+4) + e^{-w_2}\right)\partial_6
 \end{aligned} \tag{304}$$

The 3 Killing vector fields closing the $\mathfrak{su}(2)$ Lie subalgebra of the isotropy algebra $\mathbb{H}$ have the following explicit form:

$$\begin{aligned}
 \mathfrak{k}_7 &= \frac{1}{2}e^{w_2}w_6\partial_2 + \frac{1}{2}e^{w_2}(\sqrt{2}w_5+w_3w_6)\partial_3 + \left(\frac{e^{-w_2}w_5}{\sqrt{2}} - \frac{1}{2}e^{w_2}w_4w_6\right)\partial_4 \\
 &\quad - \frac{e^{-w_2}(w_3+e^{2w_2}w_4)}{\sqrt{2}}\partial_5 + \left(e^{-w_2} - \frac{1}{4}e^{w_2}(w_6^2+4)\right)\partial_6 \\
 \mathfrak{k}_8 &= \frac{1}{2}e^{w_1}w_5\partial_1 + \frac{e^{w_1}w_3w_6}{2\sqrt{2}}\partial_2 + \frac{e^{w_1}(w_3^2+2)w_6}{2\sqrt{2}}\partial_3 - \frac{e^{w_1}w_6(w_6^2+2w_3w_4+4)}{4\sqrt{2}}\partial_4 \\
 &\quad + \left(e^{-w_1} - \frac{1}{4}e^{w_1}(w_5^2+2w_6^2+2w_3w_4+4)\right)\partial_5 + \frac{e^{w_1}(4w_4-w_3(w_6^2+4))}{4\sqrt{2}}\partial_6 \\
 \mathfrak{k}_9 &= \frac{e^{w_1-w_2}(w_3+e^{2w_2}w_4)}{2\sqrt{2}}\partial_1 + \frac{1}{4}e^{w_1-w_2}(e^{2w_2}(\sqrt{2}w_4-w_5w_6)-\sqrt{2}w_3)\partial_2 \\
 &\quad + \frac{e^{-w_1-w_2}(e^{2w_2}(4-e^{2w_1}(w_5^2+\sqrt{2}w_3w_6w_5-w_6^2))-2e^{2w_1}(w_3^2+2))}{4\sqrt{2}}\partial_3 \\
 &\quad + \frac{e^{-w_1-w_2}(e^{2w_1}(-4w_5^2+4w_6^2-e^{2w_2}(8w_4^2-4\sqrt{2}w_5w_6w_4+(w_6^2+4)^2))+16)}{16\sqrt{2}}\partial_4 \\
 &\quad + \frac{1}{8}e^{w_1-w_2}w_6(4-e^{2w_2}(w_6^2+4))\partial_5 + \frac{1}{8}e^{w_1-w_2}w_5(e^{2w_2}(w_6^2+4)-4)\partial_6
 \end{aligned} \tag{305}$$

Finally, the Killing vector field associated with the $\mathfrak{u}(1)$ subalgebra of the $\mathbb{H}$ isotropy algebra has the following explicit form:

$$\begin{aligned}
\mathfrak{k}_{10} = & \frac{(e^{w_1-w_2}w_3 - e^{w_1+w_2}w_4)}{2\sqrt{2}} \partial_1 - \frac{1}{4}e^{w_1-w_2} \left(\sqrt{2}w_3 + e^{2w_2}(\sqrt{2}w_4 - w_5w_6) \right) \partial_2 \\
& + \frac{e^{-w_1-w_2} \left(e^{2w_2} \left(e^{2w_1} \left(w_5^2 + \sqrt{2}w_3w_6w_5 - w_6^2 \right) + 4 \right) - 2e^{2w_1}(w_3^2 + 2) \right)}{4\sqrt{2}} \partial_3 \\
& + \frac{e^{-w_1-w_2} \left(e^{2w_1} \left(-4w_5^2 + 4w_6^2 + e^{2w_2} \left(8w_4^2 - 4\sqrt{2}w_5w_6w_4 + (w_6^2 + 4)^2 \right) \right) - 16 \right)}{16\sqrt{2}} \partial_4 \\
& + \frac{1}{8}e^{w_1-w_2}w_6 \left(e^{2w_2} \left(w_6^2 + 4 \right) + 4 \right) \partial_5 - \frac{1}{8}e^{w_1-w_2}w_5 \left(e^{2w_2} \left(w_6^2 + 4 \right) + 4 \right) \partial_6
\end{aligned} \quad (306)$$

Last but not least, we need the moment maps associated with each of the above Killing vector fields. To this effect, we utilize the general method described in Section 6.1.2 and we apply Formula (230). Hence, we write:

$$\mathfrak{P}_\Lambda(\mathbf{W}) = \frac{1}{2} \text{Tr} \left(T_{10}^v \cdot \mathbb{L}^{-1}(\mathbf{W}) \cdot T_\Lambda^v \cdot \mathbb{L}(\mathbf{W}) \right); \quad \Lambda = 1, \dots, 10 \quad (307)$$

and the explicit result that we obtain is displayed below:

$\mathfrak{P}_1$	$= \frac{1}{16} \left(4\sqrt{2}w_4 - 4w_5w_6 - \sqrt{2}w_3(w_6^2 + 4) \right)$
$\mathfrak{P}_2$	$= \frac{4w_4 + w_3(w_6^2 + 4)}{8\sqrt{2}}$
$\mathfrak{P}_3$	$= \frac{1}{32} \left(e^{w_1-w_2} \left(\sqrt{2}(w_6^2 + 4)w_3^2 + 4w_5w_6w_3 + 2\sqrt{2}(w_5^2 + 4) \right) - 2\sqrt{2}e^{w_2-w_1}(w_6^2 + 4) \right)$
$\mathfrak{P}_4$	$= \frac{1}{64}e^{-w_1-w_2} \left(16\sqrt{2} - e^{2(w_1+w_2)} \left(8\sqrt{2}w_4^2 - 8w_5w_6w_4 + \sqrt{2}(w_5^2 + 4)(w_6^2 + 4) \right) \right)$
$\mathfrak{P}_5$	$= \frac{1}{32}e^{w_1} \left(-4\sqrt{2}w_4w_5 + \sqrt{2}w_3(w_6^2 + 4)w_5 - 4w_3w_4w_6 + 2(w_5^2 - 4)w_6 \right) - \frac{1}{4}e^{-w_1}w_6$
$\mathfrak{P}_6$	$= \frac{1}{16}e^{-w_2} \left(2\sqrt{2}(w_3 - e^{2w_2}w_4)w_6 + w_5(e^{2w_2}(w_6^2 + 4) + 4) \right)$
$\mathfrak{P}_7$	$= \frac{1}{16}e^{-w_2} \left(2\sqrt{2}(w_3 + e^{2w_2}w_4)w_6 + w_5(4 - e^{2w_2}(w_6^2 + 4)) \right)$
$\mathfrak{P}_8$	$= -\frac{1}{4}e^{-w_1}w_6 - \frac{1}{32}e^{w_1} \left(-4\sqrt{2}w_4w_5 + \sqrt{2}w_3(w_6^2 + 4)w_5 - 4w_3w_4w_6 + 2(w_5^2 - 4)w_6 \right)$
$\mathfrak{P}_9$	$= \frac{1}{64}e^{-w_1-w_2} \left[-4e^{2w_2}(w_6^2 + 4) - 2e^{2w_1} \left((w_6^2 + 4)w_3^2 + 2\sqrt{2}w_5w_6w_3 + 2(w_5^2 + 4) \right) \right. \\ \left. + e^{2(w_1+w_2)} \left(8w_4^2 - 4\sqrt{2}w_5w_6w_4 + (w_5^2 + 4)(w_6^2 + 4) \right) + 16 \right]$
$\mathfrak{P}_{10}$	$= \frac{1}{64}e^{-w_1-w_2} \left[-4e^{2w_2}(w_6^2 + 4) - 2e^{2w_1} \left((w_6^2 + 4)w_3^2 + 2\sqrt{2}w_5w_6w_3 + 2(w_5^2 + 4) \right) \right. \\ \left. - e^{2(w_1+w_2)} \left(8w_4^2 - 4\sqrt{2}w_5w_6w_4 + (w_5^2 + 4)(w_6^2 + 4) \right) - 16 \right]$

(308)

In Equation (308), we have separated the moment-maps into the three groups. The first group of six are the moment maps of the $\mathbb{K}$ translations. The subsequent group of three yields the moment maps of the $\mathfrak{su}(2)$ compact generators, while the last group of just one is the moment map of the $\mathfrak{u}(1)$ generator associated with the Kähler structure.

7. On the Partition Function and Gibbs Distributions in General and for SH_2 in Particular

Before addressing the determination of the partition function and the Gibbs distributions for the case of the Siegel half plane, which is now accessible, since we have prepared all the necessary instruments, we pause for a moment to consider a very important general property of the temperature vectors β . We might have anticipated the forthcoming discussion from previous sections, yet we chose to postpone it to the present juncture since we are now in a favorable position to illustrate it with a concrete and non-trivial example.

Once again, the metric equivalence of the non-compact symmetric spaces with a solvable Lie group manifold $\mathcal{S}_{U/H}$ and the double synergic decompositions (205) and (206) play an essential role.

The solvable Lie algebra generators can always be reexpressed as suitable linear combinations of the K_i coset generators and of the H_α compact subalgebra generators. Let us name $\mathcal{T}_i$ ($i = 1, \dots, 6$) the solvable Lie algebra generators in our $\mathbb{SH}_2$ case. With reference to Table 2 we have:

$$\mathcal{T}_i = \underbrace{\mathcal{Q}_{ij}}_{6 \times 6} K_j + \underbrace{\mathcal{Q}_{i\alpha}}_{6 \times 4} H_\alpha \quad : \quad \begin{cases} \mathcal{T}_1 = K_1 \\ \mathcal{T}_2 = K_2 \\ \mathcal{T}_3 = \frac{1}{\sqrt{2}} H_0 + \frac{1}{\sqrt{2}} H_3 + K_3 \\ \mathcal{T}_4 = -\frac{1}{\sqrt{2}} H_0 + \frac{1}{\sqrt{2}} H_3 + K_4 \\ \mathcal{T}_5 = H_2 + K_5 \\ \mathcal{T}_6 = H_1 + K_6 \end{cases} \quad (309)$$

The same linear transformation applies to the moment maps and allows to find the moment maps of the Killing vectors associated to the solvable Lie algebra generators. Hence instead of using the basis $(\mathbb{K}, \mathbb{H})$ for the moment-maps and the temperature vector β , we can use the basis $(\text{Solv}, \mathbb{H})$ and write the argument of the exponential in the partition function as

$$\hat{\beta} \cdot \hat{\mathfrak{P}}(\mathbf{W}) = \hat{\beta}^\Lambda \hat{\mathfrak{P}}_\Lambda(\mathbf{W}) \quad (310)$$

where now

$$\begin{aligned} \hat{\mathfrak{P}}_\Lambda(\mathbf{W}) &= \frac{1}{2} \text{Tr} \left(H_0 \cdot \mathbb{L}^{-1}(\mathbf{W}) \cdot \hat{T}_\Lambda \cdot \mathbb{L}(\mathbf{W}) \right); \quad \Lambda = 1, \dots, 10 \\ \hat{T}_\Lambda &= \{ \mathcal{T}_1, \dots, \mathcal{T}_6, H_1, \dots, H_4 \} \end{aligned} \quad (311)$$

Consider next the formal definition of the partition function

$$Z(\hat{\beta}) = \int \exp \left[-\hat{\beta} \cdot \hat{\mathfrak{P}}(\mathbf{W}) \right] \mu(\mathbf{W}) \quad (312)$$

$$\mu(\mathbf{W}) \equiv \mathcal{K} \wedge \mathcal{K} \wedge \mathcal{K} \simeq \underbrace{\sqrt{\det[g(\mathbf{W})]}}_{\text{just a constant}} d^6 \mathbf{W} \quad (313)$$

where $\mu(\mathbf{W})$ in Equation (313) is the integration measure that is invariant with respect to isometries of the coset manifold U/H . Recalling next the metric equivalence of the coset manifold with the solvable Lie group $\mathcal{S}_{U/H}$ that has free transitive action on the base manifold, we consider any group element $s_{\mathbf{u}} \in \mathcal{S}_{U/H}$ whose corresponding parameters we name $\mathbf{U}$. By definition, the abstract group element is represented in five dimensions by the matrix $\mathbb{L}(\mathbf{U})$, and we have:

$$s_{\mathbf{u}} : \mathbb{L}(\mathbf{W}) \longrightarrow \mathbb{L}(\mathbf{U}) \cdot \mathbb{L}(\mathbf{W}) = \mathbb{L}(s_{\mathbf{u}}(\mathbf{W})) \quad (314)$$

Since the integration measure is invariant under the solvable Lie group translations that are isometries, we can change integration variables and write:

$$Z(\hat{\beta}) = \int \exp \left[-\hat{\beta} \cdot \hat{\mathfrak{P}}(s_{\mathbf{u}}(\mathbf{W})) \right] \mu(s_{\mathbf{u}}(\mathbf{W})) = \int \exp \left[-\hat{\beta} \cdot \hat{\mathfrak{P}}(s_{\mathbf{u}}(\mathbf{W})) \right] \mu(\mathbf{W}) \quad (315)$$

Focusing on the argument of the exponential integrand, we have:

$$\begin{aligned}\widehat{\beta} \cdot \widehat{\mathfrak{P}}(\mathfrak{s}_{\mathbf{u}}(\mathbf{W})) &= \widehat{\beta}^{\Lambda} \times \frac{1}{2} \operatorname{Tr} \left(H_0 \cdot \mathbb{L}^{-1}(\mathbf{W}) \cdot \mathbb{L}^{-1}(\mathbf{U}) \cdot \widehat{T}_{\Lambda} \cdot \mathbb{L}(\mathbf{U}) \cdot \mathbb{L}(\mathbf{W}) \right) \\ &= \widehat{\beta}^{\Lambda} \times \operatorname{Adj}(\mathfrak{s}_{\mathbf{u}})_{\Lambda}^{\Sigma} \times \frac{1}{2} \operatorname{Tr} \left(H_0 \cdot \mathbb{L}^{-1}(\mathbf{W}) \cdot \widehat{T}_{\Sigma} \cdot \mathbb{L}(\mathbf{W}) \right) \\ &= \widehat{\beta}^{\Lambda} \times \operatorname{Adj}(\mathfrak{s}_{\mathbf{u}})_{\Lambda}^{\Sigma} \widehat{\mathfrak{P}}_{\Sigma}(\mathbf{W})\end{aligned}\quad (316)$$

The conclusion of the above formal calculation is that the partition function of generalized thermodynamics à la Souriau on Kähler non-compact symmetric spaces U/H has the following extremely important symmetry:

$$\begin{aligned}Z(\widehat{\beta}) &= Z(\operatorname{Adj}^T(\mathfrak{s}) \cdot \widehat{\beta}); \quad \forall \mathfrak{s} \in \mathcal{S}_{U/H} \\ \left(\operatorname{Adj}^T(\mathfrak{s}) \cdot \widehat{\beta} \right)^{\Sigma} &\equiv \widehat{\beta}^{\Lambda} \operatorname{Adj}(\mathfrak{s})_{\Lambda}^{\Sigma}\end{aligned}\quad (317)$$

The relevance of the above symmetry is that the co-adjoint action (transpose) of the solvable Lie group can always rotate a generic temperature vector β into a new one β_c that has non-vanishing components only along the compact subalgebra generators H_{α} , as we might explicitly illustrate for the Siegel case $\mathbb{SH}_2$, but we skip the somewhat lengthy calculations. This is not the end of the story. There is still the symmetry with respect to the compact isotropy subgroup that we can utilize. Consider the partition function reduced to a compact temperature β_c , namely:

$$\begin{aligned}Z(\beta_c) &= \int \exp[-\beta_c^{\alpha} \mathfrak{P}_{\alpha}(\mathbf{W})] \mu(\mathbf{W}) \\ \beta_c^{\alpha} \mathfrak{P}_{\alpha}(\mathbf{W}) &= \sum_{\alpha=1}^{\dim \mathbb{H}} \beta_c^{\alpha} \times \frac{1}{2} \operatorname{Tr} \left(H_0 \cdot \mathbb{L}^{-1}(\mathbf{W}) \cdot H_{\alpha} \cdot \mathbb{L}(\mathbf{W}) \right)\end{aligned}\quad (318)$$

The transformation of the compact isotropy subgroup H are just isometries as all other transformations of U , hence they leave the integration measure $\mu(\mathbf{W})$ invariant and act on the solvable coset representative $\mathbb{L}(\mathbf{W})$ in the canonical way as follows:

$$\forall h \in H; \quad h \cdot \mathbb{L}(\mathbf{W}) = \mathbb{L}(\mathfrak{h}(\mathbf{W})) \cdot \mathfrak{H}(h, \mathbf{W}); \quad \mathfrak{H}(h, \mathbf{W}) \in H \quad (319)$$

where $\mathfrak{h}(\mathbf{W})$ are the new solvable coordinates after the h -isometry transformation and $\mathfrak{H}(h, \mathbf{W})$ is the H -compensator, which, by definition, also lies in H and depends both on the point $\mathbf{W}$ and on the chosen h group element.

With the same strategy utilized above, we change the integration variable $\mathbf{W} \rightarrow \mathfrak{h}(\mathbf{W})$, and we rewrite:

$$\begin{aligned}Z(\beta_c) &= \int \exp[-\beta_c^{\alpha} \mathfrak{P}_{\alpha}(\mathfrak{h}(\mathbf{W}))] \mu(\mathfrak{h}(\mathbf{W})) \\ &= \int \exp[-\beta_c^{\alpha} \mathfrak{P}_{\alpha}(\mathfrak{h}(\mathbf{W}))] \mu(\mathbf{W})\end{aligned}\quad (320)$$

Next, using Equations (318) and (319), we obtain:

$$\beta_c^{\alpha} \mathfrak{P}_{\alpha}(\mathfrak{h}(\mathbf{W})) = \sum_{\alpha=1}^{\dim \mathbb{H}} \beta_c^{\alpha} \times \frac{1}{2} \operatorname{Tr} \left(H_0 \cdot \mathfrak{H}(h, \mathbf{W}) \cdot \mathbb{L}^{-1}(\mathbf{W}) \cdot h^{-1} \cdot H_{\alpha} \cdot h \cdot \mathbb{L}(\mathbf{W}) \cdot \mathfrak{H}^{-1}(h, \mathbf{W}) \right) \quad (321)$$

$$= \beta_c^{\alpha} \operatorname{Adj}(h)_{\alpha}^{\beta} \mathfrak{P}_{\beta}(\mathbf{W}) \quad (322)$$

In order to establish the equality of the r.h.s in (321) with that in (322) we utilized three properties:

1. The cyclic invariance of the trace.
2. The crucial fact that H_0 is the center of the compact Lie algebra $\mathbb{H}$, which is the very reason why the considered manifold is Kählerian, so that H_0 is invariant against any adjoint transformation of H group.
3. The adjoint H representation in the space of its Lie algebra $\mathbb{H}$:

$$\forall h \in H \quad : \quad h^{-1} \cdot H_\alpha \cdot h = \text{Adj}(h)_\alpha^\beta H_\beta \quad (323)$$

The final crucial fact is that by means of a suitable h -transformation we can always bring any $\mathbb{H}$ Lie algebra element into the Cartan subalgebra $\mathcal{C} \subset \mathbb{H}$ and then reduce the β_c to β_c^0 that has non vanishing component only along the Cartan generators, one being H_0 the remaining ones being the Cartan generators of H' in the decomposition (30).

7.1. Canonical Form of the Partition Functions and of the Gibbs Probability Distributions, in General

In this way, the space of allowed temperatures turns out to be the U (co)adjoint orbit of a proper subset $\Omega_c \mathcal{C} \subset \mathbb{H}$ of the compact Cartan subalgebra $\mathcal{C}$. Ω_c is provided by those compact Cartan temperatures that have the correct sign in order to guarantee convergence of the Gaussian integrals. As we advocate further on, for the non-maximally split manifolds, where the solvable Lie algebra admits the Paint-Group automorphism group, which is a proper subgroup $G^{\text{Paint}} \subset H \subset U$, the essential Cartan temperatures might be further reduced. We postpone this study to the next publication and confine ourselves to the preliminary observations of Appendix D relative to the case $\mathcal{M}^{[2,2]}$.

Learning the lesson that the temperature vector β can be reduced to its minimal Cartan form β_0 by means of U -isometry transformations, the general form of the Gibbs probability distributions becomes very simple and clear. It always contains as many parameters as the dimension of the U Lie group but it can be written in the following very compact and useful way:

$$G(\beta_0; g \mid \mathbf{W}) \equiv \frac{\exp \left[- \sum_{i=0}^{\ell=\text{rank } \mathbb{H}'} \beta_0^i \mathfrak{P}_i(g[\mathbf{W}]) \right]}{Z(\beta^0)}$$

β_0^i = temperatures associated with the compact Cartan generators $H_{0,1,\dots,\ell}^c$ of $\mathbb{H}$
 $\mathfrak{P}_i(\mathbf{W})$ = moment maps of the compact Cartan generators $H_{0,1,\dots,\ell}^c$ of $\mathbb{H}$
 g = any group element in U
 $g[\mathbf{W}]$ = new solvable parameters after a g -isometry transformation

(324)

In other words, the Cartan temperatures define the Gibbs probability distribution centered around the origin of the coset manifold, namely the identity of the metrically equivalent solvable group. The other temperatures simply would rotate such a distribution (H -transformations) or translate it to be centered around any other point of the manifold (solvable Lie group transformations). Hence instead of introducing such parameters we can simply evaluate the original Gibbs distribution in a transformed point. Both for analytic calculations and for practical applications in Machine Learning this view-point is extremely useful. Stated differently, when reducing the temperature vector to the compact Cartan subalgebra, the suppressed β -parameters are replaced by the parameters of a generic U -transformation g of the point, appearing in the argument of the exponential distribution.

7.2. Calculation of the Partition Function for the Siegel Plane in Canonical Form

Having clarified that we just need two compact temperatures associated with two Cartan generators we choose as second Cartan generator the one listed as $H_3 = T_9^v$ in Table 2. This means that using Equation (308) for the explicit form of the moment maps, the argument of the exponential in the partition function integrand is the following:

$$\begin{aligned}\mathfrak{A} \equiv \beta_0 \cdot \mathfrak{P}(\mathbf{W}) &= \mu \mathfrak{P}_9(\mathbf{W}) + \lambda \mathfrak{P}_{10}(\mathbf{W}) \\ &= \frac{1}{64} e^{-w_1 - w_2} \left(\lambda \left(-4e^{2w_2} (w_6^2 + 4) - 2e^{2w_1} \left((w_6^2 + 4)w_3^2 + 2\sqrt{2}w_5w_6w_3 + 2(w_5^2 + 4) \right) \right. \right. \\ &\quad \left. \left. - e^{2(w_1 + w_2)} \left(8w_4^2 - 4\sqrt{2}w_5w_6w_4 + (w_5^2 + 4)(w_6^2 + 4) \right) - 16 \right) \right. \\ &\quad \left. + \mu \left(-4e^{2w_2} (w_6^2 + 4) - 2e^{2w_1} \left((w_6^2 + 4)w_3^2 + 2\sqrt{2}w_5w_6w_3 + 2(w_5^2 + 4) \right) \right. \right. \\ &\quad \left. \left. + e^{2(w_1 + w_2)} \left(8w_4^2 - 4\sqrt{2}w_5w_6w_4 + (w_5^2 + 4)(w_6^2 + 4) \right) + 16 \right) \right) \end{aligned} \quad (325)$$

where we have named $\beta_9 = \mu$ and $\beta_{10} = \lambda$. For calculation convenience it is useful to redefine $w_{1,2} = \log[\rho_{1,2}]$. In this way we get:

$$\begin{aligned}\mathfrak{A} &= \frac{N_A}{D_A} \\ N_A &= \rho_1^2 \left(- \left(\rho_2^2 \left(8w_4^2 - 4\sqrt{2}w_5w_6w_4 + (w_5^2 + 4)(w_6^2 + 4) \right) (\lambda - \mu) \right. \right. \\ &\quad \left. \left. + 2 \left((w_6^2 + 4)w_3^2 + 2\sqrt{2}w_5w_6w_3 + 2(w_5^2 + 4) \right) (\lambda + \mu) \right) \right) - 4 \left(4(\lambda - \mu) + \rho_2^2 (w_6^2 + 4) (\lambda + \mu) \right) \\ D_A &= 64\rho_1\rho_2 \end{aligned} \quad (326)$$

We have to calculate the 6-integrals of $\exp[\mathfrak{A}]$ on the nilpotent coordinate w_3, w_4, w_5, w_6 and finally on ρ_1, ρ_2 . We begin with the w_3 integral that is perfectly Gaussian and imposes the convergence condition:

$$\lambda + \mu > 0 \quad (327)$$

Next, we calculate also the integrand on w_4 which is again Gaussian and imposes the second convergence condition:

$$\lambda - \mu > 0 \quad (328)$$

We get

$$\begin{aligned}\int_{-\infty}^{\infty} dw_4 \int_{-\infty}^{\infty} dw_3 \exp[-\mathfrak{A}] &= \frac{16\pi e^{-\mathfrak{B}}}{\rho_1 \sqrt{\lambda - \mu} \sqrt{(w_6^2 + 4)(\lambda + \mu)}} \\ \mathfrak{B} &= \frac{N_B}{64\rho_1\rho_2} \\ N_B &= 16\lambda - 16\mu + \rho_2^2 \rho_1^2 w_5^2 w_6^2 (-(\lambda - \mu)) + 4\rho_2^2 (w_6^2 + 4)(\lambda + \mu) \\ &\quad + \frac{\rho_1^2 (\rho_2^2 (w_5^2 + 4)(w_6^2 + 4)^2 (\lambda - \mu) + 16(w_5^2 + w_6^2 + 4)(\lambda + \mu))}{w_6^2 + 4} \end{aligned} \quad (329)$$

The next integration on the nilpotent coordinate w_5 is just Gaussian and it does not impose further constraints on the two temperatures μ and λ . We get the following:

$$\int_{-\infty}^{\infty} dw_5 \frac{16\pi e^{-\mathfrak{B}}}{\rho_1 \sqrt{\lambda - \mu} \sqrt{(w_6^2 + 4)(\lambda + \mu)}} = \mathfrak{C}$$

$$\mathfrak{C} = \frac{64\pi^{3/2} \exp\left(-\frac{4(\rho_1^2(\lambda + \mu) + \lambda - \mu) + \rho_2^2(w_6^2 + 4)(\rho_1^2(\lambda - \mu) + \lambda + \mu)}{16\rho_1\rho_2}\right)}{\rho_1 \sqrt{(w_6^2 + 4)(\lambda - \mu)(\lambda + \mu)} \sqrt{\rho_2\rho_1(\lambda - \mu) + \frac{4\rho_1(\lambda + \mu)}{\rho_2(w_6^2 + 4)}}} \quad (330)$$

The fourth integration on the nilpotent variable w_6 can also be performed and yields an analytical result:

$$\mathfrak{F}(\rho_1, \rho_2, \lambda, \mu) \equiv \int_{-\infty}^{\infty} dw_6 \mathfrak{C}$$

$$= \frac{64\pi^{3/2} \sqrt{\lambda - \mu} \exp\left(-\frac{\lambda^2 + (\lambda - \mu)(\rho_1^2(\rho_2^2(\lambda - \mu) + \lambda + \mu) + \rho_2^2(\lambda + \mu)) - 6\lambda\mu + \mu^2}{8\rho_1\rho_2(\lambda - \mu)}\right) K_0\left(\frac{((\lambda - \mu)\rho_1^2 + \lambda + \mu)((\lambda - \mu)\rho_2^2 + \lambda + \mu)}{8(\lambda - \mu)\rho_1\rho_2}\right)}{\sqrt{\rho_2(\lambda + \mu)} \left(\frac{\rho_1(\lambda - \mu)}{\rho_2^2(\lambda - \mu) + \lambda + \mu}\right)^{3/2} (\rho_2^2(\lambda - \mu) + \lambda + \mu)^{3/2}} \quad (331)$$

where $K_0(x)$ is the Bessel function of type K and index 0.

Unfortunately the last two integrals on the remaining two variables $\rho_{1,2}$ or better on their logarithms $w_{1,2}$ cannot be done analytically and one has to perform them numerically introducing in this way compiled functions. By plotting the integrand we can however very easily verify that it always decays exponentially to zero in all directions so that the integral is always convergent (see Figure 5).

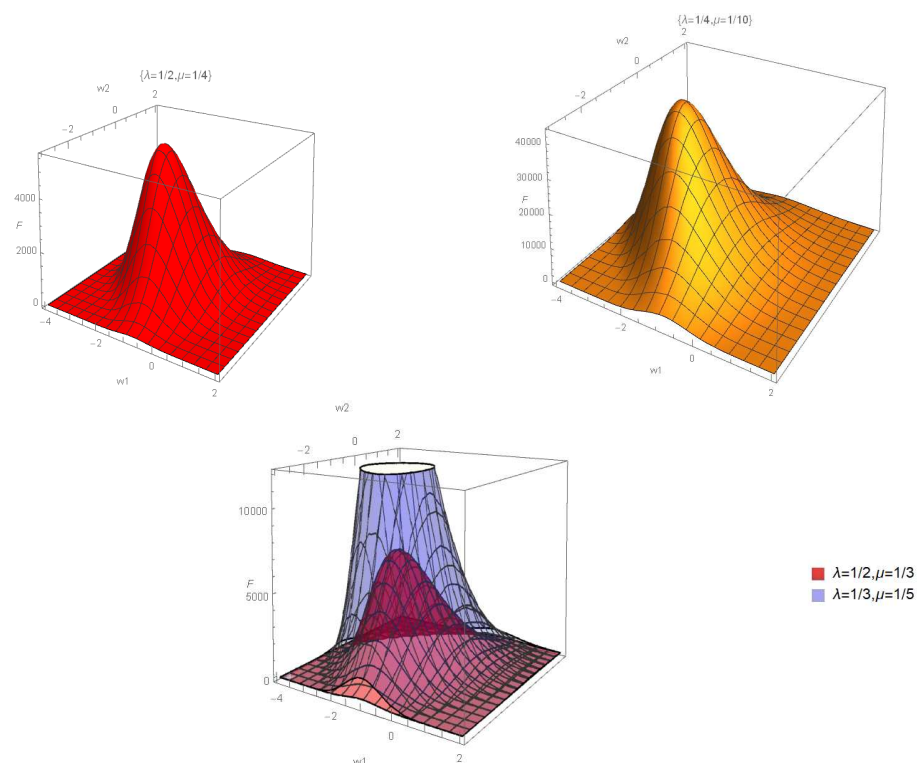


Figure 5. In this figure, we show for a few pairs of values of λ and μ the integrand $\mathfrak{F}$ in the integration variables $w_{1,2} = \log \rho_{1,2}$. The bell shape and the uniform exponential decay to zero at infinity in all directions guarantee the convergence of the two remaining integrals on $w_{1,2}$.

For this reason, one can define the partition function as a compiled function by performing numerically on a computer the last two integrals:

$$Z(\lambda, \mu) = \int_{-\infty}^{\infty} dw_1 \int_{-\infty}^{\infty} dw_2 \mathfrak{F}(\exp[w_1], \exp[w_2], \lambda, \mu) \quad (332)$$

In Figure 6, we display a plot of the partition function and of minus its logarithm, namely of the stochastic Hamiltonian.

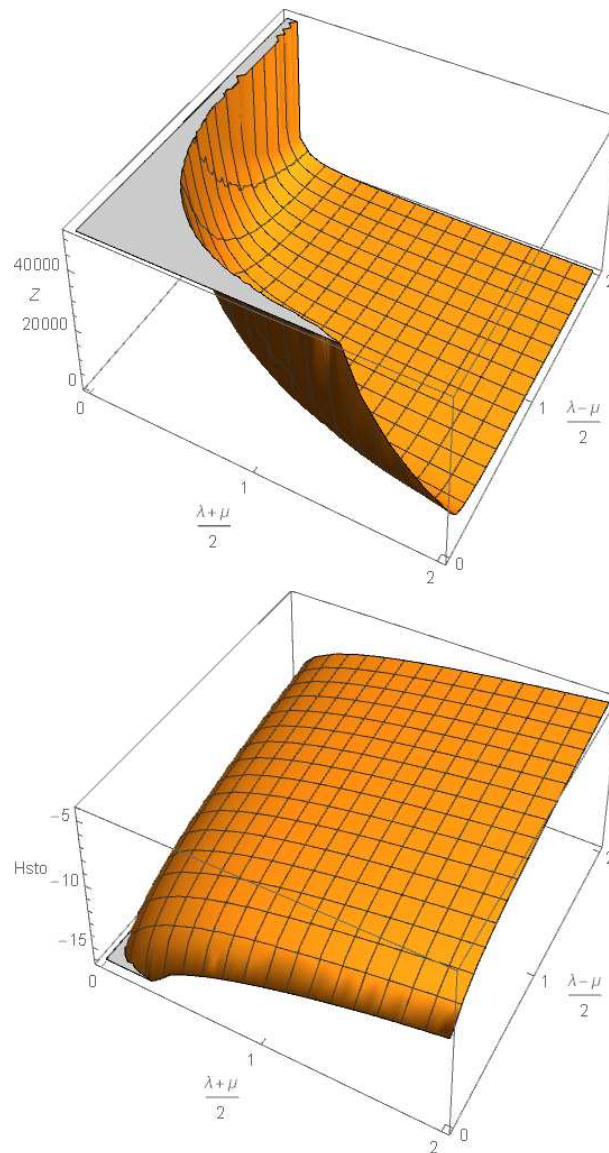


Figure 6. Plots of the numerical evaluations of the Geothermodynamical partition function and stochastic Hamiltonian à la Souriau for the Siegel plane.

8. Conclusions

Machine Learning and so-named Artificial Intelligence algorithms rely on two mathematical pillars: Differential Geometry on one side and Probability Theory on the other. The first is needed to model the spaces where data are to be encoded, the second to classify and elaborate the data by assigning a probability to their actual location in such spaces and to their correlations. Hence, the two mathematical pillars have to be reconciled with one another and solidly entangled. A statistical viewpoint on whatever set of objects always leads to thermodynamics, in a generalized sense, and, already for some decades, an abstract geometrical formulation of thermodynamics has been developed that starts from Shannon's

information entropy and leads to the identification of thermodynamical equilibrium states with Lagrangian submanifolds of symplectic manifolds, also endowing them with a canonical Riemannian structure. In the different data science literature, the same Riemannian structure has been developed under the name of Fisher's information geometry. In such a variegated conceptual landscape and in strong correlation with the newly introduced paradigm of **Cartan neural networks**, where all the hidden layers of neural networks are modeled as non-compact symmetric spaces U/H , that are all Cartan Hadamard manifolds and, consequently, are equipped with a canonical distance function, an important issue for Machine Learning is that of Gaussian-like probability distributions on the encoding spaces. Starting from the hints provided by the work of a group of French authors [30–34], who suggested the use of Gibbs states related to the Lie Group Thermodynamics proposed long ago by Souriau, in the present paper, we took on ourselves the following task:

1. Clarify the relation of Geometrical Thermodynamics à la Ruppeiner [22,24–28] and Lychagin [29] with Souriau's proposals [30,34] and with Information Geometry [43].
2. Distinguish between Souriau non-abelian thermodynamics and the geometrical thermodynamics associated with Integrable Dynamical Systems, in particular the Geodesic Dynamical System associated with the calculation of geodesics on the same U/H symmetric spaces that enter the Machine Learning game as hidden layer models.
3. Investigate the basic principle of Souriau's thermodynamics, that is, the characterization of the locus $\Omega \subset \mathbb{U}$ in the relevant Lie algebra, whose elements are possible generalized temperatures in the sense that for them the partition function integral is convergent.
4. Clarify the role of the coadjoint orbit conception, Souriau's favorite one, that turns out to be equivalent to the more practical and algorithmic conception based on coset manifolds.

We think that we have attained all the goals we aimed at. Indeed, our results can be summarized as follows.

- (1) We have established the identity of Fisher's Information metric, given as the Hessian of a certain matrix with the metric obtained as the Hessian of the stochastic Hamiltonian $\mathcal{H}^{sto}(\lambda)$, derived in Lychagin's approach as the canonical Riemannian metric on Lagrangian submanifolds of a symplectic manifold where, by definition, the symplectic 2-form ω vanishes identically. Such Lagrangian submanifolds are the thermodynamical equilibrium states, and the 1st and 2nd Principle of Thermodynamics are incorporated in their very definition. These notions are fully general and equally apply to any generalized thermodynamics, non-abelian algebra, as it happens in the thermodynamics à la Souriau.
- (2) With respect to the Poissonian structures on the dual $Solv^*$ of solvable Lie algebras $Solv$ utilized also by two of us (P.F. and A.S.) in their 2009 paper [18] on the integrability of the geodesic equations on non-compact symmetric spaces U/H and investigated by Arkhangelsky in [57], where he derived their Hamiltonians in involution, we show here that such a Poissonian structure is only half of the full story, since it is defined only on the momentum subspace of phase space. Introducing also the coordinates, which is what one should always do, there is a complete symplectic manifold with a symplectic 2-form of maximal rank, and what one describes is just the geodesic dynamical system in Hamiltonian formalism. Arkhangelsky Hamiltonians depend only on the momenta, but they are Hamiltonians in involution also with respect to the complete symplectic structure. The geometric thermodynamics associated with such integrable dynamical systems can be constructed, but it is essentially uninteresting for three reasons:

- (a) The dependence of the partition function on volume is factorized, and the equation of state resembles the trivial one of Ideal Gases.
 - (b) The degrees of freedom are few, and a statistical description seems inappropriate.
 - (c) Last but not least, the Gibbs probability distributions have a non-trivial structure only in momentum space, namely along the fibres of the tangent bundle $\mathcal{T}U/H$, not on the very base manifold U/H . All that is of little appeal for Machine Learning applications, where one looks for probability distributions (Gibbs states) on U/H .
- (3) The searched for Gaussian-like probability distributions on U/H are instead provided by the construction of Gibbs states à la Souriau. This requires a symplectic structure on the very manifold U/H and not on its tangent bundle. After demonstrating that a coadjoint U -orbit of some element $\mathbf{b} \in \mathbb{U}$ is always diffeomorphic and algebraically equivalent to a coset manifold U/H , where $H \subset U$ is the stabilizer of the element $\mathbf{b}$ in U , we abandon the coadjoint orbit conception, and we focus on non-compact symmetric spaces U/H . In order to have the symplectic structure and construct Souriau thermodynamics, H must be the stabilizer of some Lie algebra element, and this, as we show, implies that $\mathbb{H}$ has a $\mathfrak{u}(1)$ which endows the symmetric space with a Kähler structure, and the Kähler 2-form $\mathcal{K}$ is the required symplectic 2-form. In this way, we come to the conclusion that the relevant non-compact symmetric spaces are the Kähler ones corresponding only to two infinite series, the Siegel half-planes $\mathbb{S}\mathbb{H}_n$ and the Calabi–Vesentini manifolds $\mathcal{M}^{[2,q]}$ mentioned in Equation (201). The first series is composed of maximally split manifolds of increasing non-compact rank, while the second constitutes a Tits Satake universality class having the Siegel $\mathbb{S}\mathbb{H}_2$ manifold as universal Tits Satake submanifold. In application to Machine Learning, if one wants to take advantage of the Paint Group symmetry (see [1]) and its potentiality in data clustering, the Calabi–Vesentini choice is preferred.
- (4) The central point of Souriau’s generalized thermodynamics, namely the determination of the subspace $\Omega \subset \mathbb{U}$ of allowed temperature vectors was also solved by us in a simple and elegant way. Ω is just the **the adjoint U orbit of a positivity chamber in the Cartan subalgebra $\mathcal{C} \subset \mathfrak{h}$ of the compact subalgebra $\mathbb{H} \subset \mathbb{U}$** . One fixes the sign of the ℓ independent temperatures associated with ℓ generators of $\mathcal{C}$, and by an adjoint transformation of U , generates all the other possible temperature vectors that respect convergence of the partition function integral. This property, apart from solving the convergence issue, is also of practical relevance. Indeed, the true temperatures are just the compact Cartan ones; all the others are an effect of translation of the central point of a Gibbs probability distribution to any other in the manifold by means of an isometry. Hence, we can always utilize the same partition function depending on a very small number of temperatures and change the point in the Gibbs distribution from a given one to its image under any element of the isometry group U .
- (5) For the case of the Poincaré plane, we have explicitly constructed the partition function depending on three temperatures and even studied the 3-dimensional thermodynamical Riemannian metric, showing that it is non-trivial and not that of a space. For the case of the Siegel half-plane $\mathbb{S}\mathbb{H}_2$, we have reduced the partition function to an integral in two variables, whose integrand is a combination of exponentials, roots, and Bessel functions. The integral is convergent, and one can construct a compiled function of which we have shown a plot. The only problem is the computing velocity of the utilized computer.

What to Do Next

As we have shown in Appendix D, the results obtained here for the Tits Satake submanifold $\mathbb{S}\mathbb{H}_2$ are liable to be extended to the entire universality class of Calabi–Vesentini manifolds by careful use of Paint Group invariance. This is certainly the most urgent task in the program. Succeeding in that, we will have Gibbs probability distributions for each of the candidate hidden layers of a Cartan Neural Network based on Calabi–Vesentini manifolds. As Barbaresco et al. have shown [30–34], geometric thermodynamics à la Souriau can be utilized to study time/sequences, in particular those provided by radar data. This is just the tip of the iceberg. On one side, there are many other possible sequential data; on the other side, the use of Gibbs probability distributions on the hidden layers, or more generally on the manifold where various types of data can be mapped, introduces a powerful new weapon for designing algorithm architectures. Furthermore, recalling the results of [1] so far not yet used, one activity direction might be that of restricting the Gibbs probability distributions to infinite discrete subsets of the Calabi–Vesentini manifold, corresponding to orbits of the origin under the action of any of the large class of discrete subgroups of $SO(2, 2 + q)$ that were classified and constructed in the foundational paper [1].

Final Comment

When the present article was ready for posting on arXiv and for submission to a Journal we learned about the beautiful and inspiring paper by Laurent Bonnasse-Gahot and Jean Pierre Nadal [62] focused on *geometry of the internal representation*, which mathematically means Fisher’s information geometry, namely, what we have here shown to be identical with the thermodynamical geometry of Lychagin, Roop, and Ruppeiner, in whose general scheme we could also fit the Gibbs states à la Souriau on the Kähler non-compact symmetric spaces U/H . It follows that an additional challenging direction of further investigation is the possible application of [62] general methods to data mapped to Kähler non-compact symmetric spaces and to the Gibbs states defined over them.

Author Contributions: Conceptualization, P.G.F., A.S.S. and M.T.; Methodology, P.G.F., A.S.S. and M.T.; Formal analysis, P.G.F., A.S.S. and M.T.; Investigation, P.G.F., A.S.S. and M.T.; Writing—review & editing, P.G.F., A.S.S. and M.T. All authors have read and agreed to the published version of the manuscript.

Funding: The work of A. Sorin is supported in part by the Center for Integration in Science of the Israel Ministry of Aliyah and Integration.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: P.G. Fré acknowledges support by the Company Additati & Partners Consulting s.r.l. during the development of the present study.

Conflicts of Interest: Author Pietro Fré was employed by the company Additati & Partners Consulting s.r.l. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Basic Structures of Contact and Symplectic Geometry

In this appendix, we recall the basic structures of symplectic and contact geometry that are the two tightly connected arenas both for the geometric reformulation of thermodynamics and for dynamical systems, i.e., the underlying mathematical basis of all phenomena studied in stochastic processes, data science, and (geometric) deep learning. Indeed, the prototype of a symplectic manifold is the phase space of a dynamical system.

Appendix A.1. Contact Geometry

Contact Geometry is at the same time a new and ancient chapter in Mathematics, since it originates in classical results dating back to Goursat, Darboux, Lie, and other Masters of the 19th century, but has been vigorously developed in the last two to three decades by a relatively small community of mathematicians. Regarding such structures, there is an extensive literature provided by the following bibliographical references and the additional articles they cite [63–72]. To summarize, it can be said that *Contact Geometry* is a mathematical theory that aims at establishing an intrinsic geometric-topological characterization of *non-integrability* and, in a sense, is formulated in an inverted perspective from that usual in cohomological theories and integrable systems theory.

Not surprisingly, Contact Geometry comes into play whenever one intends to study *chaos and disorder* rather than order, as in *integrable systems*. Disorder and lack of information are characteristics of turbulent regimes in *Fluid Dynamics* and are essential attributes of statistical thermodynamic systems; it is therefore quite natural that Contact Geometry has relations to both Fluid Dynamics and Thermodynamics.

Contact Geometry deals exclusively with *Differentiable Manifolds of odd dimension* $\mathcal{M}_{2n+1}$ and, on the other hand, it has a symbiotic relationship with *Simplectic Manifolds* S_{2n+2} and S_{2n} in the two adjacent even dimensions, the upper and the lower one.

The fundamental notions of Contact Geometry are easily stated and require, to be assimilated, nothing more than the most basic and elementary concepts of Differential Geometry. In the concise summary we present in this section, we closely follow the excellent review article [66].

Appendix A.2. Contact Structures

Let us consider an odd-dimensional differentiable manifolds $\mathcal{M}_{2n+1}$ and its tangent bundle:

$$\mathcal{T}\mathcal{M}_{2n+1} \xrightarrow{\pi} \mathcal{M}_{2n+1}; \quad \forall p \in \mathcal{M}_{2n+1} \quad , \quad \pi^{-1}(p) \sim \mathbb{R}^{2n+1} \quad (\text{A1})$$

By definition, the transition function between any pair of local trivializations of the tangent vector bundle is provided by the *inverse Jacobian* of the transition function $\psi_{ij}(x)$ between the two corresponding open charts (U_i, φ_i) and (U_j, φ_j) in any atlas that covers the entire manifold $\mathcal{M}_{2n+1}$ (see, for instance, [9]).

The space of the sections of the tangent bundle $\Gamma[\mathcal{T}\mathcal{M}_{2n+1}, \mathcal{M}_{2n+1}]$ is made up of all the vector fields, whose local description is in terms of first-order differential operators:

$$\begin{aligned} \mathbf{X} &\in \Gamma[\mathcal{T}\mathcal{M}_{2n+1}] \quad \Downarrow \\ \mathbf{X} &= X^\mu(x) \frac{\partial}{\partial x^\mu} \quad : \quad \text{In any open chart } U \text{ with coordinates } x^1 \dots x^{2n+1} \end{aligned} \quad (\text{A2})$$

The cotangent bundle

$$\mathcal{T}^*\mathcal{M}_{2n+1} \xrightarrow{\pi_*} \mathcal{M}_{2n+1}; \quad \forall p \in \mathcal{M}_{2n+1} \quad , \quad \pi_*^{-1}(p) \sim (\mathbb{R}^{2n+1})^* \quad (\text{A3})$$

is the dual of the tangent one, in the sense that its fibre over any point $p \in \mathcal{M}_{2n+1}$ of the basis space is the dual of the fibre vector space over the same point as the tangent fibre, that is, the space of linear functionals over the latter:

$$\forall p \in \mathcal{M}_{2n+1} \quad : \quad \pi_*^{-1}(p) = \text{space of linear functionals on } \pi^{-1}(p) \quad (\text{A4})$$

By construction, the transition function between two local trivializations of the cotangent bundle is provided by the *direct Jacobian* of the two corresponding open charts (U_i, φ_i) and (U_j, φ_j) in any atlas that covers the entire base manifold $\mathcal{M}_{2n+1}$.

The space of sections of the cotangent bundle $\Gamma[\mathcal{T}^*\mathcal{M}_{2n+1}, \mathcal{M}_{2n+1}]$ is made by differential 1-forms whose local description is recalled below:

$$\begin{aligned} \omega &\in \Gamma[\mathcal{T}^*\mathcal{M}_{2n+1}, \mathcal{M}_{2n+1}] \quad \Downarrow \\ \omega &= \omega_\mu(x) dx^\mu \quad : \quad \text{In any open chart } U \text{ with coordinates } x^1 \dots x^{2n+1} \end{aligned} \quad (\text{A5})$$

In terms of these fundamental concepts, which apply to differentiable manifolds of any dimension, whether even or odd, the concept of hyperplane bundles can be introduced. A hyperplane bundle is a reduction of the tangent bundle where the fibres above each point constitute a codimension one subspace of the tangent space above the same point, the transition functions being derived accordingly:

$$\begin{aligned} \mathcal{HY} &\xrightarrow{\mathcal{P}} \mathcal{M} \quad ; \quad \forall p \in \mathcal{M}, \mathcal{P}^{-1}(p) \subset \pi^{-1}(p) \quad \text{where} \quad \mathcal{TM} \xrightarrow{\pi} \mathcal{M} \\ \dim_{\mathbb{R}} \mathcal{M} &= m \quad ; \quad \dim_{\mathbb{R}} \pi^{-1}(p) = m; \quad \dim_{\mathbb{R}} \mathcal{P}^{-1}(p) = m - 1 \end{aligned} \quad (\text{A6})$$

A simple way to construct a hyperplane bundle is by means of a section of the cotangent bundle, namely by means of a 1-form $\omega \in \Gamma[\mathcal{T}^*\mathcal{M}, \mathcal{M}]$. The desired hyperplane sub-bundle $\mathcal{HY}^\omega \subset \mathcal{TM}$ of the tangent bundle is implicitly defined by specifying the space of all of its sections $\Gamma[\mathcal{HY}^\omega, \mathcal{M}]$, namely by specifying all vector fields that are sections of $\mathcal{HY}^\omega$. In a precise mathematical language, let $\mathbf{X} \in \Gamma[\mathcal{TM}, \mathcal{M}]$ be a vector field, then we write:

$$\mathbf{X} \in \Gamma[\mathcal{HY}^\omega, \mathcal{M}] \quad \text{iff} \quad \mathbf{X} \in \ker \omega \quad \text{i.e.,} \quad \omega(\mathbf{X}) \equiv 0 \quad (\text{everywhere}) \quad (\text{A7})$$

Definition A1. Given an odd-dimensional manifold $\mathcal{M}_{2n+1}$, a **contact structure** on $\mathcal{M}_{2n+1}$ is a rank $2n$ sub-bundle $\xi \xrightarrow{\mathcal{P}} \mathcal{M}_{2n+1}$ of tangent bundle $\mathcal{TM}_{2n+1} \xrightarrow{\pi} \mathcal{M}_{2n+1}$ that can be identified with the hyperplane bundle $\mathcal{HY}^\alpha$ where the 1-form α satisfies the following condition:

$$\alpha \wedge \underbrace{d\alpha \wedge \dots \wedge d\alpha}_{n\text{-volte}} \neq 0 \quad (\text{everywhere on } \mathcal{M}_{2n+1}) \quad (\text{A8})$$

The 1-form α is named the **contact form** of the contact structure.

Definition A2. A **contact manifold** is a pair $(\mathcal{M}_{2n+1}, \xi)$ made by an odd-dimensional manifold and a contact structure $\xi \xrightarrow{\mathcal{P}} \mathcal{M}_{2n+1}$.

Some observations are obligatory in connection with the two definitions above. The first is that the same contact structure can be defined by different contact forms $\alpha, \alpha', \dots$. In fact, all multiples of a given contact form by means of a scalar function that does not vanish at any point $\lambda : \mathcal{M}_{2n+1} \rightarrow \mathbb{R}$ define the same contact structure. Secondly, it is perfectly possible that the same odd-dimensional manifold $\mathcal{M}_{2n+1}$ admits more than one contact structure. The classification of such contact structures modulo diffeomorphism connected to the identity map is an interesting and relevant mathematical problem for odd-dimensional manifolds, just as it is interesting and relevant, the classification of complex structures that exist on an even-dimensional manifold. It is therefore obligatory to single out the concept of **contactomorphism**.

Definition A3. Let $(\mathcal{M}, \xi)$ and $(\mathcal{N}, \chi)$ be two contact manifolds and let

$$\varphi : \mathcal{M} \longrightarrow \mathcal{N} \quad (\text{A9})$$

be a diffeomorphism of the former onto the latter manifold (obviously, $\mathcal{M}$ and $\mathcal{N}$ must have the same dimensions in order for φ to possibly exist). Let α be a contact form that defines ξ and let β be

a contact form that defines χ . The considered diffeomorphism φ is named a **contactomorphism** if and only if:

$$\varphi^*(\beta) = \lambda \alpha \quad (\text{A10})$$

where φ^* is the pull-back map and

$$\lambda : \mathcal{M} \longrightarrow \mathbb{R} \quad (\text{A11})$$

is a nowhere vanishing function on $\mathcal{M}$. If a contactomorphism between them does exist, the two considered manifolds are named **contactomorphic**.

In the Definition A3, the two manifolds $\mathcal{M}$ and $\mathcal{N}$ might be the same manifold. In this case, what we are considering is the transformation of a contact structure into another one by means of a contactomorphism.

Definition A4. Given two contact structures ξ and χ on the same manifold $\mathcal{M}_{2n+1}$ they must be identified as the same contact structure if a contactomorphism does exist that maps one into the other.

In conclusion the relevant mathematical problem is that of classifying contact structures on a manifold $\mathcal{M}_{2n+1}$ modulo contactomorphisms.

Appendix A.3. Integrability and Frobenius Theorem

We do not go into the details of Frobenius theorem (see for example esempio [73] e [66]) but we merely summarize the concepts underlying its formulation. We begin by noting that any vector field $\mathbf{X}$ on a differentiable manifold $\mathcal{M}$ of any dimension defines its own integral curves $\mathcal{I}_{\mathbf{X}}$, i.e., those curves that at each of their points admit the local value of the vector field $\mathbf{X}$ as a tangent vector. Since any $p \in \mathcal{M}$ lies on some integral curve $\mathcal{I}_{\mathbf{X}}$, we are guaranteed that a single vector field induces a *foliation* of the manifold $\mathcal{M}$ into one-dimensional submanifolds.

On the other hand, it is a more complicated matter to determine whether or not a rank $r > 1$ sub-bundle $\mathcal{E} \longrightarrow \mathcal{M}$ of the tangent bundle does or does not induce a *foliation* of $\mathcal{M}$. To this effect, utilizing a not completely rigorous, yet intuitive and qualitatively correct way of speaking, by means of the wording *foliation* one means the covering of the manifold with a family of *leaves*, i.e., submanifolds all diffeomorphic to each other, $\mathcal{F}_{\mathbf{v}} \subset \mathcal{M}$ having dimension equal to the rank r of the sub-bundle $\mathcal{E}$, each of which can be regarded as the level hypersurface of r functions $u_i(p)$ ($i = 1, \dots, r$) that originate from the integration of a basis of sections $\mathbf{X}_i$ of the sub-bundle $\mathcal{E} \longrightarrow \mathcal{M}$:

$$\begin{aligned} \mathcal{F}_{\mathbf{v}} &= \{p \in \mathcal{M} \mid u_i(p) = v_i\}; \quad \mathbf{v} \equiv \{v_1, \dots, v_r\}; \quad v_i = \text{real constants} \\ \nabla u_i(p) &= \mathbf{X}_i|_p \end{aligned} \quad (\text{A12})$$

When this situation is realized one says that the sub-bundle $\mathcal{E} \longrightarrow \mathcal{M}$ is integrable. **Frobenius Theorem** establishes the necessary conditions for such an integrability.

Theorem A1. Let $\mathcal{M}$ be a differentiable manifold and let $\mathcal{E} \longrightarrow \mathcal{M}$ be a rank $r > 1$ sub-bundle of the tangent bundle $\mathcal{T}\mathcal{M}$. The necessary and sufficient condition in order for $\mathcal{E}$ to be integrable is the following one:

$$\forall \mathbf{X}, \mathbf{Y} \in \Gamma[\mathcal{E}, \mathcal{M}] \quad : \quad [\mathbf{X}, \mathbf{Y}] \in \Gamma[\mathcal{E}, \mathcal{M}] \quad (\text{A13})$$

In the case where $\mathcal{E} \longrightarrow \mathcal{M}$ is a hyperplane bundle defined by a 1-form ω , Frobenius integrability condition can be formulated as follows

$$\omega \wedge d\omega = 0 \quad (\text{A14})$$

In order to appreciate the equivalence of the condition (A14) with condition (A13) it suffices to recall Cartan's formula for the value of $d\omega$ on two arbitrary vector fields $\mathbf{X}, \mathbf{Y}$:

$$d\omega(\mathbf{X}, \mathbf{Y}) = \frac{1}{2} (\mathbf{X}\omega(\mathbf{Y}) - \mathbf{Y}\omega(\mathbf{X}) - \omega([\mathbf{X}, \mathbf{Y}])) \quad (\text{A15})$$

Let $\mathbf{Z} \notin \Gamma[\mathcal{E}, \mathcal{M}]$ and $\mathbf{X}, \mathbf{Y} \in \Gamma[\mathcal{E}, \mathcal{M}]$; next evaluate the 3-form $\omega \wedge d\omega$ on the triplet $\mathbf{Z}, \mathbf{X}, \mathbf{Y}$ of vector fields. We find:

$$\omega \wedge d\omega(\mathbf{Z}, \mathbf{X}, \mathbf{Y}) \propto \underbrace{\omega(\mathbf{Z})}_{\neq 0} \omega([\mathbf{X}, \mathbf{Y}]) \quad (\text{A16})$$

Hence Equation (A14) implies $\omega([\mathbf{X}, \mathbf{Y}]) = 0$ that on its turn means $[\mathbf{X}, \mathbf{Y}] \in \Gamma[\mathcal{E}, \mathcal{M}]$.

In this way we see that a contact structure defined by a contact 1-form is the exact opposite of an integrable sub-bundle. Indeed, one can show that it corresponds to maximum nonintegrability.

In Section 4.2, considering the integrable dynamical systems constructed on *normed solvable Lie algebras* and solvable groups, we get deeper into the meaning of Frobenius theorem. For integrable systems the foliation of the differentiable manifold is provided by the level set surfaces of a maximal set of Hamiltonian functions in involution whose corresponding vector fields span the integrable sub-bundle of the tangent bundle.

Appendix A.4. Isotropic Submanifolds of a Contact Manifold and Non Integrability

Let us introduce the following:

Definition A5. Let $(\mathcal{M}_{2n+1}, \xi)$ be a contact manifold and $\mathcal{L} \subset \mathcal{M}_{2n+1}$ be one of its submanifolds. Let us consider the tangent bundle of such a submanifold $\mathcal{TL} \xrightarrow{\pi_{\mathcal{L}}} \mathcal{L}$ and the contact structure $\xi \xrightarrow{\pi_{\xi}^*} \mathcal{M}$. The submanifold $\mathcal{L}$ is named **isotropic** if and only if

$$\forall p \in \mathcal{L} \quad : \quad \pi_{\tau}^{-1}(p) \subset \pi_{\xi}^{-1}(p) \quad (\text{A17})$$

Equivalently, if the contact structure ξ is defined by contact 1-form α , the submanifold $\mathcal{L}$ is **isotropic** if any vector field $\mathbf{X}$ that is tangent to $\mathcal{L}$, is contained in $\ker \alpha$:

$$\mathbf{X} \in \Gamma[\mathcal{TL}, \mathcal{L}] \quad \Rightarrow \quad \alpha(\mathbf{X}) = 0 \quad (\text{A18})$$

Let us introduce the further:

Definition A6. Let $(\mathcal{M}_{2n+1}, \xi)$ be a contact manifold and $\widetilde{\mathcal{M}}_{2m+1} \subset \mathcal{M}_{2n+1}$ be an odd-dimensional submanifold with codimension $2(n - m) \geq 0$. Let α be the contact 1-form that defines the contact structure ξ and let ι be the inclusion map:

$$\iota \quad : \quad \widetilde{\mathcal{M}}_{2m+1} \longrightarrow \mathcal{M}_{2n+1} \quad (\text{A19})$$

In this case $(\widetilde{\mathcal{M}}_{2m+1}, \chi)$ is named a **contact sbmanifold** of $(\mathcal{M}_{2n+1}, \xi)$ if the contact structure χ on $\widetilde{\mathcal{M}}_{2m+1}$ is defined by the **contact form** $\iota^* \alpha$, namely if

$$\chi = \ker \iota^* \alpha \quad (\text{A20})$$

A result of the highest relevance both for applications in Fluid Dynamics and for the Geometrization of Thermodynamics is the following theorem due to Arnol'd:

Theorem A2. Let $(\mathcal{M}_{2n+1}, \xi)$ be a contact manifold in $2n + 1$ -dimensions and $\mathcal{L} \subset \mathcal{M}_{2n+1}$ an isotropic submanifold. Then $\dim \mathcal{L} \leq n$.

In order to prove the Theorem A2 the following Lemma is needed

Lemma A1. Let $(\mathcal{M}_{2n+1}, \xi)$ be a contact manifold whose contact structure ξ is defined as $\ker \alpha$, in terms of a contact 1-form α . As a consequence of the condition (A8) included in the definition, it follows that $d\alpha|_{\xi} \neq 0$ and for each point $p \in \mathcal{M}_{2n+1}$ the $2n$ -dimensional fibre $\xi_p \subset T_p \mathcal{M}_{2n+1}$ is a vector space endowed with an antisymmetric 2-form of maximal rank *massimale* (namely without vanishing eigenvalues) which is exactly provided by the restriction to ξ_p of $d\alpha$ i.e., $d\alpha|_{\xi_p}$. Hence the contact structure is a symplectic bundle with respect to the 2-form $d\alpha|_{\xi}$.

The proof of the lemma is almost evident from its own formulation.

Proof. In order to prove the theorem we consider the inclusion map:

$$\iota : \mathcal{L} \longrightarrow \mathcal{M}_{2n+1} \quad (\text{A21})$$

and also the *pull-back* of the contact 1-form on the isotropic submanifold. By definition $\iota^* \alpha = 0$. Hence we have $\iota^* d\alpha = 0$. In any point $p \in \mathcal{L}$, the tangent space $\mathcal{T}_p \mathcal{L}$ is a subspace of the symplectic space ξ_p on which the symplectic 2-form vanishes $d\alpha|_{\xi_p}$. Using elementary linear algebra we conclude that such a subspace has maximal dimension one-half of the dimension of ξ_p . Indeed it suffices to put, by means of change of basis the antisymmetric 2-form in canonical form:

$$\left(\begin{array}{c|c} \mathbf{0}_{n \times n} & \mathbf{1}_{n \times n} \\ \hline -\mathbf{1}_{n \times n} & \mathbf{0}_{n \times n} \end{array} \right)$$

and the statement becomes evident. This proves the theorem. $\square$

What are the consequences of this theorem? It states that if we have a contact structure ξ , induced by a contact 1-form α , then we can exclude a foliation of the contact manifold into hypersurfaces $\Sigma_h \subset \mathcal{M}_{2n+1}$ codimension one:

$$\mathcal{M}_{2n+1} \simeq \Sigma_h \times \mathbb{R}_h \quad (\text{A22})$$

such that for each $h \in \mathbb{R}$ the tangent bundle of Σ_h be contained in the contact structure. Indeed if this were true so any leave Σ_h of the foliation would be an isotropic submanifold of dimension $2 \times n$ which is exactly what is ruled out by the theorem.

Definition A7. An isotropic submanifold $\mathcal{L} \subset \mathcal{M}_{2n+1}$ of a $(2n + 1)$ -dimensional contact manifold that has the maximal allowed dimension n is named a **Legendrian submanifold**.

AN EXAMPLE RELEVANT FOR FLUID DYNAMICS

In Fluid Dynamics the relevant manifold $\mathcal{M}_3$, is locally isomorphic to our three dimensional Euclidean space $\mathbb{E}^3 \cong \mathbb{R}^3$ since fluids of interest flow in a portion $\mathcal{M}_3 \subset \mathbb{R}^3$ of such a space that can be compact (the fluid is confined in some container) or partially non compact (river, lake, ocean, atmosphere). In any case we have $n = 1$. Hence if the flow, which is a vector field $\mathbf{U} \in \Gamma[\mathcal{T}\mathcal{M}_3, \mathcal{M}_3]$ induces a contact structure on $\mathcal{M}_3$ (in the next subsection we will see that this happens when $\mathbf{U}$ is proportional to Reeb vector of a contact form α), then the unique Legendrian submanifolds are one-dimensional so that the ambient space does not admit a foliation into 2-dimensional submanifolds that contain the flows or

that are transverse to them. This is the geometrical foundation of turbulence: the flows are chaotic.

AN EXAMPLE RELEVANT FOR THERMODYNAMICS

As we are going to see later on the relevant contact manifold for thermodynamics is not the Euclidean physical space rather the space of thermodynamical variables (both extensive and intensive) that are in number of: $2n + 1 \equiv 2m + 3$ ($n = m + 1$). The simplest minimal case corresponds to $m = 1$ and the 5 thermodynamic variables are $\{U, S, V, T, P\}$, namely *internal energy, entropy, volume, temperature and pressure*. As we will explain the thermodynamical space $\mathcal{M}_{2m+3}$ is always endowed with a canonical contact 1-form that summarizes the principles of classical thermodynamics and available equilibrium thermodynamical states correspond to all possible Legendrian submanifolds $\mathcal{L}_E \subset \mathcal{M}_{2m+3}$ that have, as a consequence of Theorem A2 and of Definition A7 the following dimension:

$$\dim \mathcal{L}_E = m + 1 \quad (\text{A23})$$

In the simplest case $m + 1 = 2$ so that the Legendrian subvarieties are portions of an $\mathbb{R}^2$ -plane. It is in such planes that one studies the phase diagrams of chemicals. In the case of multiphase, multicomponent mixtures we have $m > 1$ and phase diagrams develops in spaces of higher dimensions.

BIG DATA SPACES

It is to be investigated whether certain Big Data spaces might be endowed with a hidden contact structure responsible for chaotic motions and phase transitions.

Appendix A.5. The Reeb Vector

Let us now come to the definition of the Reeb vector which shifts the definition of a contact structure from the cotangent to the tangent bundle of the considered $\mathcal{M}_{2n+1}$ manifold.

Definition A8. Associated with a contact form α we always have the so named **Reeb vector field** $\mathbf{R}_\alpha$, defined by two conditions:

$$\begin{aligned} \alpha(\mathbf{R}_\alpha) &= \lambda(\mathbf{x}) = \text{nowhere vanishing function on } \mathcal{M}_{2n+1} \\ \forall \mathbf{X} \in \Gamma[\mathcal{T}\mathcal{M}_{2n+1}, \mathcal{M}_{2n+1}] &: \quad d\alpha(\mathbf{R}_\alpha, \mathbf{X}) = 0 \end{aligned} \quad (\text{A24})$$

If the contact manifold $\mathcal{M}_{2n+1}$ is endowed with a Riemannian metric g (as it is the case of the Euclidean space $\mathbb{R}^{2n+1}$), then the contact form α and its Reeb vector field $\mathbf{R}_\alpha$ are related to each other by the operation of raising and lowering of world indices. Suppose we start from the Reeb vector field:

$$\mathbf{R} = R^\mu \frac{\partial}{\partial x^\mu} \quad (\text{A25})$$

The corresponding α is retrieved by setting:

$$\alpha = \Omega^{[\mathbf{R}]} \equiv g_{\mu\nu} R^\mu dx^\nu \quad (\text{A26})$$

and the condition (A8) that such a 1-form should be a contact form becomes the following equation on the Reeb field components:

$$e^{\lambda\mu_1\nu_1\mu_2\nu_2\ldots\mu_n\nu_n} R_\lambda \partial_{\mu_1} R_{\nu_1} \partial_{\mu_2} R_{\nu_2} \ldots \partial_{\mu_n} R_{\nu_n} \neq 0 \quad \text{nowhere vanishes} \quad (\text{A27})$$

In the opposite direction, if one start from the contact form α , the components of the Reeb vector field are obtained by setting:

$$R_\alpha^\mu = g^{\mu\nu} \alpha_\nu \quad (\text{A28})$$

One should note that the nowhere vanishing function λ mentioned in the Definition A8 is simply the squared norm of the Reeb vector field or of the contact form that make sense only in a Riemannian space and there they coincide:

$$\lambda = \| \mathbf{R} \|^2 = \| \Omega^{[\mathbf{R}]} \|^2 \equiv g_{\mu\nu} R^\mu R^\nu \quad (\text{A29})$$

CONTACT STRUCTURES FOR $n = 1$ AND HYDRODYNAMICAL FLOWS

Let us now consider the case relevant to the hydrodynamics of three-dimensional contact manifolds $(\mathcal{M}_3, \xi_\alpha)$, where, by the notation ξ_α , we refer to the contact form α that defines the contact structure ξ . The consequence of Theorem A2, as we have already pointed out, is that in these contact varieties, the Legendrian subvarieties are all one-dimensional, that is, they are *curves* or, as it is usual to refer to them in this context *knots*.

Therefore in three dimensions, there are two types of knots the **Legendrian knots**, whose tangent vector belongs to $\ker \alpha$ and the **transverse knots** whose tangent vector is parallel to the Reeb field vector at every point in their trajectory.

Furthermore in $D = 3$ condition (A27) becomes:

$$\epsilon^{\lambda\mu\nu} R_\lambda \partial_\mu R_\nu \neq 0 \quad (\text{A30})$$

THE STANDARD CONTACT STRUCTURE ON $\mathbb{R}^3$

Three-dimensional flat space, whose coordinates we denote x, y, z is equipped with a standard contact structure that admits the following contact 1-form

$$\alpha_s = dz + x dy \quad (\text{A31})$$

A picture of the local planes defining the contact structure (A31) is shown in Figure A1.

In hydrodynamics, a vector field $\mathbf{U}$ that we identify with the velocity field of a fluid is potentially interesting for the study of chaotic regimes if it is the Reeb field of a contact 1-form α . If our work arena is a Riemannian variety endowed with a metric g cioè $(\mathcal{M}_3, g)$ we can always invert the procedure and define the contact form α by writing:

$$\alpha = \Omega^{[\mathbf{U}]} \equiv g_{ij} U^i(x, t) dx^j \quad (\text{A32})$$

where $U^i(x, t)$ are the three components of the velocity field at time t and the Latin indices $i, j = 1, 2, 3$ identify the three coordinates $x^i = \{x, y, z\}$. In this way the first of the two conditions (A24) is automatically satisfied: $i_{\mathbf{U}} \Omega^{[\mathbf{U}]} = \| \mathbf{U} \|^2 > 0$. It remains to be seen if $\Omega^{[\mathbf{U}]}$ is a true contact form, namely if $\Omega^{[\mathbf{U}]} \wedge d\Omega^{[\mathbf{U}]} \neq 0$. It is at this level that one finds the relation with Beltrami equations which makes sense only in dimension $D = 3$ and hence for $n = 1$. In the thermodynamic case $n \geq 2$: hence we stop the discussion here and we refer the interested reader to [21].

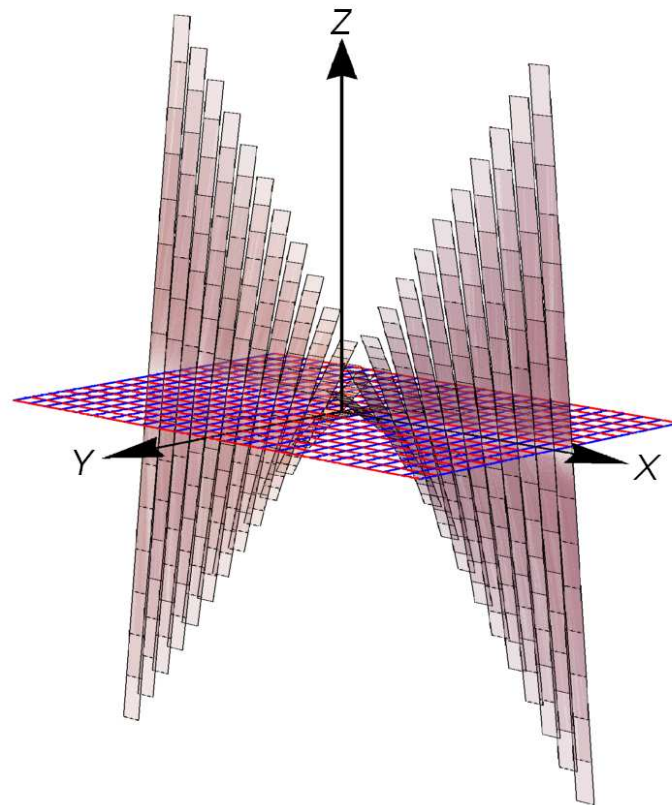


Figure A1. Schematic vision of the standard contact structure in $\mathbb{R}^3$.

Appendix A.6. Darboux Theorem and the Case of Thermodynamics

A classical theorem due to Darboux, the proof of which we omit by referring the reader to [66] where it is presented, highlights the important fact that the standard contact structure of $\mathbb{R}^3$ mostrata nella Equation (A31) and illustrated graphically in Figure A1 is not an arbitrary choice, bensì corresponds to the canonical local form of any contact structure on any contact manifold.

Theorem A3. Let $(\mathcal{M}_{2n+1}, \xi)$ be $(2n + 1)$ -dimensional contact manifold and let α be a contact 1-forma that defines $\xi = \ker \alpha$. Let $p \in \mathcal{M}_{2n+1}$ be any point of the manifold and $U \subset \mathcal{M}_{2n+1}$ an open neighborhood of p . Then we can always find a local homomorphism $\varphi : U \rightarrow \mathbb{R}^{2n+1}$ such that, naming $\{x_0, x_i, y_i\}, (i = 1, \dots, n)$ the coordinates of $\varphi(U) \subset \mathbb{R}^{2n+1}$ we get:

$$\alpha|_U = dx_0 + \sum_{i=1}^n y_i dx_i \quad (\text{A33})$$

In the case $n = 1$ Equation (A33) reproduces Equation (A31).

Darboux's canonical form is the most inspiring one to suggest a connection between contact geometry and thermodynamics. As we shall see in more detail later, in classical thermodynamics it is natural to distinguish two types of variables, the extensive ones such as U, S, V, N (internal energy, entropy, volume, number of particles or of moles) and the intensive ones such as T, P, μ (temperature, pressure, potenziale chemical potential). We can identify the coordinates $x_i = \{S, V, N, \dots\}$ with the extensive variables and the coordinates $y_i = \{T, P, \mu \dots\}$ with the intensive ones, further identifying the privileged extensive variable $x_0 = U$ with internal energy. For a reason that will become clear later, after the discussion of probability measures, it is convenient to generally rename the intensive

variables as λ_i since in the general view of information/probability theory they play the role of Lagrange multipliers. Thus the contact differential 1-form

$$\alpha_{termo} \equiv dx_0 + \sum_{\mu=1}^3 \lambda^\mu dx_\mu \quad (A34)$$

can be used to formulate the first and second Principles of Thermodynamics stating that α_{termo} should vanish. Such vanishing must be understood in a correct mathematical way: it vanishes on Legendrian submanifolds $\mathcal{L}_E \subset \mathcal{M}_{2m+3}$ which represent equilibrium thermodynamical states whose embedding functions in the ambient space $\mathcal{M}_{2m+3}$ are the physical laws defining such states.

Appendix A.7. Symplectic and Poisson Manifolds

In the further study of the immersion of Legendrian submanifolds in the thermodynamic contact manifold $\mathcal{M}_{2m+3}$ an important role is played by the symbiotic relationship of all contact manifolds $\mathcal{M}_{2n+1}$ with the adjacent symplectic manifolds that are also Poissonian.

Hence let us begin with the following:

Definition A9. A symplectic manifold is a pair $(S\mathcal{M}_{2n+2}, \omega)$ of a smooth manifold $S\mathcal{M}_{2n+2}$ in dimension $2n+2$ and one differential 2-form ω that is closed, non degenerate (namely admitting no vanishing eigenvalue) and of maximal rank:

$$d\omega = 0; \quad \omega \wedge \omega \wedge \cdots \wedge \omega \neq 0 \quad \text{everywhere on } S\mathcal{M}_{2n+2} \quad (A35)$$

On a symplectic manifold we have a naturally defined antisymmetric 2-form on the space of sections of the tangent bundle, i.e., on vector fields:

$$\begin{aligned} \omega &: \Gamma[\mathcal{T}S\mathcal{M}_{2n+2}, S\mathcal{M}_{2n+2}] \times \Gamma[\mathcal{T}S\mathcal{M}_{2n+2}, S\mathcal{M}_{2n+2}] \longrightarrow C^{(\infty)}(S\mathcal{M}_{2n+2}) \\ \forall X, Y &\in \Gamma[\mathcal{T}S\mathcal{M}_{2n+2}, S\mathcal{M}_{2n+2}] \quad , \quad \omega(X, Y) \in C^{(\infty)}(S\mathcal{M}_{2n+2}) \end{aligned} \quad (A36)$$

Poisson manifolds are instead defined as it follows:

Definition A10. A Poisson manifold $(\mathcal{P}\mathcal{M}_\ell, \{, \})$ is a pair of a smooth manifold $\mathcal{P}\mathcal{M}_\ell$ of dimension ℓ and a Poisson bracket $\{, \}$ which is a binary operation on the space of smooth functions defined over the manifold:

$$\{, \} : C^{(\infty)}(\mathcal{P}\mathcal{M}_\ell) \times C^{(\infty)}(\mathcal{P}\mathcal{M}_\ell) \longrightarrow C^{(\infty)}(\mathcal{P}\mathcal{M}_\ell) \quad (A37)$$

endowed with the following three properties:

- (1) *Antisymmetry* $\{f, g\} = -\{g, f\}, \quad \forall f, g \in C^{(\infty)}(\mathcal{P}\mathcal{M}_\ell)$
- (2) *Jacobi Identity* $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0, \quad \forall f, g, h \in C^{(\infty)}(\mathcal{P}\mathcal{M}_\ell)$
- (3) *Leibniz rule* $\{f, g \cdot h\} = \{f, g\}h + g\{f, h\}, \quad \forall f, g, h \in C^{(\infty)}(\mathcal{P}\mathcal{M}_\ell)$

The first two properties mentioned in Definition A10 guarantee that the space of functions on a Poisson variety becomes a Lie algebra when it is equipped with a Poisson bracket. On the other hand, the third property implies that every function $f \in C^{(\infty)}$ is associated by the Poisson bracket with a derivation of the commutative algebra of

functions on the manifold, which is, by definition, a vector field $\mathbf{X}_f$; this latter is named **the Hamiltonian vector field** of the function f .

Locally, in each open chart $\{x_1, \dots, x_j\}$, the Poisson bracket takes the following form::

$$\{f, g\} = \pi^{ij}(x) \frac{\partial f}{\partial x^i} \frac{\partial g}{\partial x^j}; \quad \pi^{ij}(x) = -\pi^{ji}(x) \quad (\text{A38})$$

where the antisymmetric contravariant vector $\pi^{ij}(x)$ is usually called a **bivector**. Thus the Hamiltonian vector field $\mathbf{X}_f$ is easily identified as:

$$\mathbf{X}_f = \pi^{ij} \partial_i f \partial_j \quad (\text{A39})$$

Suppose now that the dimension of the Poisson variety is $\ell = 2n + 2$ and that the bivector $\pi^{ij}(x)$ is an everywhere invertible matrix. Posing $\omega = \pi_{k\ell}^{-1} dx^k \wedge dx^\ell$ we obtain a symplectic 2-form of maximal rank that is closed as a consequence of the Jacobi identities satisfied by the bivector. In this way we see that such a Poisson manifold is a symplectic manifold and in particular we can write:

$$\{f, g\} = \omega(\mathbf{X}_f, \mathbf{X}_g) \quad (\text{A40})$$

Definition A11. Let $(\mathcal{SM}_{2n+2}, \omega)$ be a symplectic manifold. A **Liouville vector field** X is a vector field for which the following condition holds:

$$\mathcal{L}_X \omega = \omega \quad (\text{A41})$$

where $\mathcal{L}_X \omega \equiv i_X d\omega + d(i_X \omega)$ denotes the Lie derivative of the 2-form ω along the vector field X .

Note that, being ω closed, if X is a Liouville vector field, we have: $d(i_X \omega) = \mathcal{L}_X \omega = \omega$.

Appendix A.8. The Relation Between Contact Manifolds and Symplectic Manifolds

Let us consider a symplectic manifold $(\mathcal{SM}_{2n+2}, \omega)$ and let us assume that it admits at least one Liouville vector field $\mathbf{L}$. Furthermore let $\Sigma_{\mathbf{L}} \subset \mathcal{SM}_{2n+2}$ be the hypersurface which is transverse to the Liouville field $\mathbf{L}$. Then we realize that $\Sigma_{\mathbf{L}}$ is a contact manifold with contact 1-form $\alpha = i_{\mathbf{L}} \omega$. Since $\Sigma_{\mathbf{L}}$ is transverse to $\mathbf{L}$, the 1-form α vanishes along $\mathbf{L}$ and, on the contrary, it never vanishes on the tangent bundle $T\Sigma_{\mathbf{L}}$. In order to verify that α is a *bona fide* contact form we just have to perform the following calculation:

$$\begin{aligned} \alpha \wedge \underbrace{d\alpha \wedge \dots \wedge d\alpha}_{n\text{-times}} &= i_{\mathbf{L}} \omega \wedge \underbrace{di_{\mathbf{L}} \omega \wedge \dots \wedge di_{\mathbf{L}} \omega}_{n\text{-times}} = i_{\mathbf{L}} \omega \wedge \underbrace{\omega \wedge \dots \wedge \omega}_{n\text{-times}} = \frac{1}{n+1} i_{\mathbf{L}} \left(\underbrace{\omega \wedge \dots \wedge \omega}_{(n+1)\text{-times}} \right) \\ &= \text{Vol}_{\Sigma_{\mathbf{L}}} \end{aligned} \quad (\text{A42})$$

The last equality holds because ω^{n+1} is the volume form of the ambient symplectic manifold and the hypersurface $\Sigma_{\mathbf{L}}$ is, by hypothesis, transverse to the Liouville vector field.

Reversely, given a contact manifold $(\mathcal{M}_{2n+1}, \xi)$ with contact 1-form α and Reeb vector field $\mathbf{R}$, any hypersurface $\Sigma \subset \mathcal{M}_{2n+1}$ that is transverse to the Reeb vector, automatically acquires the structure of a symplectic manifold with symplectic 2-form $\tilde{\omega} = d\alpha|_{\Sigma}$.

Therefore, we can have odd-dimensional contact manifolds lying in between two adjacent even-dimensional symplectic manifolds as it is shown in the following diagram:

$$\begin{array}{ccccc}
 (\mathcal{SM}_{2n}, \tilde{\omega} = d\alpha) & \xhookrightarrow{\quad} & (\mathcal{M}_{2n+1}, \alpha = i_{\mathbf{L}}\omega) & \xhookrightarrow{\quad} & (\mathcal{SM}_{2n+2}, \omega) \\
 \Downarrow & & \Downarrow & & \Downarrow \\
 \text{symplectic} & & \text{contact} & & \text{symplectic} \\
 \text{transverse to Reeb field} & & \text{transverse to Liouville field} & &
 \end{array} \quad (\text{A43})$$

The pattern described in Equation (A43) is reminiscent of what happens with sasakian manifolds that lie in between two Kähler manifolds of adjacent even dimensions. This is not surprising since Kähler manifolds are particular instances of symplectic manifolds where the symplectic form is simply the Kähler 2-form.

We stop here, for the moment with the exposition of fundamental geometric concepts, since we have to turn to the other leg of our constructions, namely, Probability Theory. We shall come back to geometry in later chapter about diffusion theory that establishes a solid link between Markov random processes and Riemannian Geometry.

Appendix B. Fundaments of Probability Theory

In this appendix we recall the fundamental concepts, which will be indispensable to us, with regard to measurement and probability theory.

Appendix B.1. σ -Algebras and Probability Measures

Definition A12. Given a set Ω one defines σ -algebra on Ω a family $\mathcal{A}$ of subsets $A_i \subset \Omega$ such that:

1. $\emptyset \in \mathcal{A}$ and $\Omega \in \mathcal{A}$.
2. If $A \in \mathcal{A}$ then its complement $A^c \equiv \Omega - A$ also belongs to the same family $A^c \in \mathcal{A}$.
3. If the elements of a denumerable family of sets $\{A_i\}_{i \in \mathbb{N}}$ belong to $\mathcal{A}$ then also their union belongs to it:

$$A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$$

Definition A13. The pair $(\Omega, \mathcal{A})$ where Ω is a set and $\mathcal{A}$ is a σ -algebra on Ω is named a measurable space.

In Probability Theory the elements of the σ -algebra, $X \in \mathcal{A}$ are named **events** while the points $p \in \Omega$ are named **experiments**.

Definition A14. One names **Part Algebra** of a set Ω and denotes it with the symbol $\mathfrak{P}(\Omega)$ the set of all subsets equipped with the boolean algebraic operations of union, intersection and complement.

Let us now consider two sets Ω_1 and Ω_2 ; any map:

$$\phi : \Omega_1 \rightarrow \Omega_2 \quad (\text{A44})$$

induces a pullback map $\phi_{\star}^{-1}$ on the corresponding Part Algebras:

$$\phi_{\star}^{-1} : \mathfrak{P}(\Omega_2) \rightarrow \mathfrak{P}(\Omega_1) \quad (\text{A45})$$

One can verify that the map $\phi_\star^{-1}$ satisfies the following properties

$$\begin{aligned}\phi_\star^{-1}(X \cup Y) &= \phi_\star^{-1}(X) \cup \phi_\star^{-1}(Y) \\ \phi_\star^{-1}(X \cap Y) &= \phi_\star^{-1}(X) \cap \phi_\star^{-1}(Y) \\ \phi_\star^{-1}(X^c) &= \phi_\star^{-1}(X)^c\end{aligned}\tag{A46}$$

Hence $\phi_\star^{-1}$ is a morphism of boolean algebras.

Once this general concepts have been established one can introduce the following notion of *probability measure* by means of the following:

Definition A15. Let $(\Omega, \mathcal{A})$ be a measurable space. A probability measure on $(\Omega, \mathcal{A})$ is a map

$$\mathbf{p} : \mathcal{A} \rightarrow [0, 1] \subset \mathbb{R}\tag{A47}$$

that satisfies the following properties:

- $\mathbf{p}(\emptyset) = 0$ and $\mathbf{p}(\Omega) = 1$
- $\mathbf{p}(\bigcup_i X_i) = \sum_i \mathbf{p}(X_i)$ for all denumerable unions of disjoint parts, i.e., such that $X_i \cap X_j = \emptyset$ se $i \neq j$.

When we have a triplet $(\Omega, \mathcal{A}, \mathbf{p})$ we say that we have a stochastic space and the value $\mathbf{p}(X)$ is the probability of the event X .

Appendix B.2. Stochastic Functions, Stochastic Vectors and Distributions

Let us begin with:

Definition A16. If $\mathcal{T}$ is a separable topological space, one names **Borel Algebra** of $\mathcal{T}$, denoted $\mathcal{B}(\mathcal{T})$ the σ -algebra made by all denumerable unions, intersections and complements of open subsets $U \subset \mathcal{T}$.

In particular, on all varieties $\mathbb{R}^n$ we have the ball-topology which, for the real line $\mathbb{R}$, reduces to the topology of open intervals $]x, y[\subset \mathbb{R}$, and the correspondent Borel algebra is clearly defined. Hence, using as σ -algebra the natural Borel algebra $\mathcal{B}(\mathbb{R})$, we have that the pair $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, makes a measurable space

Let us then consider a stochastic space $(\Omega, \mathcal{A}, \mathbf{p})$ and a map:

$$\psi : \Omega \rightarrow \mathbb{R}\tag{A48}$$

which to any point $p \in \Omega$ of the set Ω associates a real number (its coordinate). Because of what we discussed above, the pullback map

$$\psi_\star^{-1} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{A}\tag{A49}$$

is a morphism of boolean algebras that explicitly associates an element $X \in \mathcal{A}$ to every element of the Borel algebra of the real line. Thus, composing the maps we define:

$$\mathbf{p}_\psi \equiv \mathbf{p} \circ \psi_\star^{-1}\tag{A50}$$

which is a map from the Borel algebra of the real line to the interval $[0, 1]$:

$$\mathbf{p}_\psi : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]\tag{A51}$$

This is what we name a **stochastic function**. In practice to every open interval $]x, y[$ the stochastic function associates a number between 0 and 1 which is the probability that while doing a measuring experiment the measured value happens to be in the considered interval. In this way one can consider stochastic functions that are discontinuous, step-wise and the like, yet they are Lebesgue integrable thanks to the measurability of the support space.

PROBABILITY DENSITY

An interesting case is when the stochastic function can be described in terms of a probability density given by an integrable function $\rho_\psi(q)$ on the real line such that:

$$\mathbf{p}_\psi([a, b]) = \int_a^b \rho_\psi(q) dq \quad (\text{A52})$$

For the probability density to be well defined, it is necessary that the probability density $\rho_\psi(q)$ be properly normalized:

$$\int_{-\infty}^{+\infty} \rho_\psi(q) dq = 1 \quad (\text{A53})$$

Under these conditions, one can define the average value of any function $f(q)$ of the stochastic variable $q \in \mathbb{R}$ writing

$$\langle f \rangle \equiv \int_{-\infty}^{+\infty} f(q) \rho_\psi(q) dq \quad (\text{A54})$$

STOCHASTIC VECTOR

In a similar way we can define stochastic vectors.

Consider a finite dimensional vector space $\mathbb{V}$:

$$\dim_{\mathbb{R}} \mathbb{V} = r < \infty \quad (\text{A55})$$

and a basis $\mathbf{e}_i$ ($i = 1 \dots, r$) of vectors such that

$$\forall \mathbf{X} \in \mathbb{V} : \mathbf{X} = \sum_{i=1}^r X^i(\chi) \mathbf{e}_i \quad (\text{A56})$$

where the components of the vector are thought of as functions of $\chi \in \Omega$, the space of events over which we defined the probability measure. By reasoning entirely analogous to that above, each component $X^i(\chi)$ can be thought of as a probability density $X_\psi^i(\mathbf{q})$ on a space $\mathbb{R}^n$ where n is the number of coordinates necessary to identify a point $\chi \in \Omega$, namely the dimension of the set Ω , if this latter can be thought of as a differentiable manifold. As in the previous case what we are constructing is, for each component w^i of the stochastic vector, a map:

$$\psi^i : \Omega \rightarrow \mathbb{R}^n \quad (\text{A57})$$

which, by pullback, induces a map:

$$\psi_\star^{-1|i} : \mathcal{B}(\mathbb{R}^n) \rightarrow \mathcal{A} \quad (\text{A58})$$

By composition of maps we obtain

$$\mathbf{p}_{\psi^i} \equiv \mathbf{p} \circ \psi_\star^{-1|i} \quad (\text{A59})$$

which is a map from the Borel algebra of $\mathbb{R}^n$ to the interval $[0, 1]$:

$$\mathbf{p}_{\psi^i} : \mathcal{B}(\mathbb{R}^n) \rightarrow [0, 1] \quad (\text{A60})$$

In this way we have defined a stochastic vector, namely a map:

$$\Psi : \Omega \rightarrow \mathbb{V} \quad (\text{A61})$$

Also for stochastic vectors the most favorable and smooth situation occurs when each of the vector components is substituted by an integrable probability density:

$$\mathbf{X}(\mathbf{q}) = \sum_{i=1}^r \rho_{\Psi}^i(\mathbf{q}) \mathbf{e}_i; \quad \int \int \cdots \int_{\mathbb{R}^n} \rho_{\Psi}^i(\mathbf{q}) \underbrace{d^n \mathbf{q}}_{\equiv d\mu(\mathbf{q})} = 1 \quad (\text{A62})$$

where with $d\mu(\mathbf{q})$ we have denoted the integration measure on the space Ω which might be more elaborate and contain a factor $\sqrt{\det g}$ when Ω is a Riemannian manifold endowed with a metric.

Appendix C. A Summary of Classical Thermodynamics and Statistical Mechanics

In this appendix in order to establish the notations and systematically organize the geometrical treatments of both classical thermodynamics and statistical mechanics, we introduce in a very synthetic way the basic concepts of both disciplines [74,75] making when necessary reference to the general conceptual frameworks discussed in the previous appendix and in the main text. Indeed the main motivation of the present summary is to show how Shannon Information entropy and classical thermodynamical entropy do indeed coincide emphasizing that a thermodynamical, geometrical view is closely inherent to any Big Data system.

Appendix C.1. Thermodynamical Potentials and State Functions

In macroscopic thermodynamics one utilizes the following **extensive quantities**:

1. U = *internal energy* of a thermodynamical system
2. S = *entropy*
3. V = *volume* occupied by the system under consideration, for instance a mixture of gases or a certain quantity of a liquid or of a solid.
4. N = *the total number of particles* composing the system, for instance the number of molecules of a gas that can be measured in various units, among which the most frequently utilized is the *number of moles* of the chemical compound under investigation. Alternatively when the system is a mixture of more than one component one utilizes:
5. N_i = *the total number of particles of the i -th component* of the mixture, typically measured in *number or fractions of moles*.

Extensive quantities means that partitioning the system into subsystems the considered quantity is subdivided into percentual fractions. In other words, for instance the entropy of a system composed of two subsystems A and B is the entropy of A plus the entropy of B :

$$S_{A \cup B} = S_A + S_B \quad (\text{A63})$$

Similarly can be said of the internal energy U , of the volume V and of the numbers of particles N_i or fractions of moles.

The **intensive quantities** of classical thermodynamics do not have the same property and they are instead characterized by the fact that in the equilibrium states they **assume the same value in every part of the system**, large or small. They are

1. T = *temperature* that determines the average energy per particle.
2. P = *pressure* which, as in mechanics, is the force per unit area.

3. μ = *chemical potential* which is the intensive variable canonically conjugate to the number of particles N .
4. μ_i = *chemical potentials of the various components in the various phases* as happens in multicomponent and multiphase mixtures.

Appendix C.2. Thermodynamical Constants

Before we begin our summary of classical thermodynamics and statistical mechanics, it is appropriate to recall the definition and numerical value of universal constants, both fundamental and empirical.

Definition A17. One *mole* of a chemical substance $\mathcal{X}$ is the quantity of atoms or of molecules of that substance $\mathcal{X}$ necessary to form a mass M numerically equal in grams to the weight $w_{\mathcal{X}}$ of an atom or a molecule of that substance $\mathcal{X}$ expressed in atomic mass units.

Thanks to such a definition one mole of any chemical substance $\mathcal{X}$ always contains the same number of atoms or of molecules that is the Avogadro number:

$$N_A = 6.002214076 \times 10^{23} \quad (\text{A64})$$

In view of the Definition (A20) Avogadro number can be seen as the conversion factor from the standard mass unit, namely the gram and the atomic mass unit $\mathbf{u}$

$$1\text{g} = N_A \mathbf{u} \quad (\text{A65})$$

The Equation of State (A115) that shortly after we rigorously derive from the partition function for a classical system of free identical particles, coincides with the Equation of State of diluted gases, empirically known since long time in the following form (It was experimentally determined by Émile Clapeyron in 1834):

$$P V = n R T \quad (\text{A66})$$

where P is the pressure, n denotes the **number of moles** of the gas that are present in the considered volume V , and R is a universal physical constant with the following value:

$$R = 8.31446261815324 \frac{\text{J}}{\text{mol} \times \text{K}} \quad (\text{A67})$$

Definition A18. Boltzmann constant, that appears in all formulae of statistical mechanics, is defined as the ratio of the ideal gas constant R and Avogadro number (A119):

$$k_B \equiv \frac{R}{N_A} = 1.380649 \times 10^{-16} \text{erg K}^{-1} \quad (\text{A68})$$

Thanks to Definition A19, the empirical form of the equation of state (A66) and that derived from the Statistical Mechanics of free particles do coincide since by substituting (A118) into Equation (A115), we obtain the ratio N/N_A between the number of particles and Avogadro number, which is by definition the number of moles of the considered chemical substance:

$$n = \frac{N}{N_A} = \# \text{ di moli} \quad (\text{A69})$$

The First and Second Principles of Thermodynamics

Classical thermodynamics is axiomatically formulated through two principles that are expressed in differential form and concern the infinitesimal changes in the quantities

introduced in the previous paragraph when, by means of heat supply or subtraction, $\pm dQ$ an infinitesimal transformation of the thermodynamic system is carried out from its equilibrium state.

THE FIRST PRINCIPLE

The first principle asserts that if we call dW the mechanical work absorbed or given up by the system, in an infinitesimal transformation, the following relationship holds

$$dQ = dU - dW \quad (\text{A70})$$

Typically, mechanical work produces a change in the volume V of the system, and since pressure is force per unit area, we have $dW = PdV$ so that the canonical formulation of the first principle is the following:

$$dQ = dU - P dV \quad (\text{A71})$$

THE SECOND PRINCIPLE

The second principle of thermodynamics can be formulated in several equivalent ways. The most concise is as follows:

$$\oint_C \frac{dQ}{T} = 0 \quad (\text{A72})$$

which translates to saying that if we integrate the differential form $\frac{dQ}{T}$ along a closed path in the space of state variables, that is, by performing through exchanges of work and heat a transformation that takes the system from the initial state to a final state equal to the initial one, then we obtain zero. This implies that, unlike heat differential dQ , the combination $\frac{dQ}{T}$ is an exact differential, namely, it is the differential of a new state function that we have already anticipated and which takes the name of entropy S . Hence, we can write:

$$dQ = T dS \quad (\text{A73})$$

where S is a function of state, for instance $S = S(T, V, N)$.

Thus, one introduces the following thermodynamic potentials, which are all functions of state:

(1) **Internal Energy**

$$U = U(T, V, N) \quad (\text{A74})$$

(2) **Helmholtz Free Energy**

$$F \equiv U - TS \quad (\text{A75})$$

(3) **Enthalpy**

$$H \equiv U + PV \quad (\text{A76})$$

(4) **Gibbs potential**

$$G \equiv H - TS = U + PV - TS = F + PV \quad (\text{A77})$$

The various thermodynamic potentials are all related to each other by relationships and are useful in describing thermodynamic transformations of various kinds, but their most important justification is the relationship of three of them, with the partition function, respectively, of the three possible ensembles in the microscopic, statistical mechanics description of macroscopic thermodynamic systems.

Appendix C.3. The Three Ensembles of Statistical Mechanics

Let us begin with the first of the three ensembles, which is perhaps the most fundamental of the three because it aims to directly explain disorder in terms of the enormous number of microscopic configurations corresponding to the same macroscopic state.

Appendix C.3.1. The Microcanonical Ensemble

Considering a number $N \gg 1$ of particles (molecules) that are collected in a volume V and have an overall energy E , we identify the internal energy of the system with the said energy:

$$U = E \quad (\text{A78})$$

and the entropy of the same system with:

$$S(U, V) = k_B \log(\mathcal{N}_E) \quad (\text{A79})$$

where k_B is the Boltzmann constant, and by definition

$$\mathcal{N}_E = \text{\#microscopic states that have an overall energy } E \quad (\text{A80})$$

From the first principle of thermodynamics, if we know the entropy function $S(U, V)$ we retrieve the temperature since:

$$\left. \frac{\partial S}{\partial U} \right|_{V=\text{const}} = \frac{1}{T} \quad (\text{A81})$$

Using the same logical procedure, we instead obtain:

$$T \left. \frac{\partial S}{\partial V} \right|_{U=\text{const}} = P \quad (\text{A82})$$

and all the equations of thermodynamics can be reconstructed from the knowledge of the function $S(U, V)$ defined by means of the identification in Equation (A79).

Appendix C.3.2. The Canonical Ensemble

In the microcanonical ensemble, the total energy $U = E$, the volume V , and the number of particles N are kept constant. In the canonical ensemble, on the other hand, the fixed energy condition is relaxed and only the volume V and the number of particles N constituting the system are kept fixed. Instead of energy, the temperature parameter T is introduced. With these elements, one then constructs the so-called **canonical partition function**. In the following way. Let $\Sigma(N, V)$ be the set of all possible microscopic states (at the level of classical mechanics, we can say configurations in phase space) that can be constructed with $N \gg 1$ particles (typically molecules) in volume V . Each state, i.e., each element $\sigma \in \Sigma(N, V)$ is endowed with a specific energy that we will denote E_σ . The canonical partition function is then written as the following sum:

$$Z_N(T, V) \equiv \sum_{\sigma \in \Sigma(N, V)} \exp \left[-\frac{E_\sigma}{k_B T} \right] \quad (\text{A83})$$

The connection with classical thermodynamics is made through the following identification:

$$F(T, V, N) = A(T, V, N) \equiv -k_B T \log[Z_N(T, V)] \quad (\text{A84})$$

where $F(T, V, N)$ is Helmholtz free energy introduced in Equation (A75) and $A(T, V, N)$ is the name we give to the combination to its left. The rationale for this identification becomes

clear if we reason in terms of probabilities and refer to the general scheme described in Appendix B in particular recalling Equations (43) and (44). The set of all states $\Sigma(N, V)$ is, in the present case, the measurable space Ω and the energy E_σ of a state $\sigma \in \Omega$ is the relevant stochastic vector $\mathbf{X}$. In this case the vector space $\mathbb{V}$ is one-dimensional because the energy E is a scalar quantity at least in classical mechanics. Finally we can identify:

$$\lambda = \frac{1}{k_B T} \quad (\text{A85})$$

and we see that the probability of the system being in state σ , given by:

$$p(\sigma) = \frac{\exp\left[-\frac{E_\sigma}{k_B T}\right]}{Z_N(T, V)} \quad (\text{A86})$$

coincides with that defined in the general Formula (43). For the same valid reason, in the general case, this probability density is correctly normalized to 1 if we sum over all states.

Note that the measurable set $\Sigma(N, V)$ is typically a variety of large dimensions and therefore the variables $\mathbf{q}$ that identify its points are many. For example, when the constituent particles of the system can be considered classical particles, $\Sigma(N, V)$ is the phase space for a system of N particles with 6^N dimensions. Each configurations σ in the phase space has a given energy E_σ .

We can now estimate the average energy of the system in the ensemble by writing:

$$\langle E \rangle = \sum_{\sigma \in \Sigma(N, V)} E_\sigma p(\sigma) \equiv k_B T^2 \partial_T \log[Z_N(T, V)] \quad (\text{A87})$$

The fundamental conceptual identification is that between the thermodynamic internal energy and the average energy of the canonical ensemble calculated in Equation (A87):

$$U = \langle E \rangle \quad (\text{A88})$$

Referring again to the general formulas in Appendix B we can compare Equation (A87) with the general Equation (46). Setting

$$\mathcal{H}(\lambda, N) = -\log[Z(T, N)] \quad (\text{A89})$$

and considering relation (A85) according to (46), we find:

$$\langle E \rangle = \frac{\partial}{\partial \lambda} \mathcal{H}(\lambda, N) = -\frac{\partial T}{\partial \lambda} \partial_T \log[Z(T, N)] = k_B T^2 \partial_T \log[Z_N(T, V)] \quad (\text{A90})$$

which exactly reproduces Equation (A87).

Let us compare Equation (A84) with the general Equation (47). Using the relation (A85) we get:

$$\mathcal{I}[p] = -\log[Z(T, N)] - \frac{1}{k_B T} U \quad (\text{A91})$$

and multiplying both the left and the right-hand side by $k_B T$, we find:

$$U + T(k_B \mathcal{I}[p]) = A(T, V, N) \equiv -k_B T \log[Z_N(T, V)] \quad (\text{A92})$$

Hence, recalling the definition of Helmholtz free energy (A75), we realize that the identification (A84) is the correct one if Shannon information entropy $\mathcal{I}[\mathbf{p}]$ is identified with thermodynamical entropy modulo the factor $-k_B$:

$$S(T, V, N) = -k_B \mathcal{I}[\mathbf{p}] \quad (\text{A93})$$

In classical thermodynamics, the identification between the function $A(T, V, N)$ and Helmholtz free energy $F(T, V, N)$ follows from the fact that we can show that F and A satisfy the same differential relation with internal energy.

Let us begin with the thermodynamic definition (A75). Utilizing both the first and second principles of thermodynamics, we get:

$$dF = dU - dTS - TdS \quad (\text{A94})$$

$$= TdS - PdV - dTS - TdS = -PdV - SdT \quad (\text{A95})$$

from which we work out the relations:

$$P = - \left. \frac{\partial F}{\partial V} \right|_{T=\text{const}}; \quad S = - \left. \frac{\partial F}{\partial T} \right|_{V=\text{const}} \quad (\text{A96})$$

Equation (A96) implies that the definition (A75) can be rewritten as the following differential equation:

$$F = U + T \left. \frac{\partial F}{\partial T} \right|_{V=\text{const}} \quad (\text{A97})$$

Reconsidering now Equation (A87) that provides the expression for internal energy, with obvious manipulations, we can write what follows:

$$\langle E \rangle = U = A - T \left. \frac{\partial A}{\partial T} \right|_{V=\text{const}} \quad (\text{A98})$$

which coincides with (A97) if $F = A$. Thus, the interpretation (A84) is fully justified within the setup of classical thermodynamics.

In the previous discussion, we emphasized the deeper sense of the classical identification in relation to Information Theory. In any case, having established identification (A84), the thermodynamic variables are all identified in the canonical ensemble as well. Summarizing, we have:

$$\begin{aligned} P &= \partial_V (k_B T \log[Z_N(T, V)]) \\ S &= \partial_T (k_B T \log[Z_N(T, V)]) \\ U &= k_B T^2 \partial_T \log[Z_N(T, V)] \\ F(T, V, N) &= -k_B T \log[Z_N(T, V)] \end{aligned} \quad (\text{A99})$$

Keeping the number N of particles or of moles fixed, the thermodynamic state of the system is always a point in a two-dimensional variety for which we can use either the native variables (T, V) which are the arguments of the partition function or any other pair of independent variables whose relations to (T, V) we can derive by inverting the relations (A99), whenever this is analytically possible.

Appendix C.3.3. The Grand Canonical Ensemble and the Gibbs Potential

In the formulation of statistical mechanics through the canonical grand ensemble, not only is the energy of states not fixed, but neither is the total number of particles or moles of the substance. Thus, the following grand partition function is written:

$$\mathcal{Z}(T, V, \mu) = \sum_{N=0}^{\infty} z^N Z_N(T, V) \quad (\text{A100})$$

where the variable $z = \exp\left[\frac{\mu}{k_B T}\right]$ is named the fugacity and the symbol μ is called the chemical potential. The relationship between the statistical description by the canonical grand partition function and classical thermodynamics is encoded in the identification:

$$\Phi(T, V, \mu) = -k_B T \log[\mathcal{Z}(T, V, \mu)] = U - TS - \mu N \quad (\text{A101})$$

In the grand canonical description of thermodynamics as well as in the canonical description, the internal energy of the system is not an a priori fixed datum, but rather it takes an average value that we calculate with the exact analogue of the Formula (A87)

$$U = k_B T^2 \partial_T \log[\mathcal{Z}(T, V, \mu)] \quad (\text{A102})$$

The same thing happens with the number of particles (or moles), which is not fixed but just assumes an average value calculated below:

$$\langle N \rangle = -k_B T \partial_\mu \log[\mathcal{Z}(T, V, \mu)] \quad (\text{A103})$$

Identification (A101) also allows us to write the other analogs of Equation (A99)

$$\begin{aligned} P &= \partial_V (k_B T \log[\mathcal{Z}(T, V, \mu)]) \\ U &= k_B T^2 \partial_T \log[\mathcal{Z}(T, V, \mu)] \\ S &= \partial_T (k_B T \log[\mathcal{Z}(T, V, \mu)]) \\ N &= -k_B T \partial_\mu \log[\mathcal{Z}(T, V, \mu)] \end{aligned} \quad (\text{A104})$$

and suggests the introduction of a new thermodynamic potential, called the Gibbs potential, which has the following definition

$$G(T, V, \mu) = PV + U - TS \quad (\text{A105})$$

Appendix C.4. Statistical Mechanics of Ideal Gases

After the general exposition of the previous section, we briefly present the fundamental example of the ideal gas of identical particles with masses all equal and not interacting with each other. As we recalled above, this is one of the rare cases where the partition function can be explicitly computed in closed form.

Our system consists of N non-relativistic particles, each with mass m , that are free to move inside a volume V , which, for simplicity, we regard as a cube with side $\ell = \sqrt[3]{V}$.

The phase space, in the Hamiltonian approach, is a space of dimension $2 \times 3 \times N$ whose coordinates are the N pairs $(\mathbf{p}_i, \mathbf{q}_i)$ where $\mathbf{p}_i = \{p_{i,x}, p_{i,y}, p_{i,z}\}$ is the momentum vector and $\mathbf{q}_i = \{q_{i,x}, q_{i,y}, q_{i,z}\}$ is the position vector of the i -th particle ($i = 1, \dots, N$).

The classical Hamiltonian of the system is very simple and is as follows:

$$\mathcal{H}(\mathbf{p}) = \sum_{i=1}^N \frac{\mathbf{p}_{x,i}^2 + \mathbf{p}_{y,i}^2 + \mathbf{p}_{z,i}^2}{2m} \quad (\text{A106})$$

In accordance with the general principles stated in previous sections, since the Hamiltonian represents the energy of the system, introducing Planck's constant h as the unit of measurement, the **canonical partition function** is written as follows:

$$Z_N(T, V) = \frac{1}{N!h^{3N}} \int \exp\left[-\frac{\mathfrak{H}(\mathbf{p})}{k_B T}\right] d^{3N}\mathbf{p} \times \int d^{3N}\mathbf{q} \quad (\text{A107})$$

Since the Hamiltonian does not depend on the coordinates $\mathbf{q}$, the last integral in $d^{3N}\mathbf{q}$ is calculated immediately, and we get V^N . Since the Hamiltonian is a sum of independent quadratic terms, the integral over moments is transformed into a product of integrals, and we have:

$$Z_N(T, V) = \frac{V^N}{N!} \times \prod_{i=1}^N \mathfrak{P}_{i,x} \mathfrak{P}_{i,y} \mathfrak{P}_{i,z} \quad (\text{A108})$$

, where

$$\mathfrak{P}_{i,x} = \int_{-\infty}^{+\infty} \frac{1}{h} \exp\left[-\frac{p_{i,x}^2}{2mk_B T}\right] dp_{i,x} \quad (\text{A109})$$

and similarly for $x \rightarrow y$ and $x \rightarrow z$. Everything then reduces to the calculation of a single Gaussian integral. Through the change of variable $t = p/\sqrt{2mk_B T}$, we obtain the final result:

$$Z_N(T, V) = \frac{V^N}{N!} \left(\frac{2\pi k_B T m}{h^2}\right)^{3N/2} \quad (\text{A110})$$

At this point, we are in a position to write the Helmholtz free energy as a function of temperature, volume and number of particles (or number of moles). It suffices to use the general Formula (A84) and apply it to the case of the free gas partition function calculated in (A110). We get:

$$\begin{aligned} F_{IG}(T, V, N) &= -k_B T \left(N \left(\frac{3}{2} \log\left(\frac{2\pi k_B T m}{h^2}\right) + \frac{3 \log(T)}{2} + \log(V) \right) - \log(\Gamma(N+1)) \right) \\ &\approx \frac{1}{2} k_B N T \left(-3 \log\left(\frac{2\pi k_B T m}{h^2}\right) + 2 \log(N) - 3 \log(T) - 2 \log(V) - 2 \right) \end{aligned} \quad (\text{A111})$$

where Γ denotes the Euler Gamma function e $\Gamma(N+1) = N!$ and where the second line is obtained from the first using the Stirling approximation $\log[\Gamma(N+1)] \approx N \log[N] - N$ which is very accurate for large values of N , as happens in our case.

Starting from Equation (A111) and using the general definitions (A99) and relation (A75), we obtain all thermodynamic state functions for ideal gases.

(PRESSURE)

$$P_{IG}(T, V, N) = -\partial_V F_{IG} = \frac{k_B N T}{V} \quad (\text{A112})$$

(ENTROPY)

$$S_{IG}(T, V, N) = -\partial_T F_{IG} = \frac{1}{2} k N \left(3 \log\left(\frac{2\pi k_B T m}{h^2}\right) - 2 \log(N) + 3 \log(T) + 2 \log(V) + 5 \right) \quad (\text{A113})$$

(INTERNAL ENERGY)

$$U_{IG}(T, V, N) = F_{IG}(T, V, N) + T S_{IG}(T, V, N) = \frac{3k_B N T}{2} \quad (\text{A114})$$

The Equation of State of Ideal Gases

Equation (A112) gives us the equation of state of ideal gases as a relation between pressure, volume, and temperature at a fixed number of moles N . Leaving out the subscript IG since it is clear that we are talking about ideal gases, we have:

$$P V = k_B N T \quad (\text{A115})$$

Appendix C.5. The Van Der Waals Model of a Real Gas

The first and still important example of modification of the gas equation of state is due to the Dutch physicist-mathematician van der Waals, who replaced Equation (A115) with the following one where two phenomenological parameters were introduced, (a, b) :

$$\left(P + \frac{an^2}{V^2} \right) (V - bn) = nRT; \quad (\text{A116})$$

In (A116) n denotes the number of moles of the chemical component of the gas, while, as before, P is the pressure, and V is geometrical volume in which the gas is enclosed. Furthermore, R is a universal physical constant with the following value:

$$R = 8.31446261815324 \frac{\text{J}}{\text{mol} \times \text{K}} \quad (\text{A117})$$

Definition A19. The Boltzmann constant, already utilized before, and that appears in all formulae of statistical mechanics is defined as the ratio of the ideal gas constant R and the Avogadro number (A119):

$$k_B \equiv \frac{R}{N_A} = 1.380649 \times 10^{-16} \text{ erg K}^{-1} \quad (\text{A118})$$

Definition A20. One *mole* of a chemical substance $\mathcal{X}$ is the quantity of atoms or of molecules of that substance $\mathcal{X}$ necessary to form a mass M numerically equal in grams to the weight $w_{\mathcal{X}}$ of an atom or a molecule of that substance $\mathcal{X}$ expressed in atomic mass units.

Thanks to such a definition one mole of any chemical substance, $\mathcal{X}$ always contains the same number of atoms or of molecules that is the Avogadro number:

$$N_A = 6.002214076 \times 10^{23} \quad (\text{A119})$$

In view of the Definition (A20) Avogadro number can be seen as the conversion factor from the standard mass unit, namely the gram and the atomic mass unit $\mathbf{u}$

$$1\text{g} = N_A \mathbf{u} \quad (\text{A120})$$

The two phenomenological parameters $a > 0$ and $b > 0$ were introduced to account for two effects. The first effect encrypted in parameter b takes into account the fact that each molecule does not have the entire geometric volume V at its disposal because the repulsive forces of the other molecules that are short-range create interdiction zones whose total volume is the greater, the larger the number of molecules present. The second effect modeled through the introduction of the parameter a is the reduction in the effective force exerted by molecular collisions on the walls of the container because of the attractive force, at longer radius, exerted on each molecule by all the first ones nearby. The physical dimensions of the introduced parameters can be read directly from the equation. Considering the number of moles dimensionless, we have:

$$[b] = \ell^3; [a] = m \ell^7 t^{-2} \quad (\text{A121})$$

where ℓ denotes length; m denotes mass; and t denotes time.

Skipping a lot of classical thermodynamic manipulation, we arrive at the conclusion that for the van der Waals model of a real gas Equation (78) is replaced by

$$\begin{aligned}\mathcal{A}_W(T, V) &= \frac{abn^3 - an^2V + nRTV^2}{V^2(V - bn)} \\ \mathcal{B}_W(T, V) &= n\frac{3}{2}RT - \frac{an^2}{V}\end{aligned}\quad (\text{A122})$$

Function (A122) satisfies the general constraint (76) of Lagrangian immersion. Correspondingly, the Riemannian metric that is induced on the Lagrangian variety as defined in Equation (77) takes, in the van der Waals case, the following explicit form:

$$ds_{VDW}^2 = \frac{n\left(-\frac{2TdV^2(RTV^3 - 2an(V - bn)^2)}{V^3(V - bn)^2} - 3RdT^2\right)}{2T^2}\quad (\text{A123})$$

The metric (A123) can be immediately rewritten in terms of zweibein 1-forms:

$$ds_{VDW}^2 = -\mathbf{e}^1 \otimes \mathbf{e}^1 - \mathbf{e}^2 \otimes \mathbf{e}^2\quad (\text{A124})$$

where

$$\begin{aligned}\mathbf{e}^1 &= \frac{\sqrt{\frac{3}{2}}dT\sqrt{nR}}{T} \\ \mathbf{e}^2 &= dV\sqrt{\frac{n(RTV^3 - 2an(V - \beta n)^2)}{TV^3(V - \beta n)^2}}\end{aligned}\quad (\text{A125})$$

It is very easy to calculate the unique component of the spin connection ω^{12} and the curvature 2-form that happens to be the following:

$$\begin{aligned}\mathfrak{R}_{VDW} &= R_{VDW}(T, V) \begin{pmatrix} 0 & \mathbf{e}^1 \wedge \mathbf{e}^2 \\ -\mathbf{e}^1 \wedge \mathbf{e}^2 & 0 \end{pmatrix} \\ R_{VDW}(T, V) &= \frac{2a(V - bn)^2(an(V - bn)^2 - RTV^3)}{3R(RTV^3 - 2an(V - bn)^2)^2}\end{aligned}\quad (\text{A126})$$

One immediately sees that if the two parameters a, b are set to zero, the curvature 2-form vanishes. Hence the equation of state of Ideal Gases corresponds to a *flat Lagrangian variety* deprived of phase transitions and critical phenomena. It follows that *thermodynamic curvature*, as it was first proposed by Ruppeiner [24] measures the interaction among molecules and defines critical phenomena through its own critical surfaces and critical curves.

What is essentially new compared with the equation of state for ideal gases is that the function $\mathcal{A}_W(V, n, T, a, b)$ possesses a critical point at which both its first and second derivatives with respect to the volume cancel. The two equations $\partial_V \mathcal{A}_W = 0$ and $\partial_V^2 \mathcal{A}_W = 0$ have a single solution for the temperature variable $\mathcal{T}$ and the volume variable V such that we find a single critical point for the three thermodynamic variables (P, V, T) , which is displayed in the following equation:

$$V_c = 3bn; \quad T_c = \frac{8a}{27bR} \quad \Rightarrow \quad P_c = \frac{a}{27b^2}\quad (\text{A127})$$

As can be seen, the critical temperature and pressure depend only on the parameters a, b , while the critical volume also depends on the number of moles. If we introduce dimensionless variables $\mathcal{T}, \mathfrak{v}, \mathfrak{p}$ defined as the ratios of T, V, P with respect to their critical values

$$T = \mathcal{T} T_c; \quad V = \mathfrak{v} V_c; \quad P = \mathfrak{p} P_c \quad (\text{A128})$$

in terms of such variables van der Waals equation of state becomes universal:

$$\mathfrak{p} = \frac{8\mathcal{T}}{3\mathfrak{v} - 1} - \frac{3}{\mathfrak{v}^2} \equiv \mathfrak{p}_W(\mathcal{T}, \mathfrak{v}) \quad (\text{A129})$$

It is very interesting to rewrite the function $R_{VDW}(T, V)$ in terms of the dimensionless variables introduced in Equation (A128). We get:

$$R_{VDW}(T, V) = \frac{1}{6nR} \mathcal{R}(\mathcal{T}, \mathfrak{v}) \equiv -\frac{(1 - 3\mathfrak{v})^2(8\mathcal{T}\mathfrak{v}^3 - 9\mathfrak{v}^2 + 6\mathfrak{v} - 1)}{(-4\mathcal{T}\mathfrak{v}^3 + 9\mathfrak{v}^2 - 6\mathfrak{v} + 1)^2} \quad (\text{A130})$$

The plot of function $\mathcal{R}(\mathcal{T}, \mathfrak{v})$ is shown in Figure A2. The singularity line is simply the vanishing locus in the $\{\mathfrak{v}, \mathcal{T}\}$ plane for the denominator in Equation (A130).

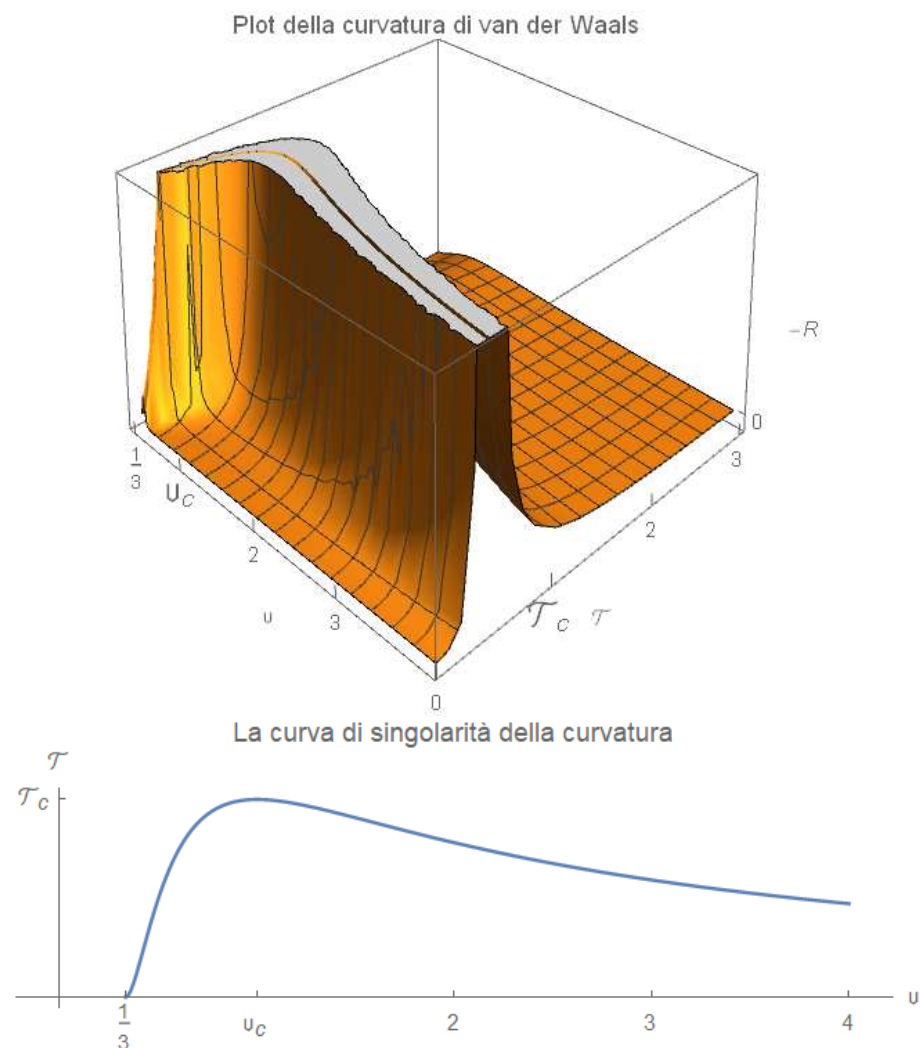


Figure A2. The first image shows the two-dimensional surface traced by the curvature scalar (A130) in the three-dimensional space spanned by the coordinates $\{\mathfrak{v}, \mathcal{T}, -\mathcal{R}\}$. One easily sees the singularity line along which thermodynamic curvature diverges to $-\infty$. Such a line projected onto the plane $\{\mathfrak{v}, \mathcal{T}\}$ is displayed in the second image.

Appendix D. The Example of the Kählerian Manifold $\mathcal{M}^{[2,2]}$ with Non Trivial Paint Group

Within the Tits Satake Universality Class (36), we choose the explicit case $q = 2$ in order to illustrate how the essential items in the formulation of generalized thermodynamics à la Souriau can be written in Paint Group covariant fashion and hence extended to the whole class. The symmetric space $\mathcal{M}^{[2,2]}$ of our example has dimension 8, which, therefore, is also the dimension of the corresponding solvable group $\mathcal{S}_8$ and of the corresponding solvable Lie algebra $Solv_8$:

$$\begin{aligned}\mathcal{M}^{[2,2]} &= \frac{SO(2,4)}{SO(2) \times SO(4)} \\ \dim[\mathcal{M}^{[2,2]}] &= 8 = \dim[\mathcal{S}_{[2,2]}] = \dim[Solv_{[2,2]}]\end{aligned}\quad (A131)$$

The chosen basis of generators of the solvable Lie algebra are displayed in Table A1 (for more details on their construction and normalization see [1]).

Table A1. The generators of the solvable Lie algebra $Solv_8$.

$T_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}; \quad T_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$	
$T_3 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}; \quad T_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$	
$T_5 = \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}; \quad T_6 = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$	
$T_7 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}; \quad T_8 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$	

Following the conventions and the theory exposed in [1,2] a generic element of the solvable Lie algebra is parameterized as follows:

$$Solv_8 \ni \mathbf{X}(\mathbf{Y}) = Y^1 T_1 + Y^2 T_2 + Y^3 T_3 + Y^4 T_4 + Y^{5,1} T_5 + Y^{5,2} T_6 + Y^{6,1} T_7 + Y^{6,2} T_8 \quad (A132)$$

The reason for the special naming of the solvable coordinates $\mathbf{Y}$ is the distinction between the long roots (generators $T_{3,4}$ associated with roots $\alpha_{3,4}$) that have no multiplicity and the short ones that have multiplicity and transform in the fundamental representation of the Paint Group G_{Paint} (see [1] for details). For the chosen example, the Paint Group is just $\text{SO}(2)$ and the solvable generators $T_{5,6}$ form the doublet of painted roots α_5 , while the solvable generators $T_{7,8}$ form the doublet of painted roots α_6 in the root system of $\mathfrak{so}(2,3) \simeq \mathfrak{sp}(4, \mathbb{R})$.

Following the conventions and notations of [1,2], the Σ exponential map from the solvable Lie algebra to the solvable group yields the generic element of the solvable group manifold in the form presented in Equation (3.56) of [1], that we repeat here for reader's convenience:

$$\mathbb{L}(\mathbf{Y})^{[2,1]} = \begin{pmatrix} e^{Y_1} & \frac{e^{Y_1} Y_3}{\sqrt{2}} & \frac{1}{2} e^{Y_1} (\sqrt{2} U_1 + Y_3 V_1) & \frac{1}{2} e^{Y_1} (\sqrt{2} U_2 + Y_3 V_2) & -\frac{1}{8} e^{Y_1} (4\mathbf{U} \cdot \mathbf{V} + \sqrt{2}(Y_3 \mathbf{V}^2 - 4Y_4)) & -\frac{1}{4} e^{Y_1} (\mathbf{U}^2 + 2Y_3 Y_4) \\ 0 & e^{Y_2} & \frac{e^{Y_2} V_1}{\sqrt{2}} & \frac{e^{Y_2} V_2}{\sqrt{2}} & -\frac{1}{4} e^{Y_2} \mathbf{V}^2 & -\frac{e^{Y_2} Y_4}{\sqrt{2}} \\ 0 & 0 & 1 & 0 & -\frac{V_1}{\sqrt{2}} & -\frac{U_1}{\sqrt{2}} \\ 0 & 0 & 0 & 1 & -\frac{V_2}{\sqrt{2}} & -\frac{U_2}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & e^{-Y_2} & -\frac{e^{-Y_2} Y_3}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 & e^{-Y_1} \end{pmatrix} \quad (\text{A133})$$

In Equation (A133) we have used the notation $U_i = Y_{5,i}$ and $V_i = Y_{6,i}$ (in our case $i = 1, 2$), which puts into evidence the existence of two Paint vectors $\mathbf{U}, \mathbf{V}$ in the case $r = 2$ and the Paint Group covariant structure of the solvable group element $\mathbb{L}(\mathbf{Y})^{[2,s]}$. Indeed, from Equation (A133), it is immediate to deduce the general form of the matrix for any value of q .

Starting from Equation (A133), we easily calculate all the further required items necessary for our argumentation. To begin with, we calculate the left-invariant 1-form:

$$\Theta \equiv \mathbb{L}^{-1} d\mathbb{L}$$

and then we project it onto the coset generators in the orthogonal decomposition of the full $\mathbb{U}$ Lie algebra:

$$\mathfrak{so}(2,4) = \underbrace{\mathfrak{so}(2) \oplus \mathfrak{so}(4)}_{\mathbb{H}} \oplus \mathbb{K} \quad (\text{A134})$$

The list of K generators is given in Table A2.

Table A2. The coset generators of the of $\text{SO}(2,4)/\text{SO}(2) \times \text{SO}(4)$.

$$K_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}; \quad K_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Table A2. Cont.

$$\begin{aligned}
K_3 &= \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix} ; K_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \end{pmatrix} \\
K_5 &= \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \end{pmatrix} ; K_6 = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix} \\
K_7 &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} ; K_8 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}
\end{aligned}$$

Since the K_i are normalized in such a way that $\text{Tr}(K_i \cdot K_j) = \delta_{ij}$, we can immediately calculate the vielbein as

$$V^i = \text{Tr}(K_i \Theta); \quad i = 1, \dots, 8 \quad (\text{A135})$$

The vielbeins are, as they should, linear combinations, with constant coefficients of the left-invariant 1-forms e^i , defined by the decomposition:

$$\Theta = \sum_{i=1}^8 e^i T_i \quad (\text{A136})$$

The latter have the following explicit appearance:

$$\begin{aligned}
e^1 &= dY_1 \\
e^2 &= dY_2 \\
e^3 &= dY_3 + Y_3(dY_1 - dY_2) \\
e^4 &= \frac{1}{4} \left(Y_{6,1}^2 (-dY_3) + Y_3 Y_{6,1}^2 dY_2 - 2\sqrt{2} Y_{6,1} dY_{5,1} - Y_{6,2}^2 dY_3 + Y_3 Y_{6,2}^2 dY_2 - 2\sqrt{2} Y_{6,2} dY_{5,2} \right. \\
&\quad \left. + (-Y_3 Y_{6,1}^2 - 2\sqrt{2} Y_{5,1} Y_{6,1} - Y_3 Y_{6,2}^2 - 2\sqrt{2} Y_{5,2} Y_{6,2} + 4Y_4) dY_1 + 4dY_4 + 4Y_4 dY_2 \right)
\end{aligned} \quad (\text{A137})$$

$$\begin{aligned}
e^{5,1} &= dY_{5,1} + \frac{Y_{6,1}(dY_3 - Y_3 dY_2)}{\sqrt{2}} + \left(Y_{5,1} + \frac{Y_3 Y_{6,1}}{\sqrt{2}} \right) dY_1 \\
e^{5,2} &= dY_{5,2} + \frac{Y_{6,2}(dY_3 - Y_3 dY_2)}{\sqrt{2}} + \left(Y_{5,2} + \frac{Y_3 Y_{6,2}}{\sqrt{2}} \right) dY_1 \\
e^{6,1} &= dY_{6,1} + Y_{6,1} dY_2 \\
e^{6,2} &= dY_{6,2} + Y_{6,2} dY_2
\end{aligned} \quad (\text{A138})$$

One finds the following relation between the vielbein and the left invariant 1-forms:

$$V^i = v_A^i e^A; \quad v = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \quad (\text{A139})$$

This result determines the form of the κ matrix in this case and consequently the expression of the $\text{SO}(2,4)$ invariant metric on the symmetric space (A131) in terms of the left-invariant 1-forms. Indeed, we have:

$$ds^2 = \kappa_{AB} e^A \otimes e^B$$

$$\kappa \equiv v^T \cdot v = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} \end{pmatrix} \quad (\text{A140})$$

to be compared with the result obtained in Equation (163) for the case of the maximally split master example $\text{SL}(3, \mathbb{R})/\text{SO}(3)$. As in all the other cases, the left-invariant 1-forms satisfy a set of Maurer–Cartan equations:

$$\begin{aligned} de^1 &= 0 \\ de^2 &= 0 \\ de^3 + \frac{1}{2}(2e^1 \wedge e^3 - 2e^2 \wedge e^3) &= 0 \\ de^4 + \frac{1}{2}(2e^1 \wedge e^4 + 2e^2 \wedge e^4 - \sqrt{2}e^{5,1} \wedge e^{6,1} - \sqrt{2}e^{5,2} \wedge e^{6,2}) &= 0 \\ de^{5,1} + \frac{1}{2}(2e^1 \wedge e^{5,1} + \sqrt{2}e^3 \wedge e^{6,1}) &= 0 \\ de^{5,2} + \frac{1}{2}(2e^1 \wedge e^{5,2} + \sqrt{2}e^3 \wedge e^{6,2}) &= 0 \\ de^{6,1} + e^2 \wedge e^{5,1} &= 0 \\ de^{6,2} + e^2 \wedge e^{5,2} &= 0 \end{aligned} \quad (\text{A141})$$

from which we read off the explicit value of the solvable Lie algebra structure constants f_{BC}^A since their general form is that given in Equations (156) and (157).

Observing Equation (A142), we immediately see how they can be generalized to any manifold of the TS universality class introducing the Paint index that runs in the

fundamental vector representation of the Paint Group $G_{\text{Paint}} = \text{SO}(q)$. The Paint coariant transcription of the Maurer–Cartan equation is the following:

$$\begin{aligned}
 de^1 &= 0 \\
 de^2 &= 0 \\
 de^3 + \frac{1}{2}(2e^1 \wedge e^3 - 2e^2 \wedge e^3) &= 0 \\
 de^4 + \frac{1}{2}\left(2e^1 \wedge e^4 + 2e^2 \wedge e^4 - \sqrt{2}\sum_{i=1}^q e^{5,i} \wedge e^{6,i}\right) &= 0 \\
 de^{5,i} + \frac{1}{2}\left(2e^1 \wedge e^{5,i} + \sqrt{2}e^3 \wedge e^{6,i}\right) &= 0 \\
 de^{6,i} + e^2 \wedge e^{5,i} &= 0
 \end{aligned} \tag{A142}$$

and it has the same form as the Maurer–Cartan equations on the Siegel plane, namely on the Tits Satake projection of the manifold. It suffices to restrict the Paint index i to the first value $i = 1$.

Appendix D.1. The Kähler 2-Form

The reason why the symmetric spaces (36) are all Kähler manifolds is the presence in the isotropy compact subgroup $H_c = \text{SO}(2) \times H'$ of the factor $\text{SO}(2) \simeq \text{U}(1)$ and the arrangement of the coset generator vector space $\mathbb{K}$ into a representation $(2 \mid \mathbf{v})$ where 2 is the doublet of $\text{SO}(2)$ and $\mathbf{v}$ is some irreducible representation of the other factor H' . All coset manifolds where such a situation is realized are Kähler manifolds, since the generator of $\text{SO}(2)$ in the $\mathbb{K}$ representation of H_c can be identified with the complex structure and leads to the explicit expression of the closed Kähler 2-form. In our specific case (A131) the $\text{SO}(2)$ -generator in the fundamental representation of $\text{SO}(2, 4)$ is the following one:

$$X^c = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix} \tag{A143}$$

and its representation on the space of K_i generators and hence on the vielbein V^i is the following one:

$$J^c = \begin{pmatrix} 0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \tag{A144}$$

The matrix J^c is obtained from the adjoint action of X^c on the K_i coset generators, namely from:

$$(J^c)_{ij} = \frac{1}{2} \text{Tr}([X^c, K_i] \cdot K_j) \quad (\text{A145})$$

and it squares to minus the identity:

$$J^c \cdot J^c = -\mathbf{1}_{8 \times 8} \quad (\text{A146})$$

hence it acts as a complex structure on the cotangent bundle (implying the same for the tangent bundle). Correspondingly, the Kähler 2-form can be written as:

$$\mathbf{K} = \sum_{a=1}^8 \sum_{b=1}^8 J_{ab}^c V^a \wedge V^b = -\frac{e^1 \wedge e^3}{\sqrt{2}} + \frac{e^1 \wedge e^4}{\sqrt{2}} + \frac{e^2 \wedge e^3}{\sqrt{2}} + \frac{e^2 \wedge e^4}{\sqrt{2}} - \frac{1}{2} \sum_{i=1}^q e^{5,i} \wedge e^{6,i} \quad (\text{A147})$$

where, once again, we have utilized the example under investigation to put the Paint invariance structure into evidence. In the present case, the index i takes only two values $i = 1, 2$ but the Formula (A147) applies to all values of q , namely to the entire Tits Satake universality class. In particular, for $q = 1$, Equation (A147) coincides, up to an overall factor, with Equation (303) corresponding to the Siegel plane. One easily verifies that $\mathbf{K}$ is closed and of maximal rank

$$d\mathbf{K} = 0; \quad \mathbf{K} \wedge \mathbf{K} \wedge \mathbf{K} \wedge \mathbf{K} = \text{const} e^1 \wedge e^2 \wedge \dots \wedge e^{6,2} \quad (\text{A148})$$

just as a consequence of the Maurer–Cartan Equation (A142).

The manifold (A131) equipped with the Kähler 2-form becomes a symplectic manifold $(\mathcal{M}^{[2,2]}, \mathbf{K})$ and because of metric equivalence, we can say that the solvable Lie group manifold $\mathcal{S}_{[2,2]}$ is a symplectic manifold $(\mathcal{S}_{[2,2]}, \mathbf{K})$.

Appendix D.2. The Hamiltonian Vector Fields and Their Moment-Maps

On the group manifold $\mathcal{S}_{[2,2]}$, we have both the left translations and the right ones, and, correspondingly, we have the left-invariant vector fields $\mathbf{t}_A^{[L]}$ that generate right translations and the right invariant ones $\mathbf{t}_A^{[R]}$ that generate left translations. Both sets of vector fields satisfy the solvable Lie algebra (157) with the structure constants defined by Equation (A142):

$$[\mathbf{t}_B^{[L]}, \mathbf{t}_C^{[L]}] = f_{BC}^A \mathbf{t}_A^{[L]}, \quad [\mathbf{t}_B^{[R]}, \mathbf{t}_C^{[R]}] = f_{BC}^A \mathbf{t}_A^{[R]} \quad (\text{A149})$$

The metric (A140) is invariant only with respect to the left-translations that are generated by the right-invariant vector fields, not with respect to both, since it is the metric of a coset manifold U/H . Correspondingly only $\mathbf{t}_A^{[R]}$ are Killing vectors for the Kähler metric and consequently the Kähler 2-form $\mathbf{K}$ admits as symplectic Killing vector fields only the set $\mathbf{t}_A^{[R]}$:

$$\begin{aligned} 0 &= \ell_{\mathbf{t}_A^{[R]}} \mathbf{K} \equiv i_{\mathbf{t}_A^{[R]}} \underbrace{d\mathbf{K}}_{=0} + d\left(i_{\mathbf{t}_A^{[R]}} \mathbf{K}\right) \\ &\Downarrow \\ i_{\mathbf{t}_A^{[R]}} \mathbf{K} &= d \underbrace{\mathfrak{P}_A(\mathbf{Y})}_{\text{moment map}} \end{aligned} \quad (\text{A150})$$

The second of Equation (A150) corresponds to the definition of the moment map functions $\mathfrak{P}_A(\mathbf{Y})$, such that

$$\begin{aligned}
\mathfrak{P} &: \mathbf{k} \longrightarrow \mathfrak{P}_{\mathbf{k}}(\mathbf{Y}) \in \mathbb{C}^\infty(\mathcal{S}_{[2,2]}) \\
\forall f(\mathbf{Y}) \in \mathbb{C}^\infty(\mathcal{S}_{[2,2]}) &: \mathbf{k}f = \{\mathfrak{P}_{\mathbf{k}}, f\} \equiv \mathbf{K}(\mathbf{k}, \mathbf{x}_f) \\
\mathbf{x}_f &\equiv \pi^{\alpha\beta}(\mathbf{Y}) \frac{\partial f}{\partial Y^\alpha} \frac{\partial}{\partial Y^\beta}; \quad \pi^{\alpha\beta} \equiv (\mathbf{K}^{-1})^{\alpha\beta}
\end{aligned} \tag{A151}$$

which are the exact analogues of Equations (125), (128), and (129). The difference, making explicit the clear-cut distinction advocated in Section 1.5 is that the symplectic manifold now is the very solvable Lie group manifold, rather than the total space of its tangent bundle, the moment maps are functions of the solvable coordinates Y^α rather than of the canonical momenta, and the symplectic form is the Kähler 2-form $\mathbf{K}$.

The very important point is that the moment maps $\mu_A(\mathbf{Y})$ that solve the differential equation in the second line of Equation (A150) have the closed form expression for all manifolds of the Tits Satake Universality class (36), discussed in Section 6.1.2 and presented in Equation (230).

We leave to a future publication the explicit calculation of the moment-maps and the study of the partition function that will closely follow the calculations already performed in the case of the Siegel plane. The main motivation for the present appendix was to show that all the structures analyzed in the main text for the Tits Satake projection of the entire universality class have obvious extensions to all members of the class by introducing Paint Group indices in the appropriate places.

References

1. Bruzzo, U.; Fré, P.G.; Trigiante, M. The Paint Group Tits Satake Theory of Hyperbolic Symmetric Spaces: The distance function, paint invariants and discrete subgroups. *arXiv* **2025**, arXiv:2503.07626.
2. Fré, P.G.; Milanesio, F.; Santoro, M.; Sanguinetti, G. Navigation through Non Compact Symmetric Spaces: A mathematical perspective on Cartan Neural Networks. *arXiv* **2025**, arXiv:2507.16871. [\[CrossRef\]](#)
3. Fré, P.G.; Milanesio, F.; Santoro, M.; Sanguinetti, G. Cartan Networks: Group theoretical Hyperbolic Deep Learning. *arXiv* **2025**, arXiv:2505.24353v1. [\[CrossRef\]](#)
4. Fré, P.; Milanesio, M.; Oyarzo, F.; Santoro, M.; Trigiante, M. Tessellation Groups, Harmonic Analysis on Non Compact Symmetric Spaces and the Heat kernel in view of Cartan Convolutional Neural Networks. *arXiv* **2025**, arXiv:2508.16015v1. [\[CrossRef\]](#)
5. Barbaresco, F. Innovative Tools for Radar Signal Processing based on Cartan's Geometry of SPD Matrices & Information Geometry. In Proceedings of the 2008 IEEE Radar Conference, Rome, Italy, 26–30 May 2008.
6. Dongho, C.; Duboskiy, P. Travelling Wave-like solutions of the Navier Stokes and related equations. *J. Math. Anal. Appl.* **1996**, *204*, 930–939.
7. Cartan, E. Sur une classe remarquable d'espaces de Riemann. *Bull. Soc. Math. Fr.* **1926**, *54*, 214–264. [\[CrossRef\]](#)
8. Helgason, S. *Differential Geometry and Symmetric Spaces*; Academic Press: Cambridge, MA, USA, 1962.
9. Fré, P.G. *Discrete, Finite and Lie Groups*; De Gruyter: Berlin, Germany; Boston, MA, USA, 2023. [\[CrossRef\]](#)
10. Magnea, U. An introduction to symmetric spaces. *arXiv* **2002**, arXiv:0205288. [\[CrossRef\]](#)
11. Kobayashi, S.; Nomizu, K. *Foundations of Differential Geometry*; Interscience Publishers: New York, NY, USA, 1963; Volume 1.
12. Alekseevsky, D. Classification of quaternionic spaces with a transitive solvable group of motions. *Math. USSR Izv.* **1975**, *9*, 297–339. [\[CrossRef\]](#)
13. Cortés, V. Alekseevskian spaces. *Diff. Geom. Appl.* **1996**, *6*, 129–168. [\[CrossRef\]](#)
14. Borel, A.; Tits, J. Groupes Réductifs. *Publications Mathématiques de l'IHES* **1965**, *27*, 55–151. [\[CrossRef\]](#)
15. de Wit, B.; Van Proeyen, A. Special geometry, cubic polynomials and homogeneous quaternionic spaces. *Commun. Math. Phys.* **1992**, *149*, 307–334. [\[CrossRef\]](#)
16. de Wit, B.; Van Proeyen, A. Broken sigma model isometries in very special geometry. *Phys. Lett.* **1992**, *B293*, 94–99. [\[CrossRef\]](#)
17. de Wit, B.; Van Proeyen, A. Isometries of special manifolds. *arXiv* **1995**, arXiv:9505097.
18. Fré, P.; Sorin, A.S. Integrability of supergravity billiards and the generalized Toda lattice equation. *Nucl. Phys.* **2006**, *B733*, 334–355. [\[CrossRef\]](#)
19. Fré, P.; Sorin, A.S. Supergravity black holes and billiards and the Liouville integrable structure associated with Borel algebras. *J. High Energy Phys.* **2010**, *2010*, 66. [\[CrossRef\]](#)

20. Fré, P.G.; Sorin, A.S. Classification of Arnold-Beltrami flows and their hidden symmetries. *Phys. Part. Nucl.* **2015**, *46*, 497–632. [[CrossRef](#)]
21. Fré, P.; Trigiante, M. Chaos from symmetry, Navier Stokes equations, Beltrami fields and the universal classifying crystallographic group. *J. Geom. Phys.* **2022**, *191*, 104884. [[CrossRef](#)]
22. Ruppeiner, G.; Seftas, A. Thermodynamic Curvature of the Binary van der Waals Fluid. *Entropy* **2020**, *22*, 1208. [[CrossRef](#)]
23. Weinhold, F. Metric geometry of thermodynamics. *Phys. Today* **1976**, *29*, 23. [[CrossRef](#)]
24. Ruppeiner, G. Thermodynamic curvature measures interactions. *Am. J. Phys.* **2010**, *78*, 1170–1180. [[CrossRef](#)]
25. Ruppeiner, G. Thermodynamic curvature from the critical point to the triple point. *Phys. Rev. E* **2012**, *86*, 021130. [[CrossRef](#)]
26. Ruppeiner, G.; Sahay, A.; Sarkar, T.; Sengupta, G. Thermodynamic geometry, phase transitions, and the Widom line. *Phys. Rev. E* **2012**, *86*, 052103. [[CrossRef](#)] [[PubMed](#)]
27. Ruppeiner, G. Thermodynamic curvature: Pure fluids to black holes. *J. Phys. Conf. Ser.* **2013**, *410*, 012138. [[CrossRef](#)]
28. Ruppeiner, G.; Mausbach, P.; May, H.-O. Thermodynamic R-diagrams reveal solid-like fluid states. *Phys. Lett. A* **2015**, *379*, 646–649. [[CrossRef](#)]
29. Lychagin, V. Contact Geometry, Measurement and Thermodynamics. In *Nonlinear PDEs, Their Geometry, and Applications, Proceedings of the Wisla 18 Summer School*; Springer Nature: Berlin/Heidelberg, Germany, 2019; pp. 3–54. [[CrossRef](#)]
30. Marle, C.-M. From Tools in Symplectic and Poisson Geometry to Souriau's Theories of Statistical Mechanics and Thermodynamics. *Entropy* **2016**, *18*, 370. [[CrossRef](#)]
31. 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](#)]
32. Barbaresco, F. Lie Group Machine Learning and Gibbs Density on Poincaré Unit Disk from Souriau Lie Group Thermodynamics and SU(1,1) Coadjoint Orbits. In *Geometric Science of Information*; Nielsen, F., Barbaresco, F., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 157–170. [[CrossRef](#)]
33. Barbaresco, F. Gaussian Distributions on the Space of Symmetric Positive Definite Matrices from Souriau's Gibbs State for Siegel Domains by Coadjoint Orbit and Moment Map. In *Geometric Science of Information*; Nielsen, F., Barbaresco, F., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; pp. 245–255. [[CrossRef](#)]
34. Marle, C.-M. On Gibbs states of mechanical systems with symmetries. *arXiv* **2021**, arXiv:2012.00582. [[CrossRef](#)]
35. Trigiante, M. Gauged supergravities. *Phys. Rep.* **2017**, *680*, 1–175. [[CrossRef](#)]
36. Fré, P.G. *Advances in Geometry and Lie Algebras from Supergravity*; Theoretical and Mathematical Physics Book Series; Springer: Berlin/Heidelberg, Germany, 2018. [[CrossRef](#)]
37. Andrianopoli, L.; Bertolini, M.; Ceresole, A.; D'Auria, R.; Ferrara, S.; Fre, P.; Magri, T. N = 2 supergravity and N = 2 superYang-Mills theory on general scalar manifolds: Symplectic covariance, gaugings and the momentum map. *J. Geom. Phys.* **1997**, *23*, 111–189. [[CrossRef](#)]
38. Fré, P.G. Lectures on resolutions à la Kronheimer of orbifold singularities, McKay quivers for Gauge Theories on D3 branes, and the issue of Ricci flat metrics on the resolved three-folds. *arXiv* **2023**, arXiv:2308.14022. [[CrossRef](#)]
39. Dwivedi, S.; Herman, J.; Jeffrey, L.; van den Hurk, T. *Hamiltonian Group Actions and Equivariant Cohomology*; SpringerBriefs in Mathematics; Springer: Berlin/Heidelberg, Germany, 2010.
40. Cabanes, Y. Multidimensional Complex Stationary Centered Gaussian Autoregressive Time Series Machine Learning in Poincaré and Siegel Disks: Application for Audio and Radar Clutter Classification. Ph.D. Thesis, Université de Bordeaux, Bordeaux, France, 2022.
41. Rao, C. Information and Accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* **1945**, *37*, 81–91.
42. Chentsov, N. Statistical Decision Rules and Optimal Inferences. *Trans. Math. Monog. Amer. Math. Soc. Provid.* **1982**, *53*.
43. Nielsen, F. An elementary introduction to information geometry. *Entropy* **2020**, *22*, 1100. [[CrossRef](#)] [[PubMed](#)]
44. Amari, S. Information Geometry and Its Applications. In *Applied Mathematical Sciences*; Springer: Tokyo, Japan, 2016.
45. Shannon, C. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
46. Koralov, L.B.; Sinai, Y.G. *Theory of Probability and Random Processes*; Springer: Berlin/Heidelberg, Germany, 2007.
47. Jaynes, E. Information Theory and Statistical Mechanics. *Phys. Rev.* **1957**, *106*, 620–630. [[CrossRef](#)]
48. Jaynes, E. Information Theory and Statistical Mechanics ii. *Phys. Rev.* **1957**, *108*, 171–190. [[CrossRef](#)]
49. Lychagin, V.; Roop, M. Phase transitions in filtration of real gases. *arXiv* **2019**, arXiv:1903.00276. [[CrossRef](#)]
50. Lychagin, V.; Roop, M. Steady filtration of Peng-Robinson gas in a porous medium. *arXiv* **2019**, arXiv:1904.08387. [[CrossRef](#)]
51. Lychagin, V.; Roop, M. On higher order structures in thermodynamics. *Entropy* **2020**, *22*, 1147. [[CrossRef](#)]
52. Kushner, A.; Lychagin, V.; Roop, M. Optimal thermodynamic processes for gases. *Entropy* **2020**, *22*, 448. [[CrossRef](#)]
53. Bianchi, M.; Bruzzo, U.; Fré, P.; Martelli, D. Resolution à la Kronheimer of $\mathbb{C}^3/\Gamma$ singularities and the Monge-Ampère equation for Ricci-flat Kähler metrics in view of d3-brane solutions of supergravity. *Lett. Math. Phys.* **2021**, *111*, 79. [[CrossRef](#)]
54. Bruzzo, U.; Fré, P.; Shahzad, U.; Trigiante, M. D3-brane supergravity solutions from Ricci-flat metrics on canonical bundles of Kähler-einstein surfaces. *Lett. Math. Phys.* **2023**, *113*, 64. [[CrossRef](#)]

55. Fré, P.G. *Gravity, a Geometrical Course*; Springer Science & Business Media: Dordrecht, The Netherlands, 2012; Volumes 1–2.
56. Alekseevsky, D.V.; Cortes, V.; Devchand, C.; Van Proeyen, A. Polyvector superPoincare algebras. *Commun. Math. Phys.* **2004**, *253*, 385–422. [\[CrossRef\]](#)
57. Arkhangel'skii, A. Completely Integrable Hamiltonian systems on a Group of Triangular Matrices. *Math. USSR Sb* **1980**, *36*, 127. [\[CrossRef\]](#)
58. Fre, P. Lectures on special Kahler geometry and electric–magnetic duality rotations. *Nucl. Phys. Proc. Suppl.* **1996**, *45BC*, 59–114. [\[CrossRef\]](#)
59. Nickel, M.; Kiela, D. Poincaré Embeddings for Learning Hierarchical Representations. In *Advances in Neural Information Processing Systems*; Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2017; Volume 30.
60. Nickel, M.; Kiela, D. Learning continuous hierarchies in the Lorentz model of hyperbolic geometry. In *Proceedings of the 35th International Conference on Machine Learning*; Dy, J., Krause, A., Eds.; PMLR: New York, NY, USA, 2018; Volume 80, pp. 3779–3788.
61. Klimovskaia, A.; Lopez-Paz, D.; Bottou, L.; Nickel, M. Poincaré maps for analyzing complex hierarchies in single-cell data. *Nat. Commun.* **2020**, *11*, 2966. [\[CrossRef\]](#)
62. Bonnasse-Gahot, L.; Nadal, J.P. Category learning in deep neural networks: Information content and geometry of internal representations. *Phys. Rev. E* **2025**, *112*, 055315. [\[CrossRef\]](#)
63. Jones, S.E.; Gilbert, A.D. Dynamo action in the abc flows using symmetries. *Geophys. Astrophys. Fluid Dyn.* **2014**, *108*, 83–116. [\[CrossRef\]](#)
64. Etnyre, J.; Ghrist, R. Contact topology and hydrodynamics: Beltrami fields and the seifert conjecture. *Nonlinearity* **2000**, *13*, 441–458. [\[CrossRef\]](#)
65. Ghrist, R. On the contact topology and geometry of ideal fluids. In *Handbook of Mathematical Fluid Dynamics*; North Holland: Amsterdam, The Netherlands, 2007; Volume 4, pp. 1–37.
66. Geiges, H. Contact geometry. *Handb. Differ. Geom.* **2006**, *2*, 315–382.
67. Cardona, R.; Miranda, E.; Peralta-Salas, D. Euler flows and singular geometric structures. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2019**, *377*, 20190034. [\[CrossRef\]](#)
68. Eva, M.; Oms, C. The geometry and topology of contact structures with singularities. *arXiv* **2021**, arXiv:1806.05638.
69. Cardona, R.; Miranda, E. On the volume elements of a manifold with transverse zeroes. *Regul. Chaotic Dyn.* **2019**, *24*, 187–197. [\[CrossRef\]](#)
70. Cardona, R.; Miranda, E. Integrable systems and closed one forms. *J. Geom. Phys.* **2018**, *131*, 204–209. [\[CrossRef\]](#)
71. Guillemin, V.; Miranda, E.; Pires, A.R. Symplectic and poisson geometry on b-manifolds. *Adv. Math.* **2014**, *264*, 864–896. [\[CrossRef\]](#)
72. Pollard, J.; Alexander, G.P. Singular contact geometry and beltrami fields in cholesteric liquid crystals. *arXiv* **2019**, arXiv:1911.10159. [\[CrossRef\]](#)
73. Arnold, V. *Metodi Matematici della Meccanica Classica*; Editori Riuniti, Edizioni Mir: Roma, Italy, 1979.
74. Huang, K. *Statistical Mechanics*; John Wiley & Sons: New York, NY, USA; London, UK; Sydney, Australia, 1963.
75. Uhlenbeck, G.; Ford George, G. *Lectures in Statistical Mechanics*; Lectures in Applied Mathematics; American Mathematical Society: Providence, RI, USA, 1963.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.