4 Path Equivalent Network

In this section we will ...

4.1 Manifold-path

Definition 4.1 (Lie group)**.** *[10] A Lie group is a smooth manifold G (without boundary) that is also a group in the algebraic sense, with the property that the multiplication map $m : G \times G \to G$ and inversion map $i : G \to G$, given by*

$$m(g, h) = gh, \quad i(g) = g^{-1}$$

are both smooth. A Lie group is, in particular, a topological group (a topological space with a group structure such that the multiplication and inversion maps are continuous).

Definition 4.2 (Identity component)**.** *If G is a Lie group, the connected component of G containing the identity is called the identity component of G, denoted as G^0.*

Definition 4.3 (G-space and G-manifold)**.** *Let G be a Lie group.*

A G-space, denoted as (X, ρ_X), is a topology space X together with a continuous action

$$\rho_X : G \times X \to X$$

satisfying $e \cdot x = x$ and $(g_1 g_2) \cdot x = g_1 \cdot (g_2 \cdot x)$.

Similarly, a G-manifold, denoted as (M, ρ_M), is a manifold M together with a smooth action

$$\rho_M : G \times M \to M$$

satisfying $e \cdot m = m$ and $(g_1 g_2) \cdot m = g_1 \cdot (g_2 \cdot m)$.

In general, we denote the action $\rho : G \times X \to X$ as $G \curvearrowright X$.

Definition 4.4 (G-equivariant map and G-equivariant learner)**.** *Given a G-space (X, ρ_X) and a G-manifold (M, ρ_M), a map $f : X \to M$ is G-equivariant map (G-map) if*

$$f(g \cdot x) = g \cdot f(x) \quad \forall g \in G, x \in X.$$

Equivalently, this diagram

$$\begin{tikzcd}
G \times X \arrow[r, "\rho_X"] \arrow[d, "id_G \times f"'] & X \arrow[d, "f"] \\
G \times M \arrow[r, "\rho_M"'] & M
\end{tikzcd}$$

commutes.

We call an G-map machine learning algorithm an G-equivariant learner.

Definition 4.5 (Manifold-path)**.** *Let G be a Lie group, X be a topological G-space, Z be a G-manifold, $E : X \to Z$ is a G-equivalent learner* [1]*. Fix $x_0 \in X$ and set $z_0 := E(x_0) \in Z$. For any continuous path $\gamma_{g_0} : [0,1] \to G$ with $\gamma_{g_0}(0) = e$, define*

$$\gamma_{z_0}(t) := \gamma_{g_0}(t) \cdot z_0 = E\left(\gamma_{g_0}(t) \cdot x_0\right) \quad t \in [0,1].$$

We call γ_{z_0} a manifold-path in Z.

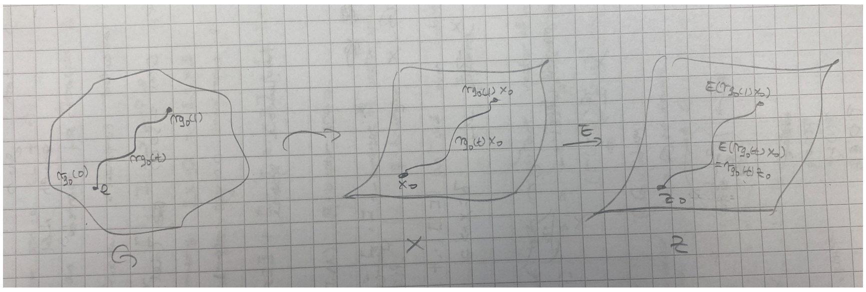

Figure 1: manifold path

Lemma 4.6. *Manifold-paths are topological paths.*

Proof. The map $\gamma_{z_0} : [0,1] \to Z$ is continuous, hence every manifold-path is an ordinary topological path. □

Proposition 4.7 (Orbit reachability)**.** *Let G^0 denote the identity component of G. For any $z_0 \in Z$,*

$$\left\{\gamma_{z_0}(1)\right\} = G^0 \cdot z_0.$$

In particular, the orbit $G^0 \cdot z_0$ is path-connected in the topological sense.

Proof. Step 1: $\left\{\gamma_{z_0}(1)\right\} \subseteq G^0 \cdot z_0$.

Let γ_{z_0} be any manifold-path. By definition:

$$\gamma_{z_0}(t) = \gamma_{g_0}(t) \cdot z_0$$

where $\gamma_{g_0} : [0,1] \to G$ is continuous with $\gamma_{g_0}(0) = e$. At the endpoint $t = 1$:

$$\gamma_{z_0}(1) = \gamma_{g_0}(1) \cdot z_0$$

Since γ_{g_0} is a continuous path connecting $e = \gamma_{g_0}(0)$ to $\gamma_{g_0}(1)$, both endpoints lie in the same connected component of G.

Since $e \in G^0$ by definition, we have:

$$\gamma_{g_0}(1) \in G^0$$

[1]In manifold learning, X is input data, the manifold is latent space, a G-equivalent learner is a encoder.

Therefore:

$$\gamma_{z_0}(1) = \gamma_{g_0}(1) \cdot z_0 \in G^0 \cdot z_0$$

This proves $\left\{\gamma_{z_0}(1)\right\} \subseteq G^0 \cdot z_0$

Step 2: $G^0 \cdot z_0 \subseteq \left\{\gamma_{z_0}(1)\right\}$

Proof:Let $g \in G^0$ be arbitrary. We need to show: $g \cdot z_0$ can be written as $\gamma_{z_0}(1)$ for some manifold-path.

Since $g \in G^0$ and G^0 is path-connected, there exists a continuous path:

$$\gamma_{g_0} : [0, 1] \to G^0 \subseteq G$$

such that $\gamma_{g_0}(0) = e$ and $\gamma_{g_0}(1) = g$.

By definition

$$\gamma_{z_0}(t) := \gamma_{g_0}(t) \cdot z_0 = E\left(\gamma_{g_0}(t) \cdot x_0\right)$$

Verify

$$\gamma_{z_0}(1) = \gamma_{g_0}(1) \cdot z_0 = g \cdot z_0$$

Therefore $g \cdot z_0 \in \left\{\gamma_{z_0}(1)\right\}$.

This proves $G^0 \cdot z_0 \subseteq \left\{\gamma_{z_0}(1)\right\}$

Claim: The orbit $G^0 \cdot z_0$ is path-connected.

Proof:

The orbit $G^0 \cdot z_0$ is the continuous image of G^0 under the map:

$$\begin{aligned} \phi : G^0 &\to Z \\ g &\mapsto g \cdot z_0 \end{aligned}$$

Since ϕ is continuous and G^0 is path-connected, by the theorem "continuous image of path-connected space is path-connected", the image $\phi\left(G^0\right) = G^0 \cdot z_0$ is path-connected.

$\square$

This shows us that the set of points we can reach by manifold-paths is exactly the orbit under the identity component, which is path-connected.

4.2 Path-equivalent (PE) map

In this section we formalize path-equivariant (PE) maps and path-equivariant networks (PENs).

4.2.1 Define PE

Definition 4.8 (Reparametrization and concatenation of paths)**.** *If γ is a path and $\phi : [0,1] \to [0,1]$ is continuous with $\phi(0) = 0$ and $\phi(1) = 1$, then the reparametrized path is $\gamma \circ \phi$.*

If γ_1, γ_2 is two paths with $\gamma_1(1) = \gamma_2(0)$ (end of first meets start of second), then the concatenation [2] *of them is:*

$$\gamma_1 \| \gamma_2 \in \mathcal{P}$$

where

$$(\gamma_1 \| \gamma_2)(t) = \begin{cases} \gamma_1(2t) & t \in [0, 1/2] \\ \gamma_2(2t-1) & t \in [1/2, 1] \end{cases}$$

Definition 4.9 (Path system)**.** *Let X be a topological space. A path system on X is a non-empty family $\mathcal{P}$ of continuous curves $\gamma : [0,1] \to X$ that is closed under reparametrization and concatenation.*

For example group paths:

$$\mathcal{P}_{\text{group}} = \{\gamma : [0,1] \to X \mid \gamma(t) = g(t) \cdot x \text{ for some } x \in X, g : [0,1] \to G \text{ continuous } \}$$

is a path system.

Definition 4.10 (Path-equivariance (PE) map)**.** *Let X be a topological space with path system $\mathcal{P}$, A be a Lie group, Z be a manifold. Consider a (usually not discrete) Lie group action $A \curvearrowright Z$. A continuous map $F : X \to Z$ is path-equivariant with respect to $\mathcal{P}$ if for every $\gamma \in \mathcal{P}$, there exists a continuous transport $a_\gamma : [0,1] \to A$ with $a_\gamma(0) = e$ such that*

$$F(\gamma(t)) = a_\gamma(t) \cdot F(\gamma(0)) \quad \forall t \in [0,1].$$

A neural network with PE property is called Path-equivariant network(PEN).

Compared with G-equivalent learner we defined before, F is more general. We introduce a new Lie group structure A to Z, and not use G structure on X. This is because A is not the same with G in general, or even, X not necessary equipped with group structure, which gives us more flexibility. For example, if input space X contains images with $G = SE(2)$, the output space Z contains feature vectors with $A = SO(2)$, in this case F is equivariant to rotations but invariant to translations, which is natural that object recognition should care about orientation but not position.

$F : X \to Z$ is a point-wise mapping, but it gives a induced path mapping.

[2]Note the order of concatenation, $\gamma_1 \| \gamma_2$ means first apply γ_1, then γ_2

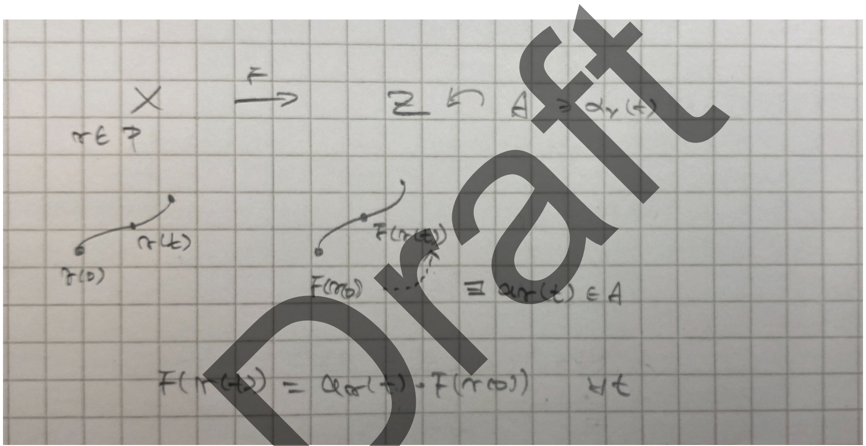

Figure 2: PE-map

Definition 4.11 (Induced path)**.** *Given path $\gamma : [0,1] \to X$, define the induced path:*

$$(F_*)(\gamma) := F \circ \gamma : [0,1] \to Z$$
$$(F_*)(\gamma)(t) = F(\gamma(t))$$

where F_* maps paths in X to paths in Z.

Proposition 4.12 (Manifold-Paths Satisfy Path-Equivariance)**.** *Let G be a Lie group, X be a topological G-space, and Z be a G-manifold. Let $E : X \to Z$ be a G-equivariant map, i.e.,*

$$E(g \cdot x) = g \cdot E(x) \quad \forall g \in G, x \in X.$$

Define the path system:

$$\begin{aligned} \mathcal{P} := \{\gamma_X : [0,1] \to X \,|\gamma_X(t) = \gamma_{g_0}(t) \cdot x \text{ for some } x \in X, \\ \gamma_{g_0} : [0,1] \to G \text{ continuous with } \gamma_{g_0}(0) = e\} \end{aligned}$$

That is, $\mathcal{P}$ consists of all group paths in X.

The map $E : X \to Z$ is path-equivariant with respect to $\mathcal{P}$ by taking $A = G$ as the acting group on Z.

Proof. Goal: For every $\gamma_X \in \mathcal{P}$, there exists a continuous transport $a_{\gamma_X} : [0,1] \to G$ with $a_{\gamma_X}(0) = e$ such that:

$$E(\gamma_X(t)) = a_{\gamma_X}(t) \cdot E(\gamma_X(0)) \quad \forall t \in [0,1].$$

Let $\gamma_X \in \mathcal{P}$ be arbitrary. By definition of $\mathcal{P}$, there exist:
- A point $x_0 \in X$
- A continuous path $\gamma_{g_0} : [0,1] \to G$ with $\gamma_{g_0}(0) = e$
such that:

$$\gamma_X(t) = \gamma_{g_0}(t) \cdot x_0 \quad \forall t \in [0,1].$$

Define the transport:

$$a_{\gamma_X}(t) := \gamma_{g_0}(t) \in G.$$

Verify properties:
1. Continuity: $a_{\gamma_X} : [0,1] \to G$ is continuous since γ_{g_0} is continuous.
2. Initial condition: $a_{\gamma_X}(0) = \gamma_{g_0}(0) = e$
We need to verify the PE condition:

$$E(\gamma_X(t)) = a_{\gamma_X}(t) \cdot E(\gamma_X(0)) \quad \forall t \in [0,1].$$

Compute the left-hand side:

$$E(\gamma_X(t)) = E\left(\gamma_{g_0}(t) \cdot x_0\right)$$

Since E is G-equivariant, we have:

$$E\left(\gamma_{g_0}(t)\cdot x_0\right) = \gamma_{g_0}(t)\cdot E\left(x_0\right)$$

Compute the right-hand side:

$$a_{\gamma_X}(t)\cdot E\left(\gamma_X(0)\right) = \gamma_{g_0}(t)\cdot E\left(\gamma_{g_0}(0)\cdot x_0\right)$$

Using $\gamma_{g_0}(0) = e$:

$$= \gamma_{g_0}(t)\cdot E\left(e\cdot x_0\right) = \gamma_{g_0}(t)\cdot E\left(x_0\right)$$

Conclusion:

$$E\left(\gamma_X(t)\right) = \gamma_{g_0}(t)\cdot E\left(x_0\right) = a_{\gamma_X}(t)\cdot E\left(\gamma_X(0)\right)$$

This holds for all $t \in [0,1]$.

Since $\gamma_X \in \mathcal{P}$ was arbitrary, we have shown that for every path in $\mathcal{P}$, there exists a continuous transport satisfying the PE condition.

Therefore, E is path-equivariant with respect to $\mathcal{P}$.

□

Corollary 4.13 (Manifold-Paths are PE Paths)**.** *For any fixed $x_0 \in X$, let $z_0 := E\left(x_0\right)$. Then for any continuous group path $\gamma_{g_0} : [0,1] \to G$ with $\gamma_{g_0}(0) = e$, the manifold-path:*

$$\gamma_{z_0}(t) := E\left(\gamma_{g_0}(t)\cdot x_0\right)$$

satisfies:

$$\gamma_{z_0}(t) = \gamma_{g_0}(t)\cdot z_0 = \gamma_{g_0}(t)\cdot\gamma_{z_0}(0).$$

Proof. This follows directly from the proof above by setting $\gamma_X(t) = \gamma_{g_0}(t)\cdot x_0$ and noting that:

$$\gamma_{z_0}(t) = E(\gamma_X(t)) = \gamma_{g_0}(t)\cdot E(x_0) = \gamma_{g_0}(t)\cdot z_0.$$

□

4.2.2 From PEN to G-NN

To recover classical group-equivariant neural networks (G-NN) from path-equivariance (PEN), we must preserve the group structure from X to Z. In the input space X, group elements $g \in G$ act on points. We can represent each group element g as the endpoint of a path starting from the identity:

$$g = \gamma(1)$$

where $\gamma : [0,1] \to G$ with $\gamma(0) = e$

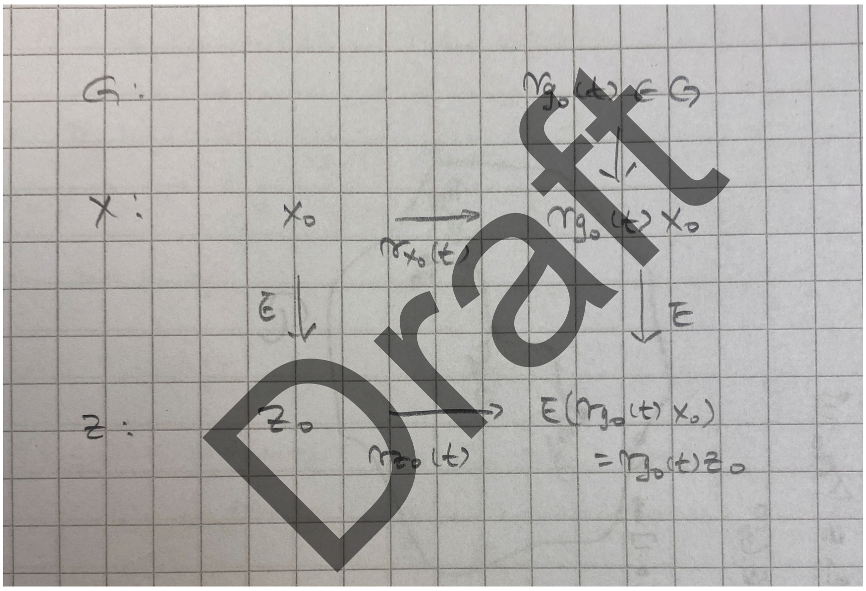

Figure 3: PE-diagram

A path-equivariant map $F : X \to Z$ induces a mapping on paths:

$$F_* : \gamma \mapsto a_\gamma$$

where γ is a path in X (induced by group path in G), and a_γ is the corresponding transport path in A.

For F to preserve the group structure, elements in Z must be determined by the corresponding group elements in X, not by the specific path taken.

If two paths in G have the same endpoint:

$$\gamma_1(1) = \gamma_2(1) = g$$

then their induced transports must also have the same endpoint:

$$a_{\gamma_1}(1) = a_{\gamma_2}(1)$$

This is precisely the no-holonomy condition.

Definition 4.14 (No-holonomy condition)**.** *Two group paths $\gamma_1, \gamma_2 : [0,1] \to G^0$ with the same endpoints ($\gamma_1(0) = \gamma_2(0) = e$ and $\gamma_1(1) = \gamma_2(1) = g$) satisfy the no-holonomy condition if their induced transports coincide: $a_{\gamma_1}(1) = a_{\gamma_2}(1)$.*

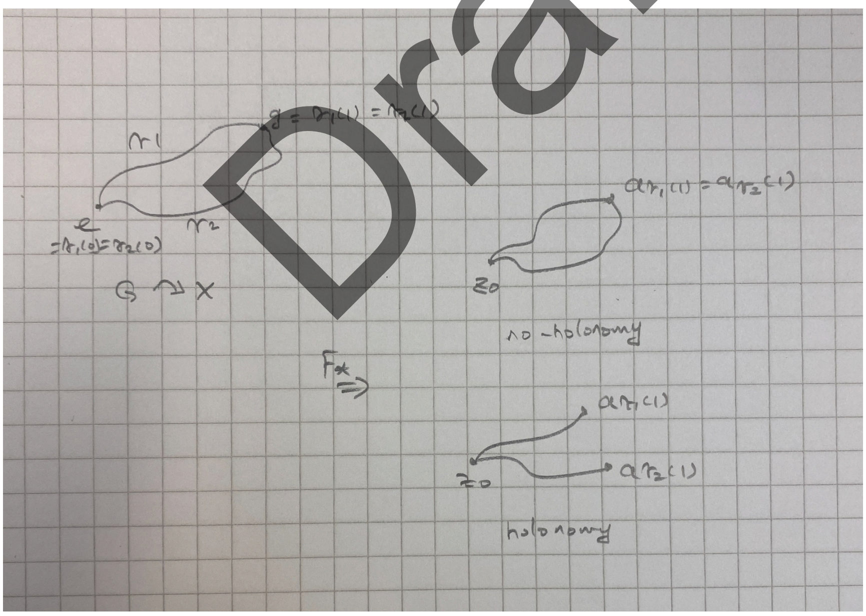

Figure 4: holonomy

Definition 4.15 (Endpoint map)**.** *Under the no-holonomy condition, we can define the endpoint map $\rho : G^0 \to A$ by*

$$\rho(g) := a_\gamma(1)$$

where γ is any path in G^0 from e to g, G^0 is identity component.

The no-holonomy condition ensures ρ is independent of the choice of path from e to g, all paths with same endpoints give the same value. So ρ is well-defined.

Proposition 4.16. *ρ is a group homomorphism.*

Proof. **Step 1**: Prove $\rho(e) = e$.

Take the constant path $\gamma(t) = e$ from e to e. By definition, $\rho(e) = a_\gamma(1) = e$

Step 2:

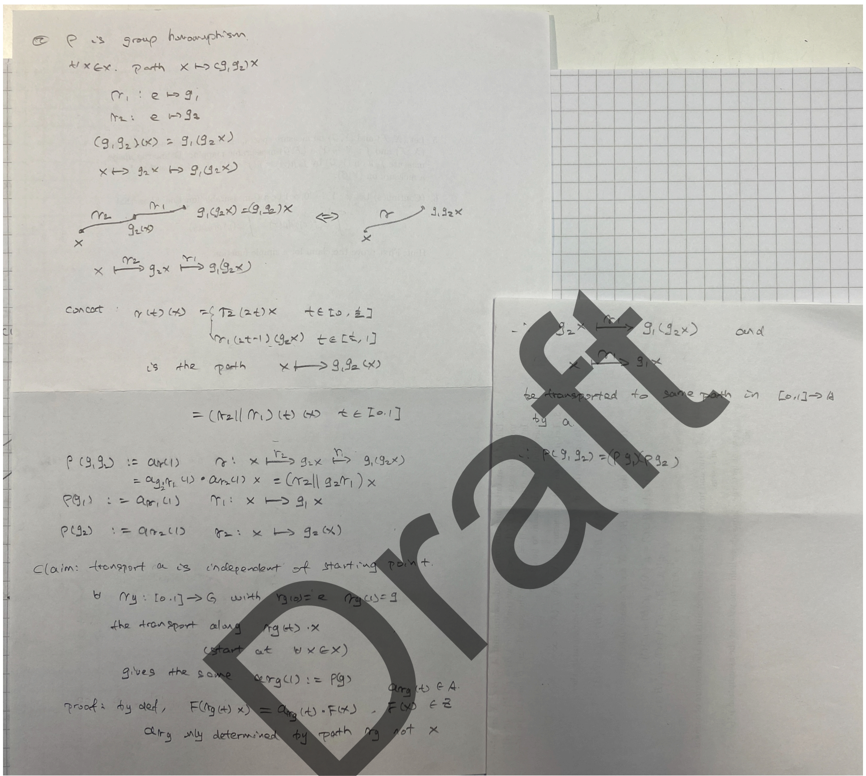

Figure 5: prove-homo

Step 3: Prove $\rho(g^{-1}) = \rho(g)^{-1}$

If γ is a path from e to g, then $\gamma^{-1}(t) := \gamma(1-t)^{-1}$ is a path from e to g^{-1}.

Transport reverses: $a_{\gamma^{-1}}(1) = a_\gamma(1)^{-1}$. Therefore $\rho(g^{-1}) = \rho(g)^{-1}$

□

Proposition 4.17 (Reduction to group equivariance on G^0)**.** *Let Let $G \curvearrowright X$ and $A \curvearrowright Z$, $F : X \to Z$ be path-equivariant with respect to group paths $\mathcal{P} = \{g(t) \cdot x : g(0) = e\}$.*

Assume the no-holonomy condition: For all paths γ_1, γ_2 from e to g:

$$a_{\gamma_1}(1) = a_{\gamma_2}(1)$$

Then:

$$F(g \cdot x) = \rho(g) \cdot F(x) \quad \forall g \in G^0, x \in X$$

where $\rho : G^0 \to A$ is the endpoint map.

This is exactly the classical group-equivariance law on G^0.

Proof. Fix $g \in G^0$ and $x \in X$. Since G^0 is path-connected, there exists a continuous path $\gamma : [0,1] \to G^0$ with $\gamma(0) = e$ and $\gamma(1) = g$.

Consider the group path in X:

$$\gamma_X(t) := \gamma(t) \cdot x$$

By definition: $\gamma_X(0) = e \cdot x = x$ and $\gamma_X(1) = g \cdot x$.

Since F is path-equivariant, there exists transport $a_\gamma : [0,1] \to A$ with $a_\gamma(0) = e$ such that:

$$F(\gamma_X(t)) = a_\gamma(t) \cdot F(\gamma_X(0))$$

At $t = 1$:

$$F(g \cdot x) = F(\gamma_X(1)) = a_\gamma(1) \cdot F(x)$$

By the no-holonomy condition and definition of ρ:

$$a_\gamma(1) = \rho(g)$$

Therefore:

$$F(g \cdot x) = \rho(g) \cdot F(x)$$

This holds for all $g \in G^0$ and $x \in X$.

□

This proposition means that if F is path-equivariant and satisfies no-holonomy, then:

- The transport map between two symmetry $\rho : G^0 \to A$ is a continuous homomorphism, so the group structure is preserved : $g_1 \in G^0 \xrightarrow{\rho} \rho(g_1) \in A$

- F becomes a classical group-equivariant map: For any $g \in G^0$ and $x \in X$: $F(g \cdot x) = \rho(g) \cdot F(x)$

From a geometric viewpoint, when G^0 is simply-connected, any two paths with the same endpoints are homotopic. Additionally, if the transport is homotopy-invariant (meaning homotopic paths induce the same transport), then these paths automatically satisfy the no-holonomy condition.

Based on this geometric insight, we can define an intermediate notion between path equivariance and classical group equivariance.

Definition 4.18 (Homotopy-equivariant network)**.** *A homotopy-equivariant network satisfies:*

$$\gamma_1 \simeq \gamma_2 \Longrightarrow a_{\gamma_1}(1) = a_{\gamma_2}(1)$$

In homotopy-equivariant network, paths in the same homotopy class induce the same output transformation

When G^0 is not simply-connected, the relationship becomes more nuanced, paths can still satisfy the no-holonomy condition. However, paths to the same endpoint are not necessarily homotopic, different homotopy paths labeled by the fundamental group $\pi_1\left(G^0\right)$. In this case, the homotopy classes represent the number of topologically distinct ways to reach g from the identity.

For $SO(3)$, we have $\pi_1(SO(3)) = \mathbb{Z}_2$, yielding two homotopy classes:

- Class 0 (untwisted paths): All paths give transport $\rho_0(R)$, for example direct rotation to angle R

- Class 1 (twisted paths): All paths give transport $\rho_1(R)$, for example rotating 360° plus reaching R

Where ρ_0, ρ_1 may differ.

The network can distinguish whether a rotation was reached via a topologically twisted path or not. This additional structure makes homotopy-equivariant networks more expressive than classical G-NNs, which would require $\rho_0(R) = \rho_1(R)$.

The relationship between these frameworks forms a hierarchy:

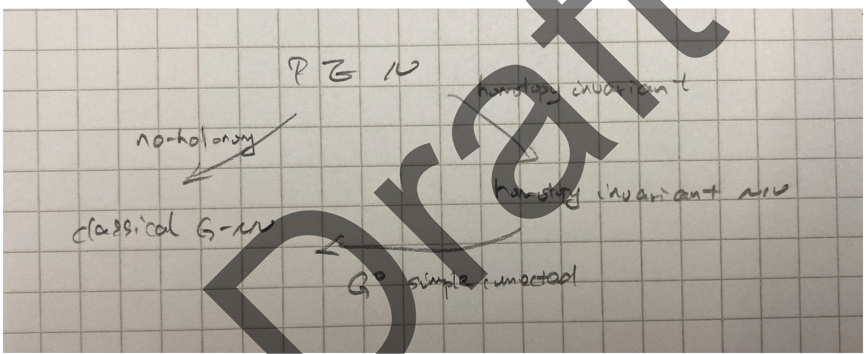

Figure 6: hierarchy

Remark 4.19 (Connectivity constraint)**.** *If Lie group G is disconnected, i.e. it can be written as:*

$$G = G^0 \sqcup G_1 \sqcup G_2 \sqcup \cdots$$

where G^0 is the identity component, $G_1, G_2, \ldots$ are other connected components, $\sqcup$ means disjoint union, we cannot reach other components from e (or G^0) by continuous paths. In this case we might need to specify discrete component group $\pi_0(G) = G/G^0$. Or, we can define content space $U = X/G$ which is not a group. We will discuss more about this idea in next chapter.

4.2.3 From Homotopy to Homology

While homotopy studies continuous deformations of paths (captured by the fundamental group π_1), homology provides a more algebraic approach to topology, measuring holes and voids of all dimensions through homology groups $H_k(X)$. This distinction has led to the development of Topological Data Analysis (TDA), a rapidly advancing field that leverages homology theory for data analysis. Several neural network architectures have been developed to preserve or utilize topological features: Topology-Preserving Autoencoders[cite], Persistence-Based Layers [cite], Topological Regularization [cite].

Intuitively (or counterintuitively), these frameworks capture different aspects of data structure: G-NN captures how data behaves under transformations, TDA captures how data looks intrinsically. However, we know that homology is also a group $H_k(X)$ which is discrete and abelian, and in previous section we argued that regularization can be thought as a soft group constraint via loss term. We conjecture that deeper connections exist between G-NN and TDA.

4.3 Content Space and Pose Space

Definition 4.20 (Content Space). *Let a Lie group G act smoothly on data manifold X. Define the content space is the quotient (orbit) space*

$$U := X/G.$$

We call G a "pose" on content U.

Remark 4.21. *U is not a group in general.*

The content space U is the collection of all orbits, each point $u \in U$ represents an entire orbit $O_x = \{g \cdot x : g \in G\}$ in X. Intuitive thinking, content space is invariant to transformations, it captures the intrinsic identity, means "what the data is". For example in face recognition tasks, data X is all face images with various rotations, $G = SO(2)$ acts by rotating images, U means face identities corresponds to different person, each $u \in U$ represents one person's face.

Definition 4.22 (Projection Map in Content Space). *Define the projection map*

$$\begin{aligned} \pi : X &\to U \\ x &\mapsto [x] \end{aligned}$$

where $[x]$ means the orbits containing x.

For example in face recognition tasks, $\pi(x)$ means one specific person x, regardless one's position.

Proposition 4.23 (Meinrenken's notes Theorem 1.21.). *If the action G is free and proper, the projection $\pi : X \to U$ is a submersion, X is a principal G-bundle over U, the fibers are the orbits.*

In this case, X is total space, $U = X/G$ is base space, $\pi : X \to U$ is the projection sends x to its orbits, $\pi^{-1}(u) \subset X$ is the fiber over u, $c : U \to X$ is a section picks one point from each fiber. If we stack all orbits together we reconstruct $X = \bigsqcup_{u \in U} \pi^{-1}(u)$.

According to principal bundle properties:

- Each fiber is a copy of G

$$\pi^{-1}(u) \cong G$$

- Around each $u \in U$, there exists a neighborhood U_α and a diffeomorphism:

$$\phi : \pi^{-1}(U_\alpha) \cong U_\alpha \times G$$

- U_α with a homeomorphism $\varphi : U_\alpha \to V \subset \mathbb{R}^n$ yield a chart (U_α, φ).

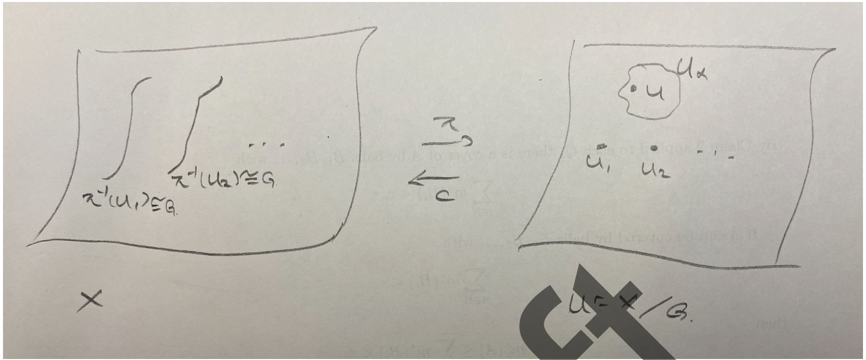

Figure 7: fiber

Definition 4.24 (Canonical Pose Map)**.** *A (local) canonical pose is a smooth section $c : U_\alpha \to X$ of π on a chart U_α.*

For each $u \in U_\alpha$, the section chooses one specific representative from the fiber $\pi^{-1}(u)$. This choice usually not unique, it serves as the reference for that chart, together with pose action g, we can represent $x \in X$.

Proposition 4.25. *Every sample $x \in X$ can be unique decomposed as*

$$x = g \cdot c(u), \quad u \in U_\alpha, g \in G.$$

The data point x can be obtained by applying transformation g to the canonical pose $c(u)$.

Proof. For any $x \in X$,

Existence:

$u = \pi(x)$ is uniquely determined by orbit of x, $c(u)$ is determined by the section c, G acts transitively on the fiber $\pi^{-1}(u)$, so $\exists g$ such that $x = g \cdot c(u)$.

Uniqueness:

Action G is free, so if $g_1 \cdot c(u) = g_2 \cdot c(u)$, then $g_1 = g_2$. Therefore the pose g is uniquely determined

□

Remark 4.26. *The canonical pose c is locally defined on a chart U_α , not all of U , because the global section $c : U \to X$ may not exist.*

Remark 4.27. *Better to give some real examples.*

Lemma 4.28 (Union of orbits are path-connected)**.** *Let G^0 be the identity component of G. Fix a chart $U_\alpha \subset U$ with section c and let*

$$C_\alpha := \left\{g \cdot c(u) : u \in U_\alpha, g \in G^0\right\} = \bigcup_{u \in U_\alpha} G^0 \cdot c(u).$$

If U_α is path-connected, then C_α is path-connected.

Proof. Consider $F : U_\alpha \times G^0 \to X, F(u, g) = g \cdot c(u)$. The action $(g, x) \mapsto g \cdot x$ is continuous, and c is a section thus continuous. So F is continuous.

U_α is path-connected by hypothesis, G^0 is path-connected by definition. So $U_\alpha \times G^0$ is path-connected.

Images of path-connected sets are path-connected under continuous maps; hence

$$C_\alpha = F\left(U_\alpha \times G^0\right)$$

is path-connected. □

Intuitively, pose means “where", it moves within an orbit, this movement can be represented by group-equivariant layers $F(gx) = \rho(g)F(x)$ which is exactly G-CNNs. Content means “what", it moves across orbits, handle with smoothness across U, no group law to enforce.

Definition 4.29 (Pose Path and Content Path)**.** *Pose path (within orbit) is defined as:*

$$\gamma_{X,G}(t) = \gamma_G(t) \cdot c(u) \quad t \in [0, 1]$$

where $\gamma_G : [0, 1] \to G$ is a continuous path in the group with $\gamma_G(0) = e$, $u \in U_\alpha$ is fixed with same content, $c(u)$ is the canonical pose (fixed point in X).

Content Paths (Between Orbits) is defined as:

$$\gamma_{X,U}(t) = c(\gamma_U(t)) \quad t \in [0, 1]$$

where $\gamma_U : [0, 1] \to U_\alpha$ is a continuous path in content space, pose is fixed at canonical pose $c(\cdot)$.

Geometric thinking, pose paths start at canonical pose $\gamma_{X,G}(0) = e \cdot c(u) = c(u)$, and stay in the same orbit $\pi^{-1}(u)$, explore different poses of the same content; content paths start at $\gamma_{X,U}(0) = c(\gamma_U(0))$, cross different orbits, and always pick the canonical representative from each orbit.

Define the combined path system:

$$\mathcal{P} = \left\{\gamma_{X,G}(t) : u \in U_\alpha, \gamma_G(0) = e\right\} \cup \left\{\gamma_{X,U}(t) : \gamma_U \text{ continuous in } U_\alpha\right\}.$$

Recall the PE Requirement: For $F : X \to Z$ with $A \curvearrowright Z$, and every $\gamma \in \mathcal{P}$:

$$\exists a_\gamma : [0, 1] \to A \text{ with } a_\gamma(0) = e$$

such that

$$F(\gamma(t)) = a_\gamma(t) \cdot F(\gamma(0)).$$

In pose/content scenario, we geometrically abstract X into $U = X/G$ space, but this space is generally not equipped with a natural group action (like $A \curvearrowright Z$ in PE). In PE, Z is a new space where we want to encode the data, the group A that acts on Z is part of our design choice, ideally, Z should “represent" U in some sense.

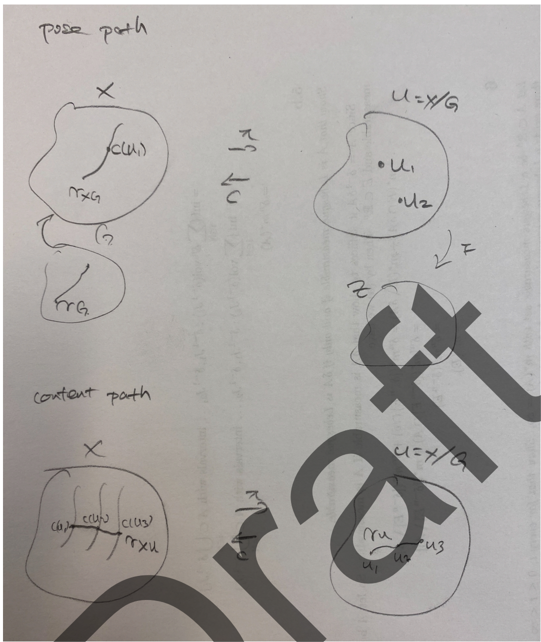

Figure 8: pose-content-path

4.3.1 PE on Pose space

Case 1: Pose-Invariant Encoder

Let $F : X \to Z$ where Z is a feature space representing content, we want F to be constant on orbits (pose-invariant).

For pose paths:

$$\gamma_{X,G}(t) = \gamma_G(t) \cdot c(u)$$

F behaves constantly on pose path $\gamma_G(t)$:

$$F\left(\gamma_{X,G}(t)\right) = F(c(u)) = \text{constant}$$

In PE condition, this is equivalent to choose $A = \{e\}$ and transport $a_{\gamma_{X,G}}(t) = e$ for all t:

$$F\left(\gamma_{X,G}(t)\right) = e \cdot F\left(\gamma_{X,G}(0)\right) = F(c(u))$$

Case 2: Pose-Preserving Encoder

For Pose Paths:

$$\gamma_{X,G}(t) = \gamma_G(t) \cdot c(u)$$

F behaves:

$$F\left(\gamma_{X,G}(t)\right) = \rho\left(\gamma_G(t)\right) \cdot F(c(u))$$

where $\rho : G \to A$ is a homomorphism.

This requires us to choose some A that can represent behaviors of G, for example same group $A = G$, then we can design $A \curvearrowright Z$ and transport $a_{\gamma_{X,G}}(t) = \rho\left(\gamma_G(t)\right) \in A$, the PE condition in this case is

$$F\left(\gamma_G(t) \cdot c(u)\right) = \rho\left(\gamma_G(t)\right) \cdot F(c(u))$$

Proposition 4.30 (Reduce to classical pose equivariance)**.** *Assume no-holonomy condition holds for pose paths: for all $\gamma_{G,1}, \gamma_{G,2} : [0,1] \to G$ with $\gamma_{G,1}(0) = \gamma_{G,2}(0) = e$ and $\gamma_{G,1}(1) = \gamma_{G,2}(1) = g$:*

$$a_{\gamma_{X,G1}}(1) = a_{\gamma_{X,G,2}}(1)$$

Then there exists a continuous homomorphism $\rho : G^0 \to A$ such that:

$$F(g \cdot c(u)) = \rho(g) \cdot F(c(u)) \quad \forall g \in G^0, \forall u \in U_\alpha$$

which gives us classical pose equivariance. Moreover, $a_{\gamma_{X,G}}(t) = \rho\left(\gamma_G(t)\right)$ for any pose path.

4.3.2 PE on Content space

For content paths:

$$\gamma_{X,U}(t) = c\left(\gamma_U(t)\right), \quad t \in [0,1]$$

But U is not a group in general, so there is no homomorphism $U \to A$.

Case 1: Content-Invariant Encoder

If F is constant across all content:

$$F\left(c\left(\gamma_U(t)\right)\right) = F\left(c\left(\gamma_U(0)\right)\right) = \text{constant}$$

In PE, we choose $A = \{e\}$ and $a_{\gamma_{X,U}}(t) = e$. This is rarely useful beceause it loses all content information.

Case 2: Content-Consistent Encoder

If F varies consistently as content changes, for content path $\gamma_{X,U}(t) = c\left(\gamma_U(t)\right)$, PE condition:

$$F\left(c\left(\gamma_U(t)\right)\right) = a_{\gamma_{X,U}}(t) \cdot F\left(c\left(\gamma_U(0)\right)\right)$$

where $a_{\gamma_{X,U}} : [0,1] \to A$ with $a_{\gamma_{X,U}}(0) = e$.

Proposition 4.31 (Holonomy)**.** *In content space, no homomorphism from U to A, so the transportation $a_{\gamma_{X,U}}(t)$ can depend on the specific path taken from u_0 to u_1, not just endpoints. In other words, the paths matter in content space, different paths with same endpoint remain different paths in latent space, which makes sense because the content identify elements we want to recognise.*

4.3.3 Soft and Hard constraint

Group Structure (Pose) provides an equivalence relation: All points in an orbit are equivalent. This equivalence allows weight sharing across group transformations and hard architectural constraints enforce equivariance exactly, for example G-CNN and steerable filters. Manifold Structure (Content) provides smoothness but no equivalence: Different contents are fundamentally distinct, we cannot use weight sharing because $c(u_1) \not\equiv c(u_2)$. Instead, smooth variation behaves like a soft constraints. Here we show that regularization can be applied to content (soft) constraint.

If PE holds exactly:

$$F(c(\gamma_U(t))) = a_{\gamma_{X,U}}(t) \cdot F(c(\gamma_U(0)))$$

Differentiate both sides with respect to t :

$$\frac{d}{dt} F(c(\gamma_U(t))) = \frac{d}{dt}\left[a_{\gamma_{X,U}}(t) \cdot F(c(\gamma_U(0)))\right]$$

Left side (chain rule):

$$\frac{d}{dt} F(c(\gamma_U(t))) = dF_{c(\gamma_U(t))}\left(\frac{dc}{du} \cdot \frac{d\gamma_U}{dt}\right) = \nabla_u F(c(u))|_{u=\gamma_{U(t)}} \cdot \dot{\gamma}_U(t)$$

Then for any content path $\gamma_U(t)$:

$$\left\|\frac{d}{dt} F(c(\gamma_U(t)))\right\| = \|\nabla_u F(c(u)) \cdot \dot{\gamma}_U(t)\| \leq \|\nabla_u F(c(u))\| \cdot \|\dot{\gamma}_U(t)\|$$

Right side:

$$\dot{a}_{\gamma_{X,U}}(t) \cdot F(c(\gamma_U(0)))$$

This gives a differential constraint:

$$\nabla_u F(c(u)) \cdot \dot{\gamma}_U(t) = \dot{a}_{\gamma_{X,U}}(t) \cdot F(c(\gamma_U(0)))$$

For any content path $\gamma_U(t)$:

$$\|\dot{a}_{\gamma_{X,U}}(t) \cdot F(c(\gamma_U(0)))\| = \|\nabla_u F(c(u)) \cdot \dot{\gamma}_U(t)\| \leq \|\nabla_u F(c(u))\| \cdot \|\dot{\gamma}_U(t)\|$$

Recall standard smoothness regularization in manifold learning (auto-encoder):

$$\mathcal{L}_{\text{smooth}} = \int_{U_\alpha} \|\nabla_u F(c(u))\|^2 \, du$$

which penalizes large gradients of F with respect to content u, thus encouraging F to vary smoothly across content space.

If $\|\nabla_u F(c(u))\|$ is samll, which means $F\left(c\left(\gamma_U(t)\right)\right)$ changes slowly with t, then $\dot{a}_{\gamma_{X,U}}(t)$ is small, which means $a_{\gamma_{X,U}}(t)$ varies slowly, gives a smoothly transport $X \to Z$.

So regularization ensures a special case of PE where the transport is smooth and bounded.

Remark 4.32. *If a global section c does not exist, cover U by charts* $\{U_\alpha\}$ *with local sections* c_α *and apply the construction per chart; on overlaps* $U_\alpha \cap U_\beta$*, the corresponding feature fields match through the induced gauge transform in A. This is exactly atlas AE.*

todo: connect to chart manifold.

todo: group / steerable / gauge CNN

5 Results

Present the results of your study here and answer the research questions, asked earlier in the thesis (in the introduction, perhaps), this study strives to answer. The scientific value of your work is measured by the results you obtain along with the arguments you give to back the answers to your research questions.

Be critical of the significance of your results. You may critically scrutinise the results and your interpretation of the results here, or you may do so later in the chapter with the discussion of your work or in the conclusions part.

This part should discuss how reliable the data used in the study are. You may discuss the reliability of the conclusions drawn from the study either in this chapter or later in the discussions part. You may have the discussion in a chapter of its own, separate from the summary or conclusions.

6 Summary/Conclusions

This is where you tie up any loose ends. Tell your reader briefly and clearly what you have done, what you have discovered, and the value of your discovery in the context of similar work done earlier. Draw clear conclusions regarding the research problem, sub-problems or hypotheses. You also discuss future lines of study and new questions your study might have posed.

As the author of the thesis, you alone are responsible for ensuring that the layout, form and structure of your thesis adheres to the guidelines outlined by your school. This template aims to help you meet these requirements.

Finally, a dummy citation [**Dyson**] to get a reference to an item in the bibliography.

References

[1] D. Bank, N. Koenigstein, and R. Giryes. "Autoencoders". In: *Machine learning for data science handbook: data mining and knowledge discovery handbook* (2023), pp. 353–374.

[2] T. S. Cohen and M. Welling. "Group Equivariant Convolutional Networks". In: *CoRR* abs/1602.07576 (2016). arXiv: 1602.07576. URL: http://arxiv.org/abs/1602.07576.

[3] G. Cybenko. "Approximation by superpositions of a sigmoidal function". In: *Mathematics of control, signals and systems* 2.4 (1989), pp. 303–314.

[4] R. A. Horn and C. R. Johnson. *Matrix Analysis*. 2nd ed. Cambridge University Press, 2012.

[5] K. Hornik. "Approximation capabilities of multilayer feedforward networks". In: *Neural networks* 4.2 (1991), pp. 251–257.

[6] C.-H. Huang, J. S. Chen, and C. C. Hu. "The Schatten p-norm on Rn". In: *J. Nonlinear Convex Anal* 21 (2020), pp. 21–29.

[7] S. Ioffe and C. Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift". In: *International conference on machine learning*. pmlr. 2015, pp. 448–456.

[8] A. Kratsios. "The universal approximation property: Characterization, construction, representation, and existence". In: *Annals of Mathematics and Artificial Intelligence* 89.5 (2021), pp. 435–469.

[9] D. Kunin et al. "Loss Landscapes of Regularized Linear Autoencoders". In: *CoRR* abs/1901.08168 (2019). arXiv: 1901.08168. URL: http://arxiv.org/abs/1901.08168.

[10] J. M. Lee. "Smooth manifolds". In: *Introduction to smooth manifolds*. Springer, 2003, pp. 1–29.

[11] M. Leshno et al. "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function". In: *Neural networks* 6.6 (1993), pp. 861–867.

[12] C.-K. Li and W. So. "Isometries of p-norm". In: *The American Mathematical Monthly* 101.5 (1994), pp. 452–453. DOI: 10.1080/00029890.1994.11996972. eprint: https://doi.org/10.1080/00029890.1994.11996972. URL: https://doi.org/10.1080/00029890.1994.11996972.

[13] C.-L. Liu. "Selected Topics in Engineering Mathematics: Advanced Operations in Linear Algebra". In: (2024).

[14] K. B. Petersen, M. S. Pedersen, et al. "The matrix cookbook". In: *Technical University of Denmark* 7.15 (2008), p. 510.

[15] A. Pinkus. “Approximation theory of the MLP model in neural networks”. In: *Acta numerica* 8 (1999), pp. 143–195.

[16] S. Rifai et al. “Contractive auto-encoders: Explicit invariance during feature extraction”. In: *Proceedings of the 28th international conference on international conference on machine learning*. 2011, pp. 833–840.

[17] D. E. Worrall et al. “Harmonic Networks: Deep Translation and Rotation Equivariance”. In: *CoRR* abs/1612.04642 (2016). arXiv: 1612.04642. URL: http://arxiv.org/abs/1612.04642.