The Complex Geometry and Representation Theory of Statistical Transformation Models

Shuhao Li

Abstract
Given a measure space $\mathcal{X}$, we can construct a number of induced structures: eg. its $L^2$ space, the space $\mathcal{P}(\mathcal{X})$ of probability distributions on $\mathcal{X}$. If, in addition, $\mathcal{X}$ admits a transitive measure-preserving Lie group action, natural actions are induced on those structures. We expect relationships between these induced structures and actions. We study, in particular, the relations between $L^2(\mathcal{X})$ and exponential transformation models on $\mathcal{X}$, which are special “submanifolds” of $\mathcal{P}(\mathcal{X})$ closed under the induced action, whose tangent bundles are Kähler manifolds (given by Molitor). Geometrically, we show the tangent bundle has, locally, the “same” Kähler metric with the Fubini-Study metric on the projectivization of $L^2(\mathcal{X})$. Moreover we show the action on the tangent bundle is equivariant with that on $L^2(\mathcal{X})$, which is a unitary representation. Finally, in some cases, when the symplectic action on the tangent bundle is Hamiltonian, we show that any coadjoint orbit in the image of its moment map induces, via Kirillov’s correspondence from orbit method, irreducible unitary representations that are sub-representations of the aforementioned representation in $L^2(\mathcal{X})$.

1 Introduction
Given a measure space $\mathcal{X}$ with measure $\mu$, we can construct a number of natural algebraic or geometric structures from the measure space structure of $\mathcal{X}$: for example, the $L^2$ space of $(\mathcal{X}, \mu)$ with the Hilbert space structure, the space $\mathcal{P}(\mathcal{X})$ of all probability distributions over $\mu$ with the Fisher information metric, and in case $\mathcal{X}$ is a Riemannian manifold with the geometric measure from the metric, the cotangent bundle $T^*\mathcal{X}$ with the symplectic structure. In particular, the field of information geometry studies statistical manifolds, which are, in some sense, finite-dimensional immersed submanifolds of the space $\mathcal{P}(\mathcal{X})$ (which is in general infinite dimensional). These statistical manifolds have the structure of a Riemannian manifold (with the Fisher metric) and in addition a compatible pair of affine connections, called the exponential and mixture connections. A class of particularly nice examples of statistical manifolds are the exponential families. In [20], Molitor proves that the tangent bundle of a statistical manifold can be naturally
given a Kähler structure using a construction given by Dombrowski in [11], where Molitor demonstrates the connection of this construction with quantum mechanics.

If, in addition, a Lie group $G$ acts transitively on $\mathcal{X}$ and the measure $\mu$ is a $G$-invariant measure, this structure-preserving $G$-action can naturally give rise to actions on those algebraic and geometric structures built upon the measure space structure of $\mathcal{X}$: $G$ acts on the $L^2$ space by left translation, i.e. a function $f : \mathcal{X} \to \mathbb{C}$ is mapped to a function $gf$ by $g \in G$ such that $(gf)(x) = f(g^{-1}x)$ for any $x \in \mathcal{X}$. This action is in fact a unitary representation of $G$. A similar $G$-action on $\mathcal{P}(\mathcal{X})$ can be defined. Finally in case $\mathcal{X}$ is a Riemannian manifold with geometric measure such that $G$ acts on $\mathcal{X}$ by isometry in the Riemannian metric, $G$ can act on the cotangent bundle $T^*\mathcal{X}$ by pullback maps. If these new structures induced from $\mathcal{X}$ only depends on the measure space structure of $\mathcal{X}$ and the action of $G$ on $\mathcal{X}$ leaves the measure unchanged, then it is intuitive that these induced actions of $G$ on the new structures are structure-preserving, which we will make precise for some of the cases in the following sections. In particular, in e.g. [5, 6], Barndorff-Nielsen studies statistical transformation models, which are nice statistical manifolds in $\mathcal{P}(\mathcal{X})$ that are closed under the actions of $G$. Given the work by Molitor, we are particularly interested in exponential transformation models, that is, statistical transformation models which are also exponential families. The action of $G$ on the exponential transformation model gives rise to a $G$-action on the tangent bundle of the exponential transformation model by pushforward maps, and this induced action can be shown to be an action by isometries in the Kähler metric. Thus it is, in particular, a symplectic action.

It is then reasonable to expect some relationships between these induced structures from the measure space $\mathcal{X}$ and the induced structure-preserving $G$-actions on the different structures. In this work, we are interested in exploring the relationship between the tangent bundle $TM$ of an exponential transformation model $M$ with the $G$-action by Kähler isometry, and the $L^2$ space of $\mathcal{X}$ with the unitary $G$-representation, under some special circumstances. We shall show, generalizing from Molitor’s work in [20], that there is a $G$-equivariant map from $TM$ to $L^2(\mathcal{X})$, such that if we compose the map with the canonical projection map $L^2(\mathcal{X}) \to \mathbb{P}L^2(\mathcal{X})$, the resulting map respects the Kähler structures of $TM$ and the Fubini-Study Kähler metric of $\mathbb{P}L^2(\mathcal{X})$. Furthermore we assume that $\mathcal{X}$ has a normal stabilizer subgroup at some point in $\mathcal{X}$, and that the group is a semisimple nilpotent (or more generally exponential solvable) Lie group. We further assume that the $G$-action on the tangent bundle is not only symplectic but also Hamiltonian. Let $\mathfrak{g}^*$ be the dual Lie algebra of $G$. Kirillov’s orbit method (see e.g. [16, 13]) establishes a one-to-one correspondence between irreducible unitary representations of $G$ and coadjoint orbits of $G$ (i.e. the orbits in the coadjoint action of $G$ on $\mathfrak{g}^*$). Now in our situation, the moment map $\mu : TM \to \mathfrak{g}^*$ of the Hamiltonian action takes $TM$ to certain coadjoint orbits, which induces unitary irreducible representations of $G$. We will show that, if we add some more constraints to the Hamiltonian action, these representations can be realized as subrepresentations of the aforementioned unitary representation of $G$ in $L^2(\mathcal{X})$. In general, following this principle, we can expect more relationships between the geometric and algebraic structures induced from $\mathcal{X}$, which could reveal interesting
representation theoretic properties of $G$.

**Acknowledgement.** I would like to thank my advisor Professor Renato Feres for introducing me to this topic, and for the countless hours guiding me to develop this work. I would like to thank Professor Klaus Mohnke for pointing out the relation between the symplectic form in the Dombrowski’s construction and the canonical symplectic form on cotangent bundles. A large portion of this work is carried out during the Freiwald Scholars Program in the summers of 2019 and 2020.

2 Background: Exponential Families

This section presents definitions needed from information geometry and Molitor’s construction of Kählerification, following [1] and [20].

2.1 Information Geometry

Let $X$ be a measure space with measure $\mu$. We let $P(X)$ be the space of all probability distributions on $X$ under $\mu$:

$$P(X) = \{p : X \to \mathbb{R} | p \text{ is measurable}, p \geq 0, \int_{X} p(x) d\mu = 1\}.$$  \hspace{1cm} (1)

**Example.** If $X$ is a finite set with the counting measure, we index elements of $X$ from 1 to $n$, and denote the probability of the $i$-th element to be $p_i$. Then we have

$$P_n := P(X) = \{p : X \to \mathbb{R} | 0 \leq p_i \leq 1, \sum_{i=1}^{n} p_i = 1\}.$$  \hspace{1cm} (2)

This set can be viewed as a $(n - 1)$-dimensional manifold with boundary.

In general, of course, $P(\Omega)$ may be infinite-dimensional. A statistical model, loosely speaking, is a parametrized (finite-dimensional) submanifold of $P(X)$.

**Definition 2.1.** Let $X$ be a measure space with measure $\mu$. A **statistical model** is a $n$-dimensional differentiable manifold $M$ together with a map $j : M \to P(X)$ such that

1. $j$ is one-to-one;

2. For $\xi \in M$, $1 \leq i \leq n$, define functions $(\partial_i j)_{\xi}$ on $X$ as follows: pick a smooth path $\gamma(t)$ realizing $(\frac{\partial}{\partial \xi})_{\xi} \in T_\xi M$; then for each $x \in X$,

$$(\partial_i j)_{\xi}(x) = \frac{d}{dt} \bigg|_{t=0} \left[j(\gamma(t))\right](x).$$

We require $(\partial_1 j)_{\xi}, \ldots, (\partial_n j)_{\xi}$ to be $\mathbb{R}$ linearly independent as functions on $X$, for each $\xi \in M$ (i.e. $dj_\xi$ is one-to-one for every $\xi \in M$).
For the most common examples of (finite-dimensional) statistical models, there is usually a global chart \( M \rightarrow E \) (which we will always assume, for simplicity), where \( E \) is an open set of \( \mathbb{R}^n \). With some abuse of notations, we will usually write a statistical model as
\[
M = \{ p(x; \xi) \mid \xi = (\xi^1, \ldots, \xi^n) \in E \subset \mathbb{R}^n \},
\]
(3)
or simply write \( \xi \in \mathbb{R}^n \) as elements of \( S \).

**Example (Finite sets).** Suppose \( \Omega = \{x_1, \ldots, x_n\} \) is given the counting measure. Let \( E \subset \mathbb{R}^{n-1} \) be the set
\[
E = \{ (\xi^1, \ldots, \xi^{n-1}) \mid \xi^i > 1 \text{ for all } i, \text{ and } \sum_{i=1}^{n-1} \xi^i < 1 \}.
\]
(4)
The statistical model \( \mathcal{P}_n^\times \) is defined as
\[
\mathcal{P}_n^\times = \{ p_\xi : \Omega \rightarrow \mathbb{R} \mid \xi \in E \},
\]
(5)
where
\[
p_\xi(x_i) = \begin{cases} 
\xi^i & i = 1, \ldots, n-1 \\
1 - \sum_{i=1}^{n-1} \xi^i & i = n
\end{cases}.
\]
(6)
\( \mathcal{P}_n^\times \) is \( \mathcal{P}_n \) without the boundary. The map \( j \) can easily be checked to satisfy the two requirements since it can be viewed as an inclusion map \( j : \mathcal{P}_n^\times \hookrightarrow \mathcal{P}_n \).

**Example (Exponential Families).** An exponential family \( \mathcal{E} \) on a measure space \( \mathcal{X} \) with measure \( \mu \) is a statistical model
\[
M = \left\{ p(x; \xi) = \exp \left\{ C(x) + \sum_{i=1}^{n} \xi^i F_i(x) - \psi(\xi) \right\} \mid \xi \in E \subset \mathbb{R}^n \right\},
\]
(7)
where \( C(x), F_1(x), \ldots, F_n(x) : \mathcal{X} \rightarrow \mathbb{R} \) are measurable functions, \( \{1, F_1(x), \ldots, F_n(x)\} \) are linearly independent, and \( \psi \) is the normalizing constant
\[
\psi(\xi) = \ln \left( \int_{\mathcal{X}} \exp \left\{ C(x) + \sum_{i=1}^{n} \xi^i F_i(x) \right\} dx \right).
\]
(8)

Many common statistical models are exponential families. For example, \( \mathcal{P}_n^\times \) is an exponential family. For any appropriate \( \xi \), we have
\[
p_\xi(x_j) = \exp \left\{ \sum_{i=1}^{n} \xi^i F_i(x) - \psi(\xi) \right\},
\]
(9)
where \( F_i(x_j) = \delta_{ij} \), and \( \psi(\xi) \) is some appropriate normalization function. Other important examples of exponential families are Gaussian, binomial, exponential, Poisson. In this paper we will always consider exponential families.

Next, we will introduce certain quantities that are meaningful in the context of probability, which give a geometric structure to the manifold. We introduce the Fisher-information matrix on a statistical model \( M \), which will give the Riemannian geometric structure on the statistical model as a manifold.
Definition 2.2. The **Fisher-information matrix** at \( \xi \in M \) is the matrix with entries 
\( E[\partial_i l_\xi(x) \partial_j l_\xi(x)] \) at the \((i,j)\) position.

Theorem 2.1. The bilinear mapping \( h_\xi : T_\xi M \times T_\xi M \to \mathbb{R} \) on \( \xi \in M \) given by
\[
h_\xi(\frac{\partial}{\partial \xi^i}, \frac{\partial}{\partial \xi^j}) = h_{ij} = E[\partial_i l_\xi(x) \partial_j l_\xi(x)]
\] (10)
is a Riemannian metric (i.e. the Fisher-information matrix is symmetric and positive-definite). This metric on \( M \) is referred to as the **Fisher metric**.

A statistical model has more structure than a general Riemannian manifold. It has a two non-metrical connections \( \nabla^{(1)} \) and \( \nabla^{(-1)} \) that are related in a special way.

Definition 2.3. We give the Christoffel symbols of \( \nabla^{(1)} \) and \( \nabla^{(-1)} \):
\[
\Gamma^{(1)}_{ij,k} = E[(\partial_i \partial_j l_\xi)(\partial_k l_\xi)]; \tag{11}
\]
\[
\Gamma^{(-1)}_{ij,k} = E[(\partial_i \partial_j l_\xi)(\partial_k l_\xi)] + E[(\partial_i l_\xi \partial_j l_\xi \partial_k l_\xi)]. \tag{12}
\]
The pair of connections satisfy an analogue of the Leibniz law, which we formally state as follows.

Definition 2.4. A **dualistic structure** on a Riemannian manifold \( M \) with metric \( g \) is a pair \((\nabla, \nabla^*)\) of connections satisfying the compatibility condition with respect to the metric: for all \( X, Y, Z \in \Gamma(TM) \),
\[
Z(g(X,Y)) = h(Z \nabla X, Y) + h(X, \nabla^* Z Y). \tag{13}
\]

Definition 2.5. A **statistical manifold** is a Riemannian manifold with a dualistic structure \((\nabla, \nabla^*)\), with both connections non-metrical.

Theorem 2.2. On a statistical model \( S \), the two connections \( \nabla^{(1)} \) and \( \nabla^{(-1)} \) form a dualistic structure with respect to the Fisher metric. That is, a statistical model is a special statistical manifold.

Definition 2.6. For a Riemannian manifold with a dualistic structure \((M, g, \nabla, \nabla^*)\), we say \( M \) is **dually flat** if both \( \nabla \) and \( \nabla^* \) are flat, in the sense that both curvature and torsion vanish.

The main result that we are going to use later is this:

Theorem 2.3. An exponential family \((\mathcal{E}, h, \nabla^{(1)}, \nabla^{(-1)})\) is dually flat.
2.2 Tangent Bundles of Exponential Families

Given an exponential family with its canonical dualistic structure \((\mathcal{E}, h, \nabla^{(1)}, \nabla^{(-1)})\), we consider the geometric structure on its tangent bundle \(T\mathcal{E}\). The goal is to show that \(T\mathcal{E}\) has a canonical Kähler structure. The construction follows [11], [20].

**Theorem 2.4** (Dombrowski Splitting Theorem, [11]). For a manifold \(M\) with an affine connection \(\nabla\),

\[
T_{v_x}(TM) \simeq T_xM \oplus T_xM, \quad \forall x \in M, v_x \in T_xM. \tag{14}
\]

The isomorphism \(\Phi : T_{v_x}(TM) \to T_xM \oplus T_xM\) can be written as follows. Let \(\pi : TM \to M\) and \(\tilde{\pi} : T(TM) \to TM\). Let \(U\) be a “nice” neighborhood of \(x\); let \(\tau : \pi^{-1}(U) \to T_xM\) be the map that takes a tangent vector \(\theta' \in T_pM\), where \(p \in U\), to an element of \(T_xM\), such that \(\tau(\theta')\) is the tangent vector obtained from parallel translation through the unique geodesic from \(p\) to \(x\). Now, let \(K : T(TM) \to TM\) be defined as

\[
K : Z \in T(TM) \mapsto \lim_{t \to 0} \frac{\tau(\tilde{Z}(t)) - \tilde{\pi}(Z)}{t}, \tag{15}
\]

where \(\tilde{Z}(t) \subset TM\) is the curve in \(TM\) that realizes \(Z\) (i.e. \(\tilde{Z}(0) = \tilde{\pi}(Z), \frac{d}{dt}\bigg|_{t=0} \tilde{Z}(t) = Z\)).

For a point \(Z_{\theta_x}\) where \(\tilde{\pi}(Z_{\theta_x}) = \theta_x \in TM\) and \(\pi(\theta_x) = x \in M\), it is mapped by \(\Phi\) to

\[
\Phi : Z_{\theta_x} \in T(TM) \mapsto (\theta_x, (\pi_x)_*\theta_x(Z_{\theta_x}), K(Z_{\theta_x})). \tag{16}
\]

The second component is called the horizontal component of \(Z_{\theta_x}\), and the third component is called the vertical component of \(Z_{\theta_x}\). With certain abuse of notation, we now denote elements in \(T(TM)\) to be the triple \((\theta_x, l_x, v_x)\), where \(\theta_x = \tilde{\pi}(\theta_x, l_x, v_x)\), and \(l_x \in T_{\theta_x}M\) (resp. \(v_x \in T_{\theta_x}M\)) the horizontal (resp. vertical) component of \((\theta_x, l_x, v_x)\).

In particular, we have the direct-sum decomposition of the tangent space at each point \(\theta_x \in TM:\)

\[
T_{\theta_x}(TM) \simeq VT_{\theta_x}(TM) \oplus HT_{\theta_x}(TM), \tag{17}
\]

where \(VT_{\theta_x}(TM) = \ker(\pi_x)_*\theta_x\) is called the vertical subspace of \(T_{\theta_x}(TM)\), and \(HT_{\theta_x}(TM) = \ker K\) is the horizontal subspace of \(T_{\theta_x}(TM)\). It is easy to show that \(VT_{\theta_x}(TM) \simeq T_{\theta_x}M\) and \(HT_{\theta_x}(TM) \simeq T_{\theta_x}M\). This decomposition is important because of the following fact:

**Proposition 2.1.** If a even-dimensional real vector space \(V\) can be decomposed into the direct sum \(V = V_v \oplus V_h\), where \(V_v \simeq V_h\), there is a canonical almost complex structure on \(V\), i.e. an automorphism \(J : V \to V\) such that \(J^2 = Id_V\).

This means that there is a canonical almost complex structure on the manifold \(T(TM)\).

If, in addition, we have a Riemannian metric \(h\) in the original space \(M\) (the affine connection \(\nabla\) doesn’t need to be the Levi-Civita connection of \(h\)), then there is a natural almost-Hermitian structure on \(TM:\)

\[
g_{\theta_x}(\theta_x, l_x, v_x), (\theta_x', l'_x, v'_x)) = h_x(l_x, l'_x) + h_x(v_x, v'_x); \tag{18}
\]

\[
\omega_{\theta_x}(\theta_x, l_x, v_x), (\theta_x', l'_x, v'_x)) = h_x(l_x, v'_x) - h_x(v_x, l'_x); \tag{19}
\]

\[
J_{\theta_x}(\theta_x, l_x, v_x) = (\theta_x, -v_x, l_x). \tag{20}
\]


This is exactly analogous to the construction of the Kähler structure on $\mathbb{R}^{2n}$.

**Proposition 2.2.** For a Riemannian manifold $M$ with a dualistic structure, the canonical almost Hermitian structure that we have constructed on $TM$ is Kählerian if and only if the dualistic structure on $M$ is dually flat.

In short, a dually flat statistical manifold has a Kählerian tangent bundle.

**Corollary 2.1.** The tangent bundle $TE$ of an exponential family $\mathcal{E}$ is a Kähler manifold.

Finally we show that the symplectic structure on $TE$ is the same as the canonical symplectic structure on the cotangent bundle $T^*\mathcal{E}$, through the isomorphism given by the Fisher information metric $h$ on $\mathcal{E}$. The exponential connection $\nabla^{(1)}$ on $\mathcal{E}$ is flat, so we can take, at each point $x \in \mathcal{E}$, a local chart $\{q^1, \ldots, q^n\}$ with vanishing Christoffel symbols. Then in the induced chart on the tangent bundle $T\mathcal{E}$, $\{q^1, \ldots, q^n, v^1, \ldots, v^n\}$ corresponds to the tangent vector $\sum_i v^i \frac{\partial}{\partial q^i} \in T_{(q^1, \ldots, q^n)}\mathcal{E}$. By Dombrowski’s construction, the symplectic form associated to this chart is

$$\omega_{Dom} = \sum_i h_{ij} \, dv^i \wedge dq^i,$$

where $h_{ij}$ are the components of the Fisher information metric.

On the other hand, the canonical symplectic form on the cotangent bundle $T^*\mathcal{E}$ is, locally

$$\omega_{T^*\mathcal{E}} = \sum_i dp_i \wedge dq^i,$$

and the isomorphism between $T\mathcal{E}$ and $T^*\mathcal{E}$ induced by the metric $h$ sends the canonical symplectic form to

$$\omega_{TM} = \sum_i h_{ij} \, dv^i \wedge dq^i = \omega_{Dom}.\quad (23)$$

In particular, we see that for any flat connection on $\mathcal{E}$, including the mixture connection, the Dombrowski construction gives the same symplectic form. Therefore for an exponential family, using either exponential connection or mixture connection would give us the same symplectic form.

On the other hand, the almost complex structures and the Riemannian metrics given by different flat connections are different, since the almost complex structure depends on the horizontal subspace determined by the connection. The Riemannian metrics are not even conformally equivalent, in general (e.g. take two vectors, one horizontal in exponential connection and one horizontal in mixture connection, with the same horizontal component but different nonzero vertical component).
3 The Tangent Bundle of a Statistical Manifold

We consider a statistical manifold $M$ on a measure space $\mathcal{X}$ with measure $\mu$. We assume that probability measures in $M$ are absolutely continuous with respect to $\mu$, and thus we can identify all probability measures in $M$ by their density with respect to $\mu$. We denote the Fisher information metric on $M$ by $h$, and the $\alpha$-connections by $\nabla^{(\alpha)}$. On the tangent bundle $TM$ of $M$, we denote the almost Hermitian structure obtained by Dombrowski’s construction (using $h$ and the exponential connection $\nabla^{(1)}$) by $(g,J,\omega)$.

In [19], Molitor notes that for the statistical manifold of all positive probability density functions $\mathcal{P}^{\times}_n$ on the set of $n$ points (with counting measure), we can define a map $\tau : T\mathcal{P}^{\times}_n \rightarrow \mathbb{P}(\mathbb{C}^n)$ that preserves the almost Hermitian structure of $T\mathcal{P}^{\times}_n$, where we consider $\mathbb{P}(\mathbb{C}^n)$ as a Kähler manifold equipped with the Fubini-Study metric. Here we consider a simple generalization of the map for the tangent bundle $TM$, and show that the almost Hermitian structure is preserved under this map. Most of the computation in this section is a simple generalization of Molitor’s computation for probability measures on finite point sets in e.g. [19, 20] with some modifications.

For an element $(p,v) \in TM$ where $v \in T_p M$, let $p_t$ be a curve in $M$ that realizes $v$, i.e. such that $p_0 = p$ and $p_t|_{t=0} = v$. We can identify $v$ as a function $v : \mathcal{X} \rightarrow \mathbb{R}$ given by

$$p(x)v(x) = \left. \frac{d}{dt} \right|_{t=0} [p_t(x)]. \quad (24)$$

Note that $\int_\mathcal{X} v(x)p(x) d\mu = 0$ if $\int_\mathcal{X} \left. \frac{d}{dt} \right|_{t=0} [p_t(x)] d\mu = \left. \frac{d}{dt} \right|_{t=0} \int_\mathcal{X} p_t(x) d\mu$, which we shall always assume for our $M$.

Let $L^2(\mathcal{X})$ be the space of complex-valued $L^2$ functions on $\mathcal{X}$ with inner product $\langle f, g \rangle = \int_\mathcal{X} \overline{f} g d\mu$, let $L^2(\mathcal{X})$ be the corresponding Hilbert space (by identifying almost everywhere equal functions), and let $SL^2(\mathcal{X})$ be the unit sphere of $L^2(\mathcal{X})$. In general, $L^2(\mathcal{X})$ is an infinite dimensional Hilbert space. We think of $L^2(\mathcal{X})$ as an infinite-dimensional manifold with tangent spaces at each point canonically isomorphic to $L^2(\mathcal{X})$ itself with the same Hermitian inner product.

For each element $q \in L^2(\mathcal{X})$, we think of it as a mapping $q : \mathcal{X} \rightarrow \mathbb{C}$, so we can define functionals $L_x : L^2(\mathcal{X}) \rightarrow \mathbb{C}$ for each $x \in \mathcal{X}$ where $L_x(q) = q(x)$. We define a mapping $f : S \rightarrow L^2(\mathcal{X})$ from an arbitrary manifold $S$ to $L^2(\mathcal{X})$ to be differentiable if and only the mappings $L_x \circ f : S \rightarrow \mathbb{C}$ are differentiable for all $x \in \mathcal{X}$. If the map $f$ is nice (e.g. the condition for the Lebesgue dominated convergence theorem is satisfied for all difference quotients), then composed with the projection $L^2(\mathcal{X}) \rightarrow L^2(\mathcal{X})$, the map $S \rightarrow L^2(\mathcal{X})$ is differentiable in the Gâteaux sense.

For simplicity, in our formulation we will circumvent issues related to the infinite-dimensionality of the spaces, leaving out analytic details related to general infinite-dimensional manifolds, e.g. general notion of differentiability on Banach manifolds, if irrelevant to our specific examples (see e.g. [4] for reference on related issues in information geometry), since the more important aspect here is the metrics of the spaces.
For example if the statistical model given by $M \to \mathcal{P}(\mathcal{X})$ is nice, the mapping $\Phi : TM \to SL^2(\mathcal{X})$ defined by

$$\Phi(p, v)(x) = \sqrt{p(x)} \exp\left(\frac{iv(x)}{2}\right), \quad x \in \mathcal{X}.$$  \hspace{1cm} (25)

is Gâteaux differentiable.

Next, we can consider the (non-positive definite) inner products of $\mathcal{L}^2(\mathcal{X})$ as tangent spaces at each point on $\mathcal{L}^2(\mathcal{X})$ as a metric on the infinite-dimensional manifold, and we write $\int_\mathcal{X} \tilde{q}_1 q_2 d\mu = g_{L^2}(q_1, q_2) + i\omega_{L^2}(q_1, q_2)$. Since integration is insensitive to measure zero sets, the same metric can be defined on $L^2(\mathcal{X})$ by projection, which is now positive definite and can be thought of as a Kähler metric. We will then blur the distinction between $L^2(\mathcal{X})$ and $L^2(\mathcal{X})$; we subsequently understand all point-evaluations and differentiations of curves to take place on $L^2(\mathcal{X})$, and the evaluations of metric can be projected to $L^2(\mathcal{X})$.

Given the map $\Phi : TM \to L^2(\mathcal{X})$, for two curves $(p_t, v_t), (\tilde{p}_t, \tilde{v}_t)$ in $TM$ (where $p_t, \tilde{p}_t \in M, v_t, \tilde{v}_t \in T_{p_t}M, \tilde{v}_t \in T_{\tilde{p}_t}M$, where we identify $v, \tilde{v}$ as the functions $v(x), \tilde{v}(x)$ on $\mathcal{X}$ as defined before, and $p_0(x) = \tilde{p}_0(x) =: p(x), v_0(x) = \tilde{v}_0(x) =: v(x)$, we obtain the pull-back of $g_{L^2}$ and $\omega_{L^2}$ by

$$\Phi^* g_{L^2} \left( \frac{d}{dt} \bigg|_{t=0} (p_t, v_t), \frac{d}{dt} \bigg|_{t=0} (\tilde{p}_t, \tilde{v}_t) \right) = g_{L^2} \left( \frac{d}{dt} \bigg|_{t=0} \Phi(p_t, v_t), \frac{d}{dt} \bigg|_{t=0} \Phi(\tilde{p}_t, \tilde{v}_t) \right) \hspace{1cm} (26)$$

Writing $w_1(x) = \frac{1}{2} \left( \frac{d}{dt} \bigg|_{t=0} p_t(x) \right) / p(x)$ and $w_2(x) = \frac{1}{2} \left( \frac{d}{dt} \bigg|_{t=0} v_t(x) \right)$, and similarly $\tilde{w}_1(x)$ and $\tilde{w}_2(x)$, we obtain

$$\Phi^* g_{L^2} \left( \frac{d}{dt} \bigg|_{t=0} (p_t, v_t), \frac{d}{dt} \bigg|_{t=0} (\tilde{p}_t, \tilde{v}_t) \right) = \Re \int_{\mathcal{X}} \left( w_1 + i w_2 \right)(\tilde{w}_1 + i \tilde{w}_2) p(x) d\mu(x),$$

$$= \int_{\mathcal{X}} (w_1 \tilde{w}_1 + w_2 \tilde{w}_2) p(x) d\mu(x), \hspace{1cm} (27)$$

and similarly

$$\Phi^* \omega_{L^2} \left( \frac{d}{dt} \bigg|_{t=0} (p_t, v_t), \frac{d}{dt} \bigg|_{t=0} (\tilde{p}_t, \tilde{v}_t) \right) = \int_{\mathcal{X}} (w_1 \tilde{w}_2 - w_2 \tilde{w}_1) p(x) d\mu(x). \hspace{1cm} (28)$$

Now we consider the projective space $\mathbb{P}L^2(\mathcal{X}) \simeq (L^2(\mathcal{X}) - \{0\})/U(1)$ of the Hilbert space $L^2(\mathcal{X})$, and we consider an infinite-dimensional version of the Fubini-Study metric (we imitate the formulation of the metric in [20], in which the metric is defined for $\mathbb{C}P^n$). Given a point $z_0 \in SL^2(\mathcal{X})$, locally the projective space looks like the orthogonal complement $z_0^\perp$ the subspace $\mathbb{C}z_0$ in $L^2(\mathcal{X})$:

$$\phi_{z_0} : \{[z] \in \mathbb{P}L^2(\mathcal{X}) \mid \langle z_0, z \rangle \neq 0 \} \to z_0^\perp : \quad [z] \mapsto \frac{1}{\langle z_0, z \rangle} z - z_0. \hspace{1cm} (29)$$
We define the Fubini-Study metric $g_{FS}$ and symplectic form $\omega_{FS}$ on $\mathbb{P}L^2(\mathcal{X})$ by taking the real and imaginary components of the inner products of $z_0^\perp$ respectively, i.e. for $\xi_1, \xi_2 \in z_0^\perp$,

$$
\langle \xi_1, \xi_2 \rangle_{L^2} = ((\phi_{z_0}^{-1})^*g_{FS})(\xi_1, \xi_2) + i((\phi_{z_0}^{-1})^*\omega_{FS})(\xi_1, \xi_2).
$$

(30)

Now denote the projection map from $L^2(\mathcal{X}) - \{0\}$ to the projective space by $\pi_P : L^2(\mathcal{X}) - \{0\} \to \mathbb{P}L^2(\mathcal{X})$. For two curves $z_1(t), z_2(t)$ on $SL^2(\mathcal{X})$ such that $z_1(0) = z_2(0) = z_0$, we compute the Fubini-Study metric between $\frac{d}{dt}|_{t=0}[z_1(t)]$ and $\frac{d}{dt}|_{t=0}[z_2(t)]$:

$$
g_{FS}(\frac{d}{dt}|_{t=0}[z_1(t)], \frac{d}{dt}|_{t=0}[z_2(t)]) + i\omega_{FS}(\frac{d}{dt}|_{t=0}[z_1(t)], \frac{d}{dt}|_{t=0}[z_2(t)])
$$

$$
= \langle \frac{d}{dt}|_{t=0}\phi_{z_0}([z_1(t)]), \frac{d}{dt}|_{t=0}\phi_{z_0}([z_2(t)]) \rangle
= \langle \dot{z}_1(0) - z_0(\dot{z}_1(0)), \dot{z}_2(0) - z_0(\dot{z}_2(0)) \rangle
= \langle \dot{z}_1(0), \dot{z}_2(0) \rangle - \langle \dot{z}_1(0), z_0 \rangle \langle z_0, \dot{z}_2(0) \rangle.
$$

(31)

We now restrict our attention to an exponential family $M$. In particular, the following result is shown by Molitor in [20]:

**Lemma 3.1** (Corollary 4.3 and 4.4 in [20]). When $M$ is dually flat (e.g. when $M$ is an exponential family), $TM$ is a Kähler manifold with the almost Hermitian structure given by Dombrowski’s construction.

Now we show that the mapping $\Phi$ composed with the projection $\pi_P : L^2(\mathcal{X}) \to \mathbb{P}L^2(\mathcal{X})$ preserves the Kähler structure:

**Proposition 3.1** (Compare Proposition 3.3 in [19]). If $M$ an exponential family, the map $\pi_P \circ \Phi : TM \to \mathbb{P}L^2(\mathcal{X})$ is such that $(\pi_P \circ \Phi)^*g_{FS} = \frac{1}{4}g$, $(\pi_P \circ \Phi)^*\omega_{FS} = \frac{1}{4}\omega$, where $g$ and $\omega$ are obtained from Dombrowski’s construction.

**Proof.** First, we compute phase of the map $\Phi$ explicitly for an exponential family $M$. Suppose that $(\theta^1, \ldots, \theta^n)$ is the natural parameter of the exponential family and the probability density functions in $M$ are of the form

$$
p_\theta(x) = \exp \left\{ C(x) + \sum_{k=1}^n \theta^k F_k(x) - \psi(\theta) \right\}
$$

(32)

for measurable $C(x), F_1(x), \ldots, F_n(x)$ on $\mathcal{X}$ such that $1, F_1, \ldots, F_n$ are linearly independent, and $\psi$ a function on the parameter space. Now consider $(p, v) \in TM$. Say the tangent vector $v$ based at $p$ has coordinate $(v^k)$ in the natural parameters. Then the function on $\mathcal{X}$ corresponding to $v$ is

$$
v(x) = \left( \frac{d}{dt}|_{t=0} \exp \left\{ C(x) + \sum_{k=1}^n (\theta^k + tv^k) F_k(x) - \psi(\theta^k + tv^k) \right\} \right)/p_\theta(x)
$$

$$
= \sum_{k=1}^n v^k F_k(x) - \partial_\theta \psi(\theta).
$$

(33)
For every vertical vector on $TM$ based at $(p, v) \in TM$, we can realize it by a curve $(p, v + tu)$ for some $u \in T_p M$. This $u$ is the vertical component of the vector. So it can be easily computed that $w_1 = 0$ and $w_2(x) = \frac{1}{2} u(x)$. Moreover,

$$\left\langle \frac{d}{dt} \bigg|_{t=0} \Phi(p, v + tu), \Phi(p, v) \right\rangle = \left\langle \sqrt{p(x)} \exp \left( \frac{iv(x)}{2} \right), \sqrt{p(x)} \exp \left( \frac{iv(x)}{2} \right) \right\rangle = \int_X \frac{iu(x)}{2} p(x) d\mu = 0. \quad (34)$$

For every horizontal vector $u$ on $TM$ based at $(p, v) \in TM$, we first pick a curve $p_t$ in $M$ such that $\frac{d}{dt} \bigg|_{t=0} p_t$ is the horizontal component of the chosen horizontal vector, and $p_0 = p$. Now the natural parameters $(\theta_1, \ldots, \theta_n)$ of the exponential family $M$ is an affine coordinate of the exponential connection. We pick the curve $(p_t, v_t)$ in $TM$ where the coordinate of the tangent vector $v_t$ is constant with respect to $t$ in the natural parameters, so this curve is horizontal. Now

$$\frac{d}{dt} \bigg|_{t=0} v_t(x) = - \frac{d}{dt} \bigg|_{t=0} \partial_v \psi(\theta_t). \quad (35)$$

Note that this is constant on $X$. So for the horizontal vector $w_1(x) = \frac{1}{2} u(x), w_2(x) \equiv -\frac{1}{2} \frac{d}{dt} \bigg|_{t=0} \partial_v \psi(\theta_t)$. Moreover,

$$\left\langle \frac{d}{dt} \bigg|_{t=0} \Phi(p_t, v_t), \Phi(p, v) \right\rangle = \left\langle \sqrt{p(x)} \exp \left( \frac{iv(x)}{2} \right) (w_1(x) + iw_2(x)), \sqrt{p(x)} \exp \left( \frac{iv(x)}{2} \right) \right\rangle = \int_X \frac{iu(x)}{2} p(x) d\mu - i \frac{1}{2} \int_X \frac{d}{dt} \bigg|_{t=0} \partial_v \psi(\theta_t) p(x) d\mu = -\frac{i}{2} \frac{d}{dt} \bigg|_{t=0} \partial_v \psi(\theta_t). \quad (36)$$

We can now compare the Fubini-Study Kähler metric with the Kähler metric from Dombrowski’s construction. We compute $(\pi_P \circ \Phi)^* g_{FS}$ and $(\pi_P \circ \Phi)^* \omega_{FS}$ on horizontal and vertical vectors using equations (27), (28), (31), and compare them with Dombrowski’s construction. Let $v, u, \tilde{u} \in T_p M$, and suppose $\theta_t, \tilde{\theta}_t$ parameterize curves in $M$ that realize $u, \tilde{u}$ respectively. Write $C = -\frac{d}{dt} \bigg|_{t=0} \partial_v \psi(\theta_t), \bar{C} = -\frac{d}{dt} \bigg|_{t=0} \partial_v \psi(\tilde{\theta}_t)$. $C$ and $\bar{C}$ are constant on $X$.

- For two vertical vectors $w, \tilde{w} \in T_{(p, v)} TM$ with vertical components $u, \tilde{u} \in T_p M$ respectively,

$$((\pi_P \circ \Phi)^* g_{FS}(w, \tilde{w}) = \frac{1}{4} \int_X u(x) \tilde{u}(x) p(x) d\mu; \quad ((\pi_P \circ \Phi)^* \omega_{FS}(w, \tilde{w}) = 0. \quad (37)$$
• For two horizontal vectors \( w, \tilde{w} \in T_{(p,v)} TM \) with horizontal components \( u, \tilde{u} \in T_p M \) respectively,

\[
(\pi_P \circ \Phi)^* g_{FS}(w, \tilde{w}) = \frac{1}{4} \left( \int_X u(x)\tilde{u}(x) p(x) d\mu + C\tilde{C} - (iC)(i\tilde{C}) \right)
= \frac{1}{4} \int_X u(x)\tilde{u}(x) p(x) d\mu;
\]

\[
(\pi_P \circ \Phi)^* \omega_{FS}(w, \tilde{w}) = \frac{1}{4} \int_X (\tilde{C}u(x) - C\tilde{u}(x)) p(x) d\mu = 0. \tag{38}
\]

• For a vertical vector \( w \in T_{(p,v)} TM \) with vertical component \( u \in T_p M \) and a horizontal vector \( \tilde{w} \in T_{(p,v)} TM \) with horizontal component \( \tilde{u} \in T_p M \),

\[
(\pi_P \circ \Phi)^* g_{FS}(w, \tilde{w}) = \frac{1}{4} \int_X \tilde{C}u(x) p(x) d\mu = 0;
\]

\[
(\pi_P \circ \Phi)^* \omega_{FS}(w, \tilde{w}) = \frac{1}{4} \int_X u(x)\tilde{u}(x) p(x) d\mu. \tag{39}
\]

The results all agree with the Kähler metric from Dombrowski’s construction (other than an \( 1/4 \) factor), so we have proved the proposition.

This computation establishes a connection between the Fisher metric in information geometry and the Fubini-Study metric in complex geometry.

4 Statistical Transformation Models and Complex Projective Spaces

We now consider a statistical transformation model \( M \) on \( \mathcal{X} \) with acting group \( G \), i.e.

• \( M \) is a statistical manifold on the sample space \( \mathcal{X} \);
• There is an action of \( G \) on the sample space \( \mathcal{X} \);
• The action of \( G \) on \( M \) induced by the pushforward of the action of \( G \) on \( \mathcal{X} \) is invariant on \( M \), i.e. for any probability measure \( P \in M \), the probability measure \( gP \) defined by

\[
(gP)(A) = P(g^{-1}A) \tag{40}
\]

for any measurable set \( A \) in \( \mathcal{X} \). Furthermore we assume the action of \( G \) on \( M \) is transitive.
In addition, we assume that \( X \) is a Riemannian manifold with the geometric measure \( \mu \) invariant under the action of \( G \) on \( X \), and we assume that all probability measures in \( M \) are absolutely continuous with respect to the geometric measure. In this way, if \( p \) is a density function (with respect to \( \mu \)) for a probability measure \( P \) in \( M \), then \( p(g^{-1}x) \) is the density function of \( gP \) with respect to \( \mu \). We usually assume that the density function is smooth on \( X \).

### 4.1 The Induced Action on Complex Projective Spaces

We consider \( L^2(X) \) as a Hilbert manifold as before. Now consider the trivial \( U(1) \)-principal bundle over \( X \). The space of \( L^2 \) sections of this bundle can be viewed as a group \( G \) with fiber-wise multiplication, and there is an action of \( G \) on the unit sphere \( SL^2(X) \) of the Hilbert manifold \( L^2(X) \). The orbit space of this action can be identified to be \( P(X) \), the space of all probability density functions over the geometric measure of \( X \), and the orbit projection map \( \pi : SL^2(X) \rightarrow P(X) \) can be identified to be \( \pi(q) = p \in P(X) \) where \( p(x) = |q(x)|^2 \). In particular, consider the subgroup of the bundle \( G \) consisting only of constant sections; this subgroup can be identified with \( U(1) \). Therefore the projection map \( \pi : SL^2(X) \rightarrow P(X) \) can be factorized into a projection \( \pi_\mathcal{P} : SL^2(X) \rightarrow \mathbb{P}L^2(X) \) and then a map \( \pi_\mathcal{P} : \mathbb{P}L^2(X) \rightarrow P(X) \).

Now we consider a lift of the action of \( G \) on \( M \) to an action of \( G \) on the image of the map \( \pi_\mathcal{P} \circ \Phi \). Note that \( \text{Im} \pi_\mathcal{P} \circ \Phi \subseteq \pi_\mathcal{P}^{-1}(M) \). Write \( N = \text{Im} \pi_\mathcal{P} \circ \Phi \). An element in \( N \) can be identified with a class \([q(x)]\) in \( \mathbb{P}L^2(X) \) where \( q(x) \in L^2(X) \). For \( g \in G \), we let \( g[q(x)] = [g(g^{-1}x)] \). This is well-defined, and respects the group action. We will further show that \( g[q(x)] \in N \) for every \( g \in G, [q(x)] \in N \). For this, we consider the action of \( G \) on \( TM \) given by \( g(p, v) = (gp, g_*v) \) where \( (p, v) \in TM \) and we regard \( g \) as a diffeomorphism on \( M \) and write \( g_*v \) to be pushforward of \( v \) under \( g \).

**Lemma 4.1.** \( N \) is invariant under the action of \( G \), and the map \( \pi_\mathcal{P} \circ \Phi \) is \( G \)-equivariant.

**Proof.** Given \( (p, v) \in TM \) and \( g \in G \), let \( p_t \) be a curve in \( M \) realizing \( v \). The tangent vector \( g_*v \in T_{gp}M \) can be identified with the function

\[
(g_*v)(x) = \frac{d}{dt}igr|_{t=0} gp_t(x) = \frac{d}{dt}igr|_{t=0} p_t(g^{-1}x) = v(g^{-1}x).
\]

(41)

Now \( \Phi(g(p, v)) = \sqrt{p(g^{-1}x)} \exp \left( \frac{iv(g^{-1}x)}{2} \right) = g\Phi(p, v) \). This shows that \( N = \text{Im} \pi_\mathcal{P} \circ \Phi \) is invariant under the action of \( G \) and that \( \pi_\mathcal{P} \circ \Phi \) is a \( G \)-equivariant map. \( \square \)

We notice that \( G \) acts by isometry (with respect to Fisher metric) on \( M \):

**Lemma 4.2 (Chapter 3.8 of [5], or Proposition 2.2 and 2.3 of [12]).** The action of \( G \) on \( M \) induced from \( G \) acting on \( X \) preserves both the Fisher metric and the exponential connection.
Motivated by this lemma and Proposition 3.1, we consider the Fubini-Study metric of $N$ under the action of $G$. For this we first consider the action of $G$ on $TM$.

**Lemma 4.3.** The action of $G$ on $TM$ preserves the Kähler metric given by Dombrowski’s construction.

**Proof.** For a vertical vector $u^v \in T_{(p,v)}TM$ with vertical component $u \in T_pM$, we consider its horizontal and vertical component after the pushforward of the diffeomorphism $g \in G$. Denote the projection map of the tangent bundle to be $\pi : TM \to M$. The vertical vector $u^v$ can be realized by a curve $(p, v + tu)$ in $TM$. The horizontal component of the pushforwarded vector is

$$\pi_*g_*(u^v) = \left. \frac{d}{dt} \right|_{t=0} \pi(g(p, v + tu)) = 0. \quad (42)$$

Let $K : TTM \to TM$ be the connector map for the exponential connection of $M$. The vertical component of the pushforwarded vector is

$$K(g_*(u^v)) = K\left( \left. \frac{d}{dt} \right|_{t=0} (gp, g_*v + tg_*(u)) \right) = g_*u. \quad (43)$$

Therefore $g_*(u^v)$ is a vertical vector with vertical component $g_*u$.

For a horizontal vector $u^h \in T_{(p,v)}TM$ with horizontal component $u \in T_pM$, $u^h$ can be realized by the horizontal lift to $TM$ of a curve $p_t$ in $M$ that realizes $u$. Denote the lifted curve to be $(p_t, v_t) \in TM$. The horizontal component of the pushforwarded vector is

$$\pi_*g_*(u^h) = \left. \frac{d}{dt} \right|_{t=0} \pi(g(p_t, v_t)) = \left. \frac{d}{dt} \right|_{t=0} gp_t = g_*u. \quad (44)$$

The vertical component of the pushforwarded vector is

$$K(g_*(u^h)) = K\left( \left. \frac{d}{dt} \right|_{t=0} g(p_t, v_t) = \frac{D^{(e)}}{dt} \right|_{t=0} g(p_t, v_t) = g_* \frac{D^{(e)}}{dt} \right|_{t=0} (p_t, v_t) = 0 \quad (45)$$

where $\frac{D^{(e)}}{dt}$ is the covariant derivative for the exponential connection on $M$, and the second to last equality holds because as a diffeomorphism of $M$, $g$ is an affine transformation with respect to the exponential connection (Lemma 4.2). Therefore $g_*(u^h)$ is a horizontal vector with horizontal component $g_*u$.

Finally, since by Lemma 4.2 $G$ acts on $M$ by isometry, we know that $h_{gp}(g_*u, g_*\bar{u}) = h_p(u, \bar{u})$ for $h$ the Fisher metric on $M$, $g \in G$, and $u, \bar{u} \in T_pM$. By Dombrowski’s construction of the Kähler metric on $TM$, the action of $G$ on $TM$ preserves the Kähler metric. \qed

Now this action of $G$ on $TM$ preserving Kähler metric transfers, via the $G$-equivariant map $\pi_F \circ \Phi$ which also preserves the Kähler structure (up to a constant factor), to an action of $G$ on $N$. Write the Fubini-Study Kähler metric as $G_{FS} = g_{FS} + i\omega_{FS}$, and the Kähler metric from Dombrowski’s construction as $G_{Dom} = g_{Dom} + i\omega_{Dom}$. 

14
Corollary 4.1. The action of $G$ on $N$ preserves the Fubini-Study Kähler metric, in the sense that $G_{FS}(g*w, g*\tilde{w}) = G_{FS}(w, \tilde{w})$ for all $w, \tilde{w} \in (\pi_P \circ \Phi)_* [T_{(p,v)}TM]$, $g \in G$, $(p,v) \in TM$.

Proof. Given two $u, \tilde{u} \in T_{(p,v)}TM$, from Proposition 3.1 we know that

\[
G_{FS}((\pi_P \circ \Phi)_* u, ((\pi_P \circ \Phi)_* \tilde{u})) = (\pi_P \circ \Phi)^* G_{FS}(u, \tilde{u}) = \frac{1}{4} G_{Dom}(u, \tilde{u})
\]  

(46)

and that

\[
G_{FS}((\pi_P \circ \Phi)_* (g*u), (\pi_P \circ \Phi)_* (g*\tilde{u})) = (\pi_P \circ \Phi)^* G_{FS}(g_* u, g_* \tilde{u}) = \frac{1}{4} G_{Dom}(g_* u, g_* \tilde{u}).
\]  

(47)

But by Lemma 4.3 $G_{Dom}(g_* u, g_* \tilde{u}) = G_{Dom}(u, \tilde{u})$. Finally, since $\pi_P \circ \Phi$ is $G$-equivariant,

\[
(\pi_P \circ \Phi)_* (g_* u) = (\pi_P \circ \Phi \circ g)_* u = (g \circ \pi_P \circ \Phi)_* u = g_* (\pi_P \circ \Phi)_* u.
\]  

(48)

and similarly for $\tilde{u}$. Therefore we can conclude that

\[
G_{FS}(g_* w, g_* \tilde{w}) = G_{FS}(w, \tilde{w}),
\]  

(49)

for all $w, \tilde{w} \in (\pi_P \circ \Phi)_* [T_{(p,v)}TM]$.

\[
\square
\]

5 Moment Maps and The Orbit Method

We first briefly recall the constructions of induced representations and Kirillov’s orbit method that gives a correspondence between the coadjoint orbits of a group and unitary irreducible representations of the group. We then proceed to prove that in certain situations, the left-translation unitary representation of $G$ in $L^2(\mathcal{X})$ contains as subrepresentations unitary irreducible representations corresponding to the coadjoint orbits in the image of the moment map of the action of the group on the tangent bundle of an exponential transformation model (if the action is Hamiltonian).

5.1 Induced Representation

We describe the construction of induced representations in a special situation pertaining to our setting (see [9]; for the general version of this construction, see [13]). Given a locally compact group $G$ and a closed subgroup $H \leq G$, we assume the homogeneous space $G/H$ possesses an invariant measure $\mu$ (which is indeed the assumption for a statistical transformation model). Suppose $H$ has a unitary representation $\sigma$ on the Hilbert space $\mathcal{H}_\sigma$. We want to construct out of this representation $\sigma$ on $H$ a unitary representation $\pi = \text{ind}^G_H(\sigma)$ on $G$. We first construct a Hilbert space $\mathcal{H}_\pi$ consisting of all Borel measurable functions $f : G \to \mathcal{H}_\sigma$ such that:
• Contravariance: for any \( h \in H, g \in G, \sigma(h^{-1})f(g) \),

\[
 f(gh) = \sigma(h^{-1})f(g) \tag{50}
\]

Note that \( \| f(g) \|^2 \) is constant on cosets by the contravariance property, so the integral (50) is well-defined.

We define an inner product on the set \( \mathcal{H}_\pi \):

\[
 \langle f, \tilde{f} \rangle = \int_{G/H} \langle f(g), \tilde{f}(g) \rangle_{\mathcal{H}_\sigma} \, d\mu(gH). \tag{51}
\]

Again, this integral is well-defined by the contravariance property. Thus the set \( \mathcal{H}_\pi \) is made into a Hilbert space, and then we can define the representation \( \pi = \text{ind}_H^G(\sigma) \) by

\[
 [\pi(g)f](x) = f(g^{-1}x). \tag{52}
\]

for any \( g \in G, x \in G/H \). By the \( G \)-invariance property of the measure \( \mu \), \( \pi \) is a unitary representation of \( G \) with Hilbert space \( \mathcal{H}_\pi \).

### 5.2 Coadjoint Orbits and Kirillov’s Orbit Method

Let \( G \) be a nilpotent Lie group. Let \( g^* \) be the dual Lie algebra of \( G \). Kirillov’s orbit method gives a bijective correspondence between orbits of the coadjoint action of \( G \) on \( g^* \) and unitary irreducible representations of \( G \). We describe this correspondence, and for details of Kirillov’s orbit method we refer to e.g. \([16, 9, 13]\) (the presentation of this section mainly follows \([13]\)).

Let \( \lambda \in g^* \). We write \( \mathcal{O}_\lambda \) for its coadjoint orbit. Let \( h \) be a maximal subalgebra in \( g \) such that \( \lambda \) restricts to 0 on \( [h, h] \) (see e.g. Theorem 2.2.1 in \([9]\) for the existence of such a subalgebra). We can then define a one-dimensional representation \( \sigma_\lambda \) as follows: for a nilpotent Lie group \( G \), \( H = \exp h \) is a connected subgroup with Lie algebra \( h \) (\( \exp \) is an analytic diffeomorphism for a nilpotent Lie group); let \( \sigma_\lambda : H \to \mathbb{C}^* \) be defined as

\[
 \sigma_\lambda(\exp X) = e^{2\pi i \lambda(X)}, \tag{53}
\]

for any \( \exp X \in H, X \in h \). This is indeed a unitary representation by the fact that \( \lambda|_{[h, h]} = 0 \) and Baker–Campbell–Hausdorff formula.

Then we can associate with \( \lambda \) the representation \( \text{ind}_H^G(\sigma_\lambda) \) on \( G \). Our construction depends on a choice of maximal subalgebra \( h \subseteq g \) such that \( \lambda \) vanishes on \([h, h]\), but in fact the choice of such \( h \) does not alter the equivalence class of the representation \( \text{ind}_H^G(\sigma_\lambda) \) (see e.g. Theorem 2.2.2 in \([9]\)). Moreover, the maximality property of the subalgebra \( h \) gives the irreducibility of the representation \( \text{ind}_H^G(\sigma_\lambda) \), and for any two elements in \( g^* \) in the same coadjoint orbit the resulting representations are equivalent. Finally, every equivalence class of unitary irreducible representation of \( G \) can be realized by this construction. Summarizing, we have the following:
Theorem 5.1 (see e.g. Theorem 7.9 in [13]). Let $G$ be a simply connected nilpotent Lie group. There is a correspondence between coadjoint orbits of $G$ and unitary irreducible representations of $G$ given by $\mathcal{O}_\lambda \mapsto \text{ind}_H^G(\sigma_\lambda)$.

5.3 The Moment Map on Tangent Bundles of Transformation Models

Theorem 5.2. Let $G$ be a simply connected nilpotent Lie group. Let $X$ be a $G$-homogeneous space with a $G$-invariant measure, such that the stabilizer subgroup $N$ at some point in $X$ is normal, i.e. $X = G/N$. For any $\lambda \in \mathfrak{g}^*$ such that $\lambda|_{\text{Lie}(N)} = 0$, the irreducible unitary representation of $G$ corresponding to the coadjoint orbit $\mathcal{O}_\lambda$ under Kirillov’s correspondence can be realized as a sub-representation of the left-translation representation of $G$ on $L^2(X)$.

Proof. Let $\lambda \in \mathfrak{g}^*$. Denote $\text{Lie}(N) = \mathfrak{n}$. Then $\lambda|_{\mathfrak{n}} = 0$. Let $\mathcal{S}$ be the collection of subalgebras $\mathfrak{h}$ in $\mathfrak{g}$ such that $\lambda|_{[\mathfrak{h}, \mathfrak{h}]} = 0$ and $\mathfrak{n} \subseteq \mathfrak{h}$. $\mathcal{S}$ is nonempty since $\mathfrak{n} \in \mathcal{S}$. Moreover if $\mathfrak{h}_1 \subseteq \mathfrak{h}_2 \subseteq \ldots$ is a chain in $\mathcal{S}$, then $\bigcup \mathfrak{h}_n$ is in $\mathcal{S}$ and is an upper bound of the chain. By Zorn’s lemma $\mathcal{S}$ contains a maximal element $\mathfrak{h}$ with respect to inclusion, which is also a maximal subalgebra in $\mathfrak{g}$ such that $\lambda|_{[\mathfrak{h}, \mathfrak{h}]} = 0$. Therefore we can use $\mathfrak{h}$ and the corresponding Lie subgroup $H = \text{exp}(\mathfrak{h})$ to build an induced representation.

Following the notations in Section 5.2, we denote by $\sigma_\lambda$ the 1-dimensional unitary representation of the subgroup $H$ given by the formula (53). The construction in Section 5.2 requires an invariant measure on $G/H$. In our case, $N$ is a subgroup of $H$, and we have assumed that $G/N$ possesses an invariant measure which we denote by $\mu$. Let $\rho : G/N \to G/H$ be the natural projection map, and $\rho_\ast(\mu)$ the pushforward measure on $G/H$ with the property that for any measurable function $f$ on $G/H$

$$\int_{G/H} f(gH) \, d(\rho_\ast\mu)(gH) = \int_{G/N} f \circ \rho(gN) \, d\mu(gN).$$  (54)

This measure is $G$-invariant, since for any $g \in G$,

$$g_\ast(\rho_\ast\mu) = (g \circ \rho)_\ast\mu = (\rho \circ g)_\ast\mu = \rho_\ast(g_\ast\mu) = \rho_\ast\mu.$$  (55)

Now the Hilbert space $\mathcal{H}$ of the induced representation $\text{ind}_H^G(\sigma_\lambda)$ in our case contains all Borel measurable functions $f : G \to \mathbb{C}$ such that for any $h \in H, g \in G$, $f(gh) = \sigma(h^{-1})f(g)$ and $\int_{G/H} ||f(g)||^2d(\rho_\ast\mu)(gH) < \infty$. Now for any $\exp X \in N$ for $X \in \mathfrak{n}$, since $\lambda(X) = 0$, we know that $\sigma(\exp X) = 1$. Therefore for any $n \in N, g \in G$, $f(gn) = \sigma(n^{-1})f(g) = f(g)$. Therefore for any $f \in \mathcal{H}$, $f$ is constant on each coset of $N$. Thus $f$ induces a Borel measurable function $\tilde{f} : G/N \to \mathbb{C}$. Moreover, by (54), the integrability property for functions in $\mathcal{H}$ is equivalent to $\int_{G/N} ||f(g)||^2 \, d\mu(gN) < \infty$. Therefore $\tilde{f} \in L^2(G/N)$. Therefore there is an embedding of $\mathcal{H}$ to the Hilbert space $L^2(G/N)$, and the unitary representation of $G$ on $\mathcal{H}$ defined by (52) coincides with the natural left-translation representation of $G$ in $L^2(G/N)$. We can then conclude that the representation $\text{ind}_H^G(\sigma_\lambda)$ can be realized as a sub-representation of the left-translation representation of $G$ on $L^2(X) = L^2(G/N)$.  □
Corollary 5.1. Under the same assumptions of $G$ and $X$, let $M$ be an exponential transformation model on $X$, and $TM$ the Kähler manifold that is the tangent bundle of $M$. If there is a map $\mu : TM \to g^*$ such that for any $\lambda \in \text{Im}[\mu] \subseteq g^*$, $\lambda|_{\text{Lie}(N)} = 0$, then the irreducible unitary representation of $G$ corresponding to the coadjoint orbit $O_\lambda$ under Kirillov’s correspondence using the orbit method can be realized as a sub-representation of the left-translation representation of $G$ on $L^2(X)$. Moreover these subrepresentations corresponding to different coadjoint orbits in the image of $\lambda$ form a direct sum in $L^2(X)$.

Proof. The direct sum property is given by the fact that irreducible subrepresentations in $L^2(X)$ have no non-trivial proper subrepresentations. \hfill \square

In particular, suppose the symplectic $G$-action on $TM$ is Hamiltonian with moment map $\mu : TM \to g^*$. Let $\mu^* : g \to C^\infty(TM)$ be the comoment map of $\mu$. If $X \in n = \text{Lie}(N)$, then its fundamental vector field $X^\flat$ on $TM$ is 0, so

$$d\mu^*(X) = \iota_X \omega_{\text{Dom}} = 0 \implies \mu^*(X) \text{ is constant.}$$

If there exists $c \in [g, g]^0 \subseteq g^*$ such that $c|_n = \mu^*|_n$, then the dual map $(\mu^* - c)^* : TM \to g^*$ of $(\mu^* - c)$ is a moment map. Moreover $(\mu^* - c)^*$ satisfies the condition in the theorem: for any $\lambda \in \text{Im}[(\mu^* - c)^*], \lambda|_n = 0$. In other words, there exists a moment map $\tilde{\mu} : TM \to g^*$ of the Hamiltonian $G$-action on $TM$ such that the conditions of the corollary holds. To summarize,

Corollary 5.2. Under the same assumptions of $G, X, M$, if the $G$-action on $TM$ is Hamiltonian with comoment map $\mu^* : g \to C^\infty(TM)$ and $c \in [g, g]^0$ such that $c|_{\text{Lie}(N)} = \mu^*|_{\text{Lie}(N)}$, then there exists a moment map $\tilde{\mu} : TM \to g^*$ of the $G$-action such that

$$\bigoplus_{O_\lambda, \lambda \in \text{Im}[\tilde{\mu}]} \pi_{O_\lambda} \subseteq L^2(X)$$

is a subrepresentation of the unitary $G$-representation in $L^2(X)$, where $\pi_{O_\lambda}$ is a $G$-subrepresentation in $L^2(X)$ equivalent to the unitary irreducible representation of $G$ corresponding to $O_\lambda$ under Kirillov’s correspondence.

Since for any comoment map $\mu^*$, $\mu^*(X)$ is constant, such comoment map is already close to satisfying the condition that any $\lambda \in \text{Im}[\mu]$ satisfies $\lambda|_{\text{Lie}(N)} = 0$. Therefore other alternative conditions for the Hamiltonian $G$-action can be required to arrive at the same conclusion.

References

[1] Shun-ichi Amari and Hiroshi Nagaoka. Methods of information geometry, volume 191. American Mathematical Soc., 2007.
[2] Vladimir Igorevich Arnol’d. *Mathematical methods of classical mechanics*, volume 60. Springer Science & Business Media, 2013.

[3] Abhay Ashtekar and Troy A Schilling. Geometrical formulation of quantum mechanics. In *On Einstein’s Path*, pages 23–65. Springer, 1999.

[4] Nihat Ay, Jürgen Jost, Hồng Vân Lê, and Lorenz Schwachhöfer. *Information geometry*, volume 64. Springer, 2017.

[5] Ole E Barndorff-Nielsen. *Parametric statistical models and likelihood*, volume 50. Springer Science & Business Media, 2012.

[6] Ole E Barndorff-Nielsen, Preben Blæsild, and Poul S Eriksen. *Decomposition and invariance of measures, and statistical transformation models*, volume 58. Springer Science & Business Media, 2012.

[7] Ovidiu Calin and Constantin Udrişte. *Geometric modeling in probability and statistics*. Springer, 2014.

[8] Manfredo Perdigao do Carmo. *Riemannian geometry*. Birkhäuser, 1992.

[9] Laurence Corwin and Frederick P Greenleaf. *Representations of nilpotent Lie groups and their applications: Volume 1, Part 1, Basic theory and examples*, volume 18. Cambridge university press, 2004.

[10] Ana Cannas Da Silva. *Lectures on symplectic geometry*. Springer, 2008.

[11] Peter Dombrowski. On the geometry of the tangent bundle. *Journal für Mathematik. Bd*, 210(1/2):10, 1962.

[12] Marco AN Fernandes and Luiz AB San Martin. Fisher information and $\alpha$-connections for a class of transformational models. *Differential Geometry and its Applications*, 12(2):165–184, 2000.

[13] Gerald B Folland. *A course in abstract harmonic analysis*, volume 29. CRC press, 2016.

[14] Janusz Grabowski, Marek Kuś, and Giuseppe Marmo. Segre maps and entanglement for multipartite systems of indistinguishable particles. *Journal of Physics A: Mathematical and Theoretical*, 45(10):105301, 2012.

[15] Brian C Hall. *Quantum theory for mathematicians*, volume 267. Springer, 2013.

[16] Aleksandr Aleksandrovich Kirillov. *Lectures on the orbit method*, volume 64. American Mathematical Soc., 2004.

[17] John Lee. *Introduction to topological manifolds*, volume 202. Springer Science & Business Media, 2010.
[18] Chernoff Marsden. *Properties of infinite dimensional Hamiltonian systems.* Springer, 1974.

[19] Mathieu Molitor. Remarks on the statistical origin of the geometrical formulation of quantum mechanics. *International Journal of Geometric Methods in Modern Physics*, 9(03):1220001, 2012.

[20] Mathieu Molitor. Exponential families, kähler geometry and quantum mechanics. *Journal of Geometry and Physics*, 70:54–80, 2013.

[21] Andrei Moroianu. *Lectures on Kähler geometry*, volume 69. Cambridge University Press, 2007.

[22] James Munkres. *Topology.* Pearson Education, 2014.

[23] Walter A Poor. *Differential geometric structures.* Courier Corporation, 2007.