EECE 517F: Advanced Topics in Deep Learning **Spring 2026**

Lecture 4: Steerable Group Convolutions

Lecturer: Renjie Liao *Date: Jan 29* *Scribes: Renjie Liao*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*
Disclaimer: *These notes are written with the help from Gemini and ChatGPT.*

4.1 Feature Types (Representations)

In steerable CNNs, each feature vector at a spatial location lives in a *fiber space* V and comes with a **type**, specified by a group representation

$$\rho : G \to GL(V). \tag{4.1}$$

Intuitively, the type tells us how the feature channels should transform when we apply a symmetry $g \in G$.

Field interpretation You can interpret a feature map as a *field* over space:

- If $V = \mathbb{R}$ (scalar type), then $f(\mathbf{x}) \in \mathbb{R}$ and the feature map is a **scalar field**.
- If $V = \mathbb{R}^d$ (vector type), then $f(\mathbf{x}) \in \mathbb{R}^d$ and the feature map is a **vector field**.

More generally, $\dim(V)$ equals the number of channels per field element (per "type instance"), and a layer typically contains multiple copies of various types.

4.1.1 Concrete Examples

- **Scalar type (invariant):** $V = \mathbb{R}$ and $\rho(g) = 1$. A scalar (*e.g.*, intensity) does not rotate; only its spatial position changes.
- **Vector type:** $V = \mathbb{R}^d$ and $\rho(g) = R(g)$, the standard rotation matrix. A 2D/3D vector (*e.g.*, gradient or velocity) rotates with the coordinate frame.
- $SO(3)$ **irreps (spherical-harmonic types):** for $l = 0, 1, 2, \dots$, take $V = \mathbb{C}^{2l+1}$ and

  $$\rho(g) = D^l(g) \in \mathbb{C}^{(2l+1)\times(2l+1)}. \tag{4.2}$$

  A feature of type-l has $2l + 1$ channels that mix under rotation according to $D^l(g)$.

4.1.2 Irreducible Representations (Irreps)

A representation $\rho : G \to GL(V)$ is called *reducible* if there exists a nontrivial subspace $W \subset V$ such that $\rho(g)W \subseteq W$ for all $g \in G$. If no such proper invariant subspace exists, the representation is *irreducible* (an **irrep**).

Why irreps matter for steerable networks Irreps are the "atomic" building blocks of representations: for many groups of interest (*e.g.*, finite groups and compact Lie groups such as $SO(3)$), any finite-dimensional representation can be decomposed into a direct sum of irreps,

$$V \cong \bigoplus_i m_i V_i, \tag{4.3}$$

where each V_i carries an irrep and $m_i \in \mathbb{N}$ is its multiplicity. In steerable CNNs, choosing feature types as (copies of) irreps makes the transformation law clean and structured: under a symmetry, channels only *mix within each irrep block.*

SO(3) example (angular momentum) For $SO(3)$, the irreps are indexed by a degree $l \in \{0, 1, 2, \dots\}$. The corresponding (complex) irrep D^l has dimension $2l + 1$ and acts on the coefficient vector of spherical harmonics of degree l. This is why a type-l field naturally has $2l + 1$ channels.

Induced action on a field (recall) If $f : \mathbb{R}^d \to V$ is a feature field, we write $\pi(g)$ for the **induced action** (also called the *induced representation*) of the group element g on the space of fields. That is, $\pi(g)$ is an operator that takes a field f and returns the transformed field $\pi(g)f$. With the type/representation ρ on the fiber, this action is

$$[\pi(g)f](\mathbf{x}) = \rho(g)\, f(g^{-1}\mathbf{x}). \tag{4.4}$$

Steerable convolutions are designed so that their output transforms according to this rule.

4.2 Steerable Group Convolutions

A steerable convolution maps an input field f_{in} (of type/representation ρ_{in}) to an output field f_{out} (of type ρ_{out}). For simplicity, we write the spatial convolution on $\mathbb{R}^2$ as

$$f_{\text{out}}(\mathbf{x}) = \int_{\mathbb{R}^2} K(\mathbf{x} - \mathbf{y})\, f_{\text{in}}(\mathbf{y})\, d\mathbf{y}. \tag{4.5}$$

Here the kernel is matrix-valued: $K : \mathbb{R}^2 \to \text{Hom}(V_{\text{in}}, V_{\text{out}})$.

4.2.1 Kernel Constraint (Equivariance)

We now derive the kernel constraint from the equivariance requirement, using the standard **group convolution** form (with h^{-1} acting on the argument).

Step 1: Group convolution form Let a field $f : X \to V_{\text{in}}$ live on a space X on which G acts (e.g., $X = \mathbb{R}^d$ with rotations). A common way to write a G-equivariant correlation/convolution is

$$[K * f](\mathbf{x}) := \int_G K(h^{-1} \cdot \mathbf{x})\, f(h)\, dh, \tag{4.6}$$

where dh is the Haar measure on G. (Here, the key structural point is the appearance of h^{-1} inside $K(\cdot)$, exactly as in the usual group convolution $k(h^{-1}g)$ on G.)

Step 2: State equivariance Let π_{in} and π_{out} be the induced actions on input and output fields: $[\pi(g)f](\mathbf{x}) = \rho(g)f(g^{-1}\mathbf{x})$. Equivariance requires

$$\pi_{\text{out}}(g)\,(K * f_{\text{in}}) = K * (\pi_{\text{in}}(g) f_{\text{in}}) \quad \forall g \in G. \tag{4.7}$$

Step 3: Expand both sides and match terms Expanding the left-hand side gives

$$\begin{aligned}[\pi_{\text{out}}(g)(K * f_{\text{in}})](\mathbf{x}) &= \rho_{\text{out}}(g)\,(K * f_{\text{in}})(g^{-1} \cdot \mathbf{x}) \\ &= \rho_{\text{out}}(g) \int_G K(h^{-1} \cdot (g^{-1} \cdot \mathbf{x}))\, f_{\text{in}}(h)\, dh.\end{aligned} \tag{4.8}$$

For the right-hand side,

$$[K * (\pi_{\text{in}}(g) f_{\text{in}})](\mathbf{x}) = \int_G K(h^{-1} \cdot \mathbf{x})\, \rho_{\text{in}}(g)\, f_{\text{in}}(g^{-1}h)\, dh. \tag{4.9}$$

Make the change of variables $h = gh'$ (so $g^{-1}h = h'$ and $dh = dh'$ by left-invariance of Haar measure), yielding

$$[K * (\pi_{\text{in}}(g) f_{\text{in}})](\mathbf{x}) = \int_G K((gh')^{-1} \cdot \mathbf{x})\, \rho_{\text{in}}(g)\, f_{\text{in}}(h')\, dh'. \tag{4.10}$$

Since $(gh')^{-1} = h'^{-1}g^{-1}$, both sides depend on $f_{\text{in}}(h')$. Requiring Eq. (4.7) to hold for all inputs implies the kernel must satisfy

$$K(g \cdot \mathbf{x}) = \rho_{\text{out}}(g)\, K(\mathbf{x})\, \rho_{\text{in}}(g)^{-1}. \tag{4.11}$$

4.2.2 Practical Parameterization via a Steerable Basis

Instead of learning K freely pixel-by-pixel, we solve the equivariance constraint $K(g\cdot\mathbf{x}) = \rho_{\text{out}}(g)K(\mathbf{x})\rho_{\text{in}}(g)^{-1}$ to obtain its solution space. We choose a basis of this solution space, $\{\Psi_1, \ldots, \Psi_B\}$, so that any kernel written as a linear combination of Ψ_b is *automatically equivariant*. We then parameterize

$$K(\mathbf{x}) = \sum_{b=1}^{B} w_b \, \Psi_b(\mathbf{x}), \tag{4.12}$$

and learn only the coefficients $\{w_b\}$.

4.3 Example: $SO(3)$ and Spherical Harmonics

This section ties together the key ingredients above (types, induced actions, equivariance constraints, and steerable bases) in the concrete case of 3D rotations.

4.3.1 The Group Action

The group $SO(3)$ is the set of all 3D rotations. It acts on points on the unit sphere S^2 by rotating their 3D coordinates. We write this action as $\mathbf{x} \mapsto g \cdot \mathbf{x}$.

4.3.2 Signals and Feature Types on the Sphere

A scalar signal on the sphere is a function $f : S^2 \to \mathbb{R}$. More generally, a feature field on the sphere is a function $f : S^2 \to V$ where V is the fiber space. Choosing a type $\rho : SO(3) \to GL(V)$ specifies how f should transform under rotations via the induced action

$$[\pi(g)f](\mathbf{x}) = \rho(g)\, f(g^{-1} \cdot \mathbf{x}). \tag{4.13}$$

4.3.3 Spherical Harmonics as a Canonical Basis

Spherical harmonics $\{Y_l^m\}$ form an orthonormal basis for scalar functions on S^2. They are indexed by

$$l \in \{0, 1, 2, \ldots\}, \qquad m \in \{-l, \ldots, l\}. \tag{4.14}$$

Rotation acts blockwise A central fact is that under rotation, spherical harmonics of different degree l do not mix. For any $g \in SO(3)$,

$$Y_l^m(g^{-1} \cdot \mathbf{x}) = \sum_{m'=-l}^{l} D^l_{m,m'}(g)\, Y_l^{m'}(\mathbf{x}), \tag{4.15}$$

where $D^l(g) \in \mathbb{C}^{(2l+1)\times(2l+1)}$ is the l-th irreducible representation (Wigner-D matrix).

A concrete type choice This immediately gives a family of natural feature types for $SO(3)$: for each l, choose $V = \mathbb{C}^{2l+1}$ and define

$$\rho(g) = D^l(g). \tag{4.16}$$

A feature of type-l thus has $2l + 1$ channels that mix under rotation according to $D^l(g)$.

4.3.4 From the Equivariance Constraint to a Steerable Basis (The $SO(3)$ Case)

To make the "steerable basis" completely concrete, we now specialize to the common setting where both the input and output types are $SO(3)$ irreps.

Setup: irrep-to-irrep kernels Let the input be a type-l_{in} field and the output be a type-l_{out} field, so

$$\rho_{\text{in}}(g) = D^{l_{\text{in}}}(g), \qquad \rho_{\text{out}}(g) = D^{l_{\text{out}}}(g). \tag{4.17}$$

Then the kernel is a matrix-valued function

$$K : S^2 \to \text{Hom}(\mathbb{C}^{2l_{\text{in}}+1}, \mathbb{C}^{2l_{\text{out}}+1}) \cong \mathbb{C}^{(2l_{\text{out}}+1)\times(2l_{\text{in}}+1)} \tag{4.18}$$

that must satisfy the equivariance constraint

$$K(g \cdot \mathbf{x}) = D^{l_{\text{out}}}(g)\, K(\mathbf{x})\, D^{l_{\text{in}}}(g)^{-1}. \tag{4.19}$$

Key representation-theoretic fact: tensor-product decomposition Eq. (4.19) says that (as a function of $\mathbf{x}$) the kernel must transform in the same way as the representation $D^{l_{\text{out}}} \otimes (D^{l_{\text{in}}})^*$. For $SO(3)$, tensor products decompose into irreps:

$$D^{l_{\text{out}}} \otimes (D^{l_{\text{in}}})^* \;\cong\; \bigoplus_{J=|l_{\text{out}}-l_{\text{in}}|}^{l_{\text{out}}+l_{\text{in}}} D^J. \tag{4.20}$$

This tells us that the angular dependence of an equivariant kernel can be expanded in spherical harmonics of degrees J in that range.

Concrete basis functions via Clebsch–Gordan coupling *Clebsch–Gordan (CG) coefficients* are the change-of-basis coefficients that decompose a tensor product of two irreducible representations into a direct sum of irreducibles. Concretely, given two $SO(3)$ irreps of degrees l_1 and l_2, the tensor product space has a natural "product" basis $\{|l_1 m_1\rangle \otimes |l_2 m_2\rangle\}$. CG coefficients $C^{JM}_{l_1 m_1,\, l_2 m_2}$ define the corresponding "coupled" basis vectors $|JM\rangle$ via

$$|JM\rangle \;=\; \sum_{m_1=-l_1}^{l_1} \sum_{m_2=-l_2}^{l_2} C^{JM}_{l_1 m_1,\, l_2 m_2}\, |l_1 m_1\rangle \otimes |l_2 m_2\rangle. \tag{4.21}$$

They satisfy the selection rule $M = m_1 + m_2$, and J ranges over $|l_1 - l_2|, \ldots, l_1 + l_2$. In our setting, the CG coefficients serve as the unique (up to convention) intertwining weights that "couple" the spherical-harmonic angular dependence (Y_J^m) with the channel indices $(m_{\text{in}}, m_{\text{out}})$ so that the resulting matrix-valued basis functions $\Psi^J(\mathbf{x})$ transform according to Eq. (4.19).

Let $C^{l_{\text{out}}\, m_{\text{out}}}_{l_{\text{in}}\, m_{\text{in}},\, Jm}$ denote Clebsch–Gordan (CG) coefficients. A standard equivariant basis on the sphere is:

$$[\Psi^J(\mathbf{x})]_{m_{\text{out}}, m_{\text{in}}} \;\triangleq\; \sum_{m=-J}^{J} C^{l_{\text{out}}\, m_{\text{out}}}_{l_{\text{in}}\, m_{\text{in}},\, Jm}\, Y_J^m(\mathbf{x}), \tag{4.22}$$

for each $J \in \{|l_{\text{out}} - l_{\text{in}}|, \ldots, l_{\text{out}} + l_{\text{in}}\}$. One can verify (using the transformation rules of Y_J^m and the defining property of CG coefficients) that these satisfy Eq. (4.19).

Steerable parameterization (explicit for $SO(3)$) With the basis (4.22), any equivariant kernel of this type can be parameterized as

$$K(\mathbf{x}) = \sum_{J=|l_{\text{out}}-l_{\text{in}}|}^{l_{\text{out}}+l_{\text{in}}} w_J\, \Psi^J(\mathbf{x}), \tag{4.23}$$

where w_J are learnable (complex) coefficients (or real coefficients with an appropriate real basis). In practice, one often also includes a learnable radial profile if the kernel lives on $\mathbb{R}^3$ rather than S^2.

4.3.5 (Optional) Where Y_l^m Comes From

Spherical harmonics can be derived as eigenfunctions of the Laplacian on the sphere by solving the eigenvalue problem

$$\Delta_{S^2} Y(\theta, \phi) = -\lambda\, Y(\theta, \phi), \tag{4.24}$$

where (θ, ϕ) are the standard spherical angles (colatitude and longitude).

Step 1: Write the spherical Laplacian On the unit sphere, the spherical Laplacian is

$$\Delta_{S^2} Y = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}. \tag{4.25}$$

Step 2: Separation of variables Assume a separable form

$$Y(\theta,\phi) = \Theta(\theta)\,\Phi(\phi). \tag{4.26}$$

Plugging into $\Delta_{S^2} Y = -\lambda Y$ and dividing by $\Theta\Phi$ gives

$$\frac{1}{\Theta\sin\theta}\frac{d}{d\theta}\left(\sin\theta\,\frac{d\Theta}{d\theta}\right) + \frac{1}{\Phi\sin^2\theta}\frac{d^2\Phi}{d\phi^2} = -\lambda. \tag{4.27}$$

Multiply both sides by $\sin^2\theta$ to separate the θ and ϕ dependence:

$$\sin\theta\,\frac{1}{\Theta}\frac{d}{d\theta}\left(\sin\theta\,\frac{d\Theta}{d\theta}\right) + \frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2} = -\lambda\sin^2\theta. \tag{4.28}$$

The left-hand side is a sum of a function of θ and a function of ϕ, so each term must equal a constant. Let

$$\frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2} = -m^2. \tag{4.29}$$

Step 3: Solve the azimuthal equation The ϕ-equation becomes

$$\frac{d^2\Phi}{d\phi^2} + m^2\Phi = 0, \tag{4.30}$$

with solutions $\Phi(\phi) = e^{im\phi}$. Single-valuedness under $\phi \mapsto \phi + 2\pi$ implies $m \in \mathbb{Z}$.

Step 4: Obtain the polar (associated Legendre) equation With the separation constant m^2, the θ-equation becomes

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\,\frac{d\Theta}{d\theta}\right) - \frac{m^2}{\sin^2\theta}\,\Theta = -\lambda\,\Theta. \tag{4.31}$$

Let $x = \cos\theta$ and define $u(x) = \Theta(\theta(x))$. Using $\frac{d}{d\theta} = -\sin\theta\,\frac{d}{dx}$, one obtains the associated Legendre differential equation

$$(1-x^2)\,u''(x) - 2x\,u'(x) + \left(\lambda - \frac{m^2}{1-x^2}\right)u(x) = 0. \tag{4.32}$$

Regularity of solutions at the poles $x = \pm 1$ forces the eigenvalues to take the discrete form

$$\lambda = l(l+1), \qquad l \in \{0,1,2,\dots\}, \qquad |m| \le l. \tag{4.33}$$

The corresponding regular solutions are (up to normalization) the associated Legendre functions $P_l^m(x)$.

Step 5: Assemble Y_l^m and normalize Putting the pieces together yields

$$Y_l^m(\theta,\phi) = N_{lm}\,P_l^m(\cos\theta)\,e^{im\phi}, \tag{4.34}$$

where N_{lm} is chosen so that $\{Y_l^m\}$ are orthonormal on S^2:

$$\int_0^{2\pi}\int_0^{\pi} Y_l^m(\theta,\phi)\,\overline{Y_{l'}^{m'}(\theta,\phi)}\,\sin\theta\,d\theta\,d\phi = \delta_{ll'}\,\delta_{mm'}. \tag{4.35}$$

A common convention is

$$N_{lm} = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}. \tag{4.36}$$