Lecture 2: Groups, Vector Bundles, and Gauge Theory

Geometry and Topology in Machine Learning Seminar

June 13th, 2025

Outline

Group

Definition (Group)

A group is a set $\mathcal{G}$ together with a binary operation

$$\circ \colon \mathcal{G} \times \mathcal{G} \longrightarrow \mathcal{G}, \qquad (g, h) \mapsto g \circ h,$$

satisfying the following axioms:

- **(i) Associativity:** $(gh)k = g(hk)$ for all $g, h, k \in \mathcal{G}$
- **(ii) Identity:** There exists a unique element $e \in \mathcal{G}$ such that $eg = ge = g$ for all $g \in \mathcal{G}$
- **(iii) Inverse:** For each $g \in \mathcal{G}$ there exists a unique inverse $g^{-1} \in \mathcal{G}$ such that $g\,g^{-1} = g^{-1}g = e$
- **(iv) Closure:**[1] For all $g, h \in \mathcal{G}$, the product gh lies in $\mathcal{G}$

[1] Note that this is actually implied by the statement $\circ \colon \mathcal{G} \times \mathcal{G} \to \mathcal{G}$

Examples of Groups used in ML

- Translation groups $\mathbb{Z}^d$ or $\mathbb{R}^d$ for CNNs. Here, the group operations are vector additions.

  As we saw in the last lecture, the convolution operator is the unique linear map that commutes with these translations, and CNNs implement this equivariance via weight sharing.
- Permutation groups S_n. Here S_n is the set of all bijections $h : \{1, ..., n\} \to \{1, ..., n\}$ and $\circ$ is the composition of functions.

  GNNs operate graph-based data by enforcing S_n equivariance under node relabeling. Permutation equivariance relaxes the classical i.i.d. assumption to the weaker notion of exchangeability in GNNs.

Examples of Groups used in ML

- Special Orthogonal Groups $SO(n)$ - the set of $n \times n$ real orthogonal matrices with determinant $+1$ forms a group under matrix multiplication. Recall that Spherical CNNs apply $SO(3)$-equivariant convolutions. More generally, the set of $n \times n$ invertible real matrices forms a group under matrix multiplication.
- Euclidean groups $E(n)$ - the set of distance preserving maps $\phi : \mathbb{R}^n \to \mathbb{R}^n$ with group operation being composition. These are especially useful for tasks like protein/molecule modeling or physics simulations, where only relative geometry drives the prediction.

Abelian group and group generator

Definition (Abelian Group)

A group $(G, \cdot)$ is called Abelian or commutative if

$$g\,h = h\,g \quad \text{for all } g, h \in G.$$

Note that S_1, S_2 and $SO(2)$ are abelian but in general $S_{n\geq 3}$, $SO(n)$ with $n \neq 2$, and $E(n)$ are in general nonabelian. $\mathbb{Z}^n$ and $\mathbb{R}^n$ are abelian for sure.

Definition (Generated Subset)

Let $(G, \cdot)$ be a group and $S \subseteq G$. We say G is *generated* by S (and write $G = \langle S \rangle$) if every element of G can be written as a finite product of elements of S and their inverses, i.e.

$$\langle S \rangle \;=\; \left\{ s_1^{\pm 1} s_2^{\pm 1} \cdots s_n^{\pm 1} : n \geq 1,\; s_i \in S \right\} = G.$$

Subgroup and Coset

Definition (Subgroup)

Let $(G, \cdot)$ be a group. A nonempty subset $H \subseteq G$ is called a subgroup (denoted $H \leq G$) if it is closed under the group operation and inverses, i.e. for all $h_1, h_2 \in H$,

$$h_1 \cdot h_2 \in H \quad \text{and} \quad h_1^{-1} \in H$$

Definition (Coset)

Let $H \leq G$ be a subgroup and $g \in G$.

- The left coset of H by g is $gH = \{g \cdot h : h \in H\}$
- The right coset of H by g is $Hg = \{h \cdot g : h \in H\}$

We use G/H to denote the collection of left cosets of H. Note that the left cosets of H in G partition G into disjoint sets.

Example. If $G = \mathbb{Z}$ and $H = 2\mathbb{Z}$, then the two cosets are odd and even integers $G/H = \{0 + 2\mathbb{Z},\ 1 + 2\mathbb{Z}\}$.

Group Actions

Definition (Group Action)

Let $(G, \cdot)$ be a group with identity e, and let Ω be a set. A (left) action of G on Ω is a map

$$\alpha\colon G \times \Omega \longrightarrow \Omega, \qquad (g, u) \mapsto g \cdot u,$$

satisfying for all $g, h \in G$ and $u \in \Omega$:

(i) $e \cdot u = u$.

(ii) $g \cdot (h \cdot u) = (gh) \cdot u\,.$

Informally speaking, a group action is simply a rule that assigns each group element a transformation of a set so that doing one transformation after another is the same as doing them in combination.

Group Actions

A group action of G on Ω is

1. **transitive** if for all $x, y \in \Omega$, there exists some $g \in G$ such that $g \cdot x = y$.
2. **free** if $g \cdot x = x$ for some $g \in G, x \in \Omega$ implies $g = e$.

Here are some examples:

1. $\mathrm{GL}_n(\mathbb{R})$ acts on $\mathbb{R}^n$ by matrix multiplication. This action is neither transitive nor free.
2. More generally, any subgroup of $GL(V)$ defines a group action on the vector space V by $\alpha : G \times V \to V, \quad (g, v) \mapsto g(v)$, i.e., each group element acts as a linear transformation on V.
3. G acts on G itself by left multiplication. This is both free and transitive.

Group Representations

Recall that the power of geometric models lies in building their hypothesis class around data symmetries from the very first linear mapping. e.g. via operations like $w^\top x$. Linear group actions, known as group representations, play a key role in ML.

Definition (Real Representation)

Let G be a group. An n-dimensional real representation of G is a map

$$\rho : G \longrightarrow \mathrm{GL}_n(\mathbb{R}), \qquad g \mapsto \rho(g),$$

such that for all $g, h \in G$

$$\rho(gh) \; = \; \rho(g)\,\rho(h).$$

The representation is called

- orthogonal if each $\rho(g)$ is an orthogonal matrix (over $\mathbb{R}$).

Homogeneous Space

Definition (Homogeneous Space)

Let G be a group acting on a nonempty set X via

$$\alpha \colon G \times X \to X, \qquad (g, x) \mapsto g \cdot x.$$

Then X is called a homogeneous space (or G-space) if the action is transitive.

Equivalently, fixing any basepoint $o \in X$ and its stabilizer subgroup $H = \{h \in G : h \cdot o = o\}$, one has a canonical G-equivariant bijection

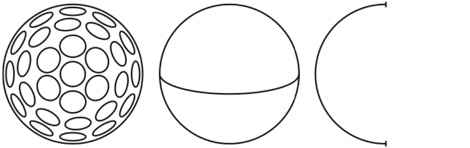
$$X \;\cong\; G/H,$$

where G/H is the set of left cosets of H in G.

Spherical CNNs

Recall: spherical CNNs replace the translation group by $SO(3)$ and equip signals $\psi, f : S^2 \to \mathbb{R}^K$with the inner product $\langle \psi, f \rangle \; := \; \int_{S^2} \sum_{k=1}^{K} \psi_k(x)\, f_k(x)\, d\mu_{S^2}$.

A key claim: a spherical CNN is exactly an equivariant CNN on the homogeneous space $SO(3)/SO(2) \cong S^2$.



Picture from [Cohen et al., 2019]:

$SO(3)$ as a principle $SO(2)$ bundle over S^2

Spherical CNNs

Consider the action

$$SO(3) \times S^2 \longrightarrow S^2, \text{ by matrix multiplication.}$$

Observe that this is a transitive group action.

Then $H =$ rotations about the z–axis is exactly the stabilizer of the north pole $(0,0,1)$. H is also known as $SO(2)$, and hence we have that

$$SO(3)/SO(2) \cong S^2.$$

Spherical CNNs

Representations. Given $H = SO(2)$ and choose $\rho_{\text{in}}, \rho_{\text{out}} : H \to GL(V_{\text{in}}), GL(V_{\text{out}})$.

An equivariant kernel[2] $\kappa\colon S^2 \to \mathrm{Hom}(V_{\text{in}}, V_{\text{out}})$ satisfies

$$\kappa\big(h \cdot x\big) \;=\; \rho_{\text{out}}(h)\,\kappa(x)\,\rho_{\text{in}}(h)^{-1}, \quad \forall\, h \in H,\; x \in S^2.$$

[2] Details will be given next lecture

Outline

Domain vs. Feature Space

Recall that every dataset we study is anchored to a geometric domain Ω—the set where the raw samples live, such as the pixel grid $\mathbb{Z}^2$, graph vertex set V, or the sphere S^2. In many cases, we can recognize Ω as a homogeneous space G/H. Neural layers, however, are not built on Ω itself. They operate on the corresponding feature space

$$\mathcal{F}(\Omega, C) = \{ x : \Omega \to C \},$$

the vector space of signals (stacks of C-dimensional feature channels) attached point-wise to Ω. All layer maps are linear or nonlinear transformations of these signal fields, subject to the desired equivariance with respect to the symmetries of the domain.

Equivariance in the feature space

Definition (G-Equivariant Function)

If C is a group representation, then G can act on $\mathcal{F}(\Omega, C)$ via $\rho(g)$. For simplicity we omit C, a map

$$f\colon \mathcal{F}(\Omega) \longrightarrow \mathcal{F}(\Omega)$$

is called G-*equivariant* if

$$f\big(\rho(g)x\big) = \rho(g)\big(f(x)\big) \quad \text{for all } g \in G,\ x \in \mathcal{F}(\Omega).$$

Equivalently, this square commutes:

$$\begin{tikzcd}
\mathcal{F}(\Omega) \arrow[r, "f"] \arrow[d, "\rho(g)"'] & \mathcal{F}(\Omega) \arrow[d, "\rho(g)"] \\
\mathcal{F}(\Omega) \arrow[r, "f"'] & \mathcal{F}(\Omega)
\end{tikzcd}$$

Invariance in the feature space

Definition (G-Invariant Function)

Let G act on the signal space $\mathcal{F}(\Omega)$ via $\rho(g)x$, and let

$$f\colon \mathcal{F}(\Omega) \longrightarrow \mathcal{Y}$$

be any map. We say f is G-*invariant* if

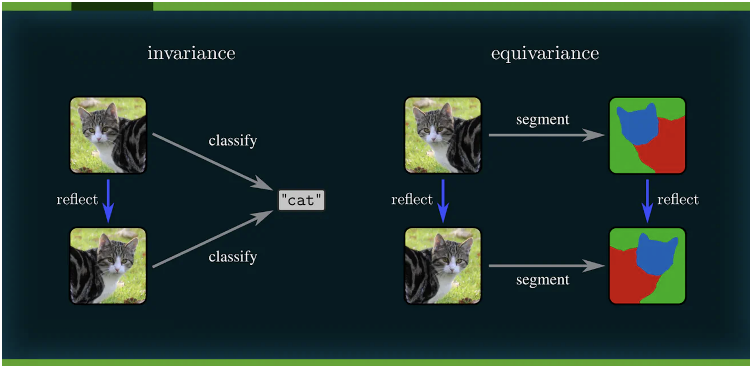
$$f\big(\rho(g)x\big) = f(x) \quad \text{for all } g \in G,\ x \in \mathcal{F}(\Omega).$$

Equivalently, this square commutes:

$$\begin{tikzcd} \mathcal{F}(\Omega) \arrow[d, "\rho(g)"'] \arrow[dr, "f"] & \\ \mathcal{F}(\Omega) \arrow[r, "f"'] & \mathcal{Y} \end{tikzcd}$$

Note that being G-invariant is the same as G-equivariant, when the action on $\mathcal{Y}$ is trivial.

Example



Picture from [Weiler et al., 2025]

Outline

Fiber Bundles

Definition (Fiber Bundle)

A fiber bundle is a quadruple (E, B, π, F) consisting of a total space E, base space B, fiber F, and a continuous surjection

$$\pi \colon E \longrightarrow B,$$

called the projection, such that for every $b \in B$ there exists an open neighborhood $U \subseteq B$ and a homeomorphism

$$\varphi_U \colon \pi^{-1}(U) \longrightarrow U \times F$$

making $\mathrm{proj}_1 \circ \varphi_U = \pi$ on $\pi^{-1}(U)$, and the maps φ_U are called local trivializations.

Transition Maps

For two trivializing open sets U_i, U_j, we locally have

$$\varphi_i : \pi^{-1}(U_i) \to U_i \times F,$$

$$\varphi_j : \pi^{-1}(U_j) \to U_j \times F.$$

The composition $\varphi_j \circ \varphi_i^{-1}$ defines a map

$$U_i \cap U_j \times F \to U_i \cap U_j \times F.$$

Because of the axioms, this map must be of the form

$$(x, f) \mapsto (x, g_{ij}(f))$$

for some function $g_{ij} : U_i \cap U_j \to \mathrm{Aut}(F)$[3].

[3] This is the group of continuous self-homeomorphisms on F with group operation being composition

Vector Bundles (bundles of vector spaces)

A vector bundle is simply a special case of fiber bundles that the fiber F is a finite-dimensional vector space $\mathbb{R}^k$ and the transition map lies in $GL(F)$, i.e., $g_{ij} : U_i \cap U_j \to GL(F)$.

Definition (Vector Bundle)

A vector bundle of rank k over a base space B is a fiber bundle

$$\pi \colon E \longrightarrow B$$

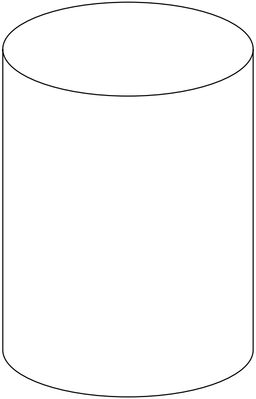
whose typical fiber F is a k-dimensional real vector space, and for which there exists an open cover $\{U_i\}$ of B and homeomorphisms

$$\varphi_{U_i} \colon \pi^{-1}(U_i) \xrightarrow{\sim} U_i \times \mathbb{R}^k$$

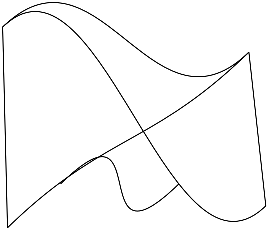
making $\pi = \mathrm{proj}_1 \circ \varphi_{U_i}$ and such that each φ_{U_i} restricts on fibers to a linear isomorphism.

Example of Vector Bundles

The following are two different real line bundles over S^1.



Cylinder



Mobius Band

Principle G-Bundles

A principle G-bundle is a specialized fiber bundle whose fibers are G-torsors, spaces on which the group G acts freely and transitively.

Definition (Principal G-Bundle)

A principal G-bundle over a base B is a fiber bundle

$$\pi : P \longrightarrow B$$

equipped with a continuous right action $P \times G \to P$ such that:

- **Fiber Preservation:** $\pi(p \cdot g) = \pi(p)$ for all $p \in P$, $g \in G$.
- **Free & Transitive Action:** G acts freely and transitively on each fiber $\pi^{-1}(b)$.
- **Local Triviality:** $\pi : P \to B$ locally looks like the projection $U \times G \to U$.

Principle G-Bundles: claim of terminology

- Note that principal G-bundles are sometimes called bundles of groups in the literature.
- But this is misleading, they are in fact bundles of G-torsors. A G-torsor is just a set on which G acts freely and transitively, meaning there is no perferred "origin" within.
- Informally, one can regard it as a group without the identity element.

Sections of Bundles

Definition (Section of a Fiber Bundle)

Let $\pi\colon E \to B$ be a fiber bundle with typical fiber F. A section of π is a continuous map

$$s\colon B \longrightarrow E \quad \text{such that} \quad \pi\big(s(b)\big) = b \quad \forall\, b \in B.$$

In other words, $s(b)$ picks out one point in each fiber E_b.

Examples:

- Vector bundle $E \to B$: A section is a continuous choice of vector in each fiber. E.g. for any vector bundle, there is always a continuous choice of the zero vector (known as the zero section).
- Principal G-bundle $P \to B$: A section is a map $s\colon B \to P$ such that $s(b)$ lies in the torsor P_b. Globally defined sections exist exactly when the bundle is trivial (i.e. $P \cong B \times G$).

Aside: Principal G-Bundles and Gauge Theory

Mathematically, gauge theory is the study of gauges, which are defined to be local sections of principal G-bundles. One can think of gauges as generalized frames that express some geometric quantities.

Associated Vector Bundle

Definition (Associated Vector Bundle)

Let $\pi\colon P \to B$ be a principal G-bundle and let $\rho\colon G \to \mathrm{GL}(V)$ be a representation on a finite-dimensional vector space V. The associated vector bundle $E = P \times_\rho V$ is the quotient

$$E \;=\; (P \times V)\big/\!\sim,$$

where $(p \cdot g,\, v) \sim (p,\, \rho(g)\, v)$ for all $g \in G$, $p \in P$, $v \in V$. Its projection $\pi_E\colon E \to B$ is given by $\pi_E([p, v]) = \pi(p)$, making E a rank-$\dim V$ vector bundle over B.

Example of Associated Vector Bundle

Let $\mathbb{C}^\times$ denote the non-zero complex numbers whose group structure is multiplication.

- Given a principal $\mathbb{C}^\times$-bundle $\pi : P \to B$, we can always "add zeroes" to π to obtain a complex line bundle.
- Given a complex line bundle, we can always remove the zero section to get a principal $\mathbb{C}^\times$-bundle.

Another name for $\mathbb{C}^\times$ is $\mathrm{GL}_1(\mathbb{C})$. In general, there is a correspondence between

$$\text{principal } \mathrm{GL}_n(\mathbb{F}) - \text{bundles and associated } \mathbb{F} - \text{vector bundles},$$

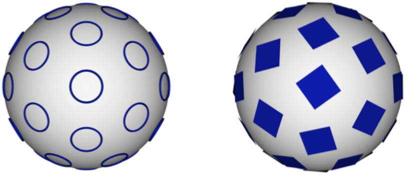
where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

A paraphrase of feature spaces

Here we will firstly show that in the spherical CNN, feature spaces on the sphere

$$\mathcal{F} = \Gamma(SO(3) \times_{SO(2)} V\} \cong \{\text{V-values fields on } S^2\}$$

where the fields mean a section of an associated bundle.



Pictures from [Cohen et al., 2019]

Then we will lift to a general theory: in a homogeneous space, feature spaces can be considered as spaces of sections of the associated vector bundle given a representation.

Vector-valued Feature Space on S^2

1. **Principal H–bundle:**

$$\pi : SO(3) \longrightarrow SO(3)/SO(2) \cong S^2,$$

with $H = SO(2)$ acting freely on the right of $SO(3)$.

2. **Associated vector bundle:** Let $\rho(h) : SO(2) \to GL(V)$ be the 2×2 rotation matrix on $V = \mathbb{R}^2$. Form

$$E = SO(3) \times_{\rho(h)} V = \big(SO(3) \times V\big) \Big/ \big((g\,h, v) \sim (g, \rho(h)v)\big),$$

which inherits a projection $E \to S^2$.

3. **Sections (feature space):**

$$\mathcal{F} = \Gamma\big(S^2, E\big) = \left\{ s : S^2 \to E \;\middle|\; \pi \circ s = \mathrm{id}_{S^2} \right\}.$$

Equivalently,

$$\mathcal{F} \cong \left\{ f : SO(3) \to V \;\middle|\; f(gh) = \rho_1(h)^{-1} f(g) \right\},$$

which identifies each section with a V-valued field on S^2.

$$\boxed{\mathcal{F} = \Gamma\big(SO(3) \times_{SO(2)} V\big) \cong \{V\text{-valued fields on } S^2\}.}$$

Outline

Towards a General Theory

Let $M = G/H$ be a homogeneous space. Fix a representation $\rho : H \to \mathrm{GL}(V)$. We'll then get the associated vector bundle

$$E \;=\; G \times_\rho V \;\longrightarrow\; M = G/H.$$

and the resulting feature spaces

$$\mathcal{F} \;=\; \Gamma(M, E) \;=\; \{\, s : M \to E \mid \pi \circ s = \mathrm{id}_M \},$$

i.e. the space of sections of E.

As we stated at the beginning of this section, on a homogeneous space $M = G/H$, feature spaces are exactly spaces of sections of the associated bundle $G \times_\rho V$.

Equivalence of Sections and Mackey Functions

Definition

Let V be an H-representation, a Mackey function is a function $f : G \to V$ such that

$$f(gh) = \rho(h)^{-1} f(g), \forall g \in G, h \in H.$$

Theorem

Let $E = G \times_\rho V \xrightarrow{\pi} G/H$ be the associated bundle for $\rho : H \to \mathrm{GL}(V)$. Then there is an isomorphism

$$\Gamma(E) \cong \mathcal{I}_G = \{\text{collection of all Mackey functions}\}.$$

Equivalence of Sections and Mackey Functions

Proof. Let $s(x) \in G$ be a fixed coset representative for $x \in G/H$. Consider the map $\Lambda : \Gamma(E) \to \mathcal{I}_G$ given by

$$\Lambda f(x) \coloneqq \rho(\mathrm{h}(g)^{-1}) f(gH),$$

where $\mathrm{h}(g) = s(gH)^{-1} g \in H$.

We claim it has an inverse $\Lambda^{-1} : \mathcal{I}_G \to \Gamma(E)$ given by

$$\Lambda^{-1} f'(x) \coloneqq f'(s(x)).$$

Equivalence of Sections and Mackey Functions

Indeed, we check that

$$\begin{aligned}
[\Lambda^{-1}[\Lambda f]](x) &= [\Lambda f](s(x)) \\
&= \rho(\mathrm{h}(s(x))^{-1}) f(s(x)H) \\
&= \rho((s(s(x)H)^{-1} s(x))^{-1}) f(x) \\
&= \rho(s(x)^{-1} s(x)) f(x) \\
&= f(x).
\end{aligned}$$

Conversely, we have that

$$\begin{aligned}
[\Lambda[\Lambda^{-1} f']](g) &= \rho(\mathrm{h}(g)^{-1}) [\Lambda^{-1} f'](gH) \\
&= \rho(\mathrm{h}(g)^{-1}) f'(s(gH)) \\
&= f'(s(gH)\,\mathrm{h}(g)) \qquad f' \text{ is a Mackey function} \\
&= f'(s(gH) s(gH)^{-1} g) \\
&= f'(g).
\end{aligned}$$

Induced Representation

We also want to describe how features transform (ie. for spherical GNNs, how the vectors are changed under rotations). The general tool to describe them in this case comes from induced representations.

Definition

Let $H \leq G$ and $\rho : H \to \mathrm{GL}(V)$ be an H-representation. The induced representation $\pi = \mathrm{Ind}_H^G \rho$ is the data:

1. The vector space $\mathcal{I}_G$
2. An action of G on $\mathcal{I}_G$ by

$$g \cdot f(k) := f(g^{-1}k)$$

References I

Cohen, T. S., Geiger, M., and Weiler, M. (2019).
A general theory of equivariant CNNs on homogeneous spaces.
Curran Associates Inc., Red Hook, NY, USA.

Weiler, M., Forré, P., Verlinde, E., and Welling, M. (2025).
Equivariant and coordinate independent convolutional networks.
WORLD SCIENTIFIC.