# A General Theory of Equivariant CNNs on Homogeneous Spaces

Taco Cohen[1], Mario Geiger[2], and Maurice Weiler[3]

[1] *Qualcomm AI Research*
[2] *PCSL Lab, EPFL*
[3] *QUVA Lab, University of Amsterdam*



**Abstract**

Group equivariant convolutional neural networks (G-CNNs) have recently emerged as a very effective model class for learning from signals in the context of known symmetries. A wide variety of equivariant layers has been proposed for signals on 2D and 3D Euclidean space, graphs, and the sphere, and it has become difficult to see how all of these methods are related, and how they may be generalized.

In this paper, we present a fairly general theory of equivariant convolutional networks. Convolutional feature spaces are described as *fields* over a homogeneous base space, such as the plane $\mathbb{R}^2$, sphere $S^2$ or a graph $\mathcal{G}$. The theory enables a systematic classification of all existing G-CNNs in terms of their group of symmetry, base space, and field type (e.g. scalar, vector, or tensor field, etc.).

In addition to this classification, we use Mackey theory to show that convolutions with equivariant kernels are the most general class of equivariant maps between such fields, thus establishing G-CNNs as a universal class of equivariant networks. The theory also explains how the space of equivariant kernels can be parameterized for learning, thereby simplifying the development of G-CNNs for new spaces and symmetries. Finally, the theory introduces a rich geometric semantics to learned feature spaces, thus improving interpretability of deep networks, and establishing a connection to central ideas in mathematics and physics.

## 1 Introduction

Through the use of convolution layers, Convolutional Neural Networks (CNNs) are able to exploit both the spatial locality of the input space as well as the translational symmetry that is inherent in many learning problems. Because convolutions are *translation equivariant* (a shift of the input leads to a shift of the output), convolution layers preserve the translation symmetry. This is important, because it means that further layers of the network can also exploit the symmetry.

| $G$ | $H$ | $G/H$ | $\rho$ | Reference |
|---|---|---|---|---|
| $\mathbb{Z}^2$ | $\{1\}$ | $\mathbb{Z}^2$ | regular | LeCun et al. [1990] |
| $p4, p4m$ | $C_4, D_4$ | $\mathbb{Z}^2$ | regular | Cohen and Welling [2016], |
| " | " | " | " | Dieleman et al. [2016] |
| $p4, p4m$ | $C_4, D_4$ | $\mathbb{Z}^2$ | irrep & regular | Cohen and Welling [2017] |
| $p6, p6m$ | $C_6, D_6$ | $\mathbb{H}^2$ | regular | Hoogeboom et al. [2018] |
| $\mathbb{Z}^3 \rtimes H$ | $D_4, D_{4h}, O, O_h$ | $\mathbb{Z}^3$ | regular | Winkels and Cohen [2018] |
| $\mathbb{Z}^3 \rtimes H$ | $V, T_4, O$ | $\mathbb{Z}^3$ | regular | Worrall and Brostow [2018] |
| SE(2) | SO(2) | $\mathbb{R}^2$ | regular | Weiler et al. [2018a] |
| " | " | " | " | Zhou et al. [2017] |
| " | " | " | " | Bekkers et al. [2018] |
| SE(2) | SO(2) | $\mathbb{R}^2$ | irrep | Worrall et al. [2017] |
| " | " | " | " | Marcos et al. [2017] |
| SE(3) | SO(3) | $\mathbb{R}^3$ | irrep | Kondor [2018] |
| " | " | " | irrep & regular | Thomas et al. [2018] |
| SO(3) | SO(2) | $S^2$ | regular | Cohen et al. [2018a] |
| " | | | trivial | Esteves et al. [2018] |
| $\mathbb{S}_n$ | $\mathbb{S}_k \times \mathbb{S}_{n-k}$ | $\mathbb{S}_n \ / \ \mathbb{S}_k \times \mathbb{S}_{n-k}$ | trivial | Kondor and Trivedi [2018] |

Table 1: A taxonomy of G-CNNs. Methods are classified by the group $G$ they are equivariant to, the subgroup $H$ that acts on the fibers, the base space $G/H$ to which the fibers are attached (implied by $G$ and $H$), and the type of field $\rho$ (regular, irreducible or trivial). These objects will be defined below.

Motivated by the success of CNNs, many researchers have worked on generalizations, leading to a growing body of work on *group equivariant networks* for signals on Euclidean space, the sphere, and graphs [Cohen and Welling, 2016, 2017, Worrall et al., 2017, Weiler et al., 2018a, Thomas et al., 2018, Kondor, 2018]. With the proliferation of equivariant layers, it has become difficult to see the relations between the various approaches, and this has led to a considerable amount of multiple discovery, often without the inventors being aware of it. Furthermore, when faced with a new modality (e.g. diffusion tensor MRI), it may not be immediately obvious how to create an equivariant network for it, and whether a given class of equivariant architectures is the most general one for a given modality.

In this paper we present a theory of equivariant convolutional networks that covers all of the existing models, as well as models that have not been developed yet. The theory describes convolutional feature spaces as spaces of fields over a homogeneous space, which enables a systematic classification of existing methods (Table 1). Moreover, we show (Theorem 6.1) that the most general equivariant linear map between fields (i.e. layer of the network) is a convolution / cross-correlation with an equivariant kernel. According to Theorems 6.2 – 6.4, the space of equivariant kernels can be described in three equivalent ways: as a space of operator-valued kernels on the group, quotient space or double quotient space, satisfying certain linear constraints. The latter two can be used to parameterize an equivariant kernel for use in a practical implementation of G-CNNs.

In order to give an elegant and general account of convolutional feature spaces (fields), we use the theory of fiber bundles. This theory formalizes the idea of parameterizing a set of spaces called the fibers (e.g. vector spaces of visual features) by another called the base space (e.g. a plane or sphere). We have made an effort to make these ideas accessible, but it may still be helpful to skip back and forth between the general theory and the concrete examples provided in Section 8.

This paper does not contain truly new mathematics (in the sense that a professional mathematician with expertise in the relevant subjects would not be surprised by our results), but instead provides a new formalism for the study of convolutional networks. This field theoretic framework allows us to systematically classify equivariant CNNs and derive special cases with ease. Moreover, by using this field theoretic language to describe generalized CNNs, we establish a bridge between deep learning and fundamental ideas in modern mathematics and physics. Indeed, almost all modern theories of physics are field theories, and so it is quite remarkable that the perceptual theories learned by convolutional networks can also be expressed in this framework.

## 1.1 Other Related Work

Besides the references in Table 1, several papers deserve special mention. Most closely related is the work of Kondor and Trivedi [2018], whose theory is analogous to ours, but only covers scalar fields (corresponding to using a trivial representation $\rho(h) = I$ in our theory). A proper treatment of general fields as we do here is more difficult, as it requires the use of fiber bundles. A framework for (non-convolutional) networks equivariant to finite groups was presented by Ravanbakhsh et al. [2017]. Our use of fields (with fibers transforming under a reduced representation) can be viewed as a formalization of convolutional capsules [Sabour et al., 2017, Hinton et al., 2018]. Other related work includes [Olah, 2014, Gens and Domingos, 2014, Sifre and Mallat, 2013, Oyallon and Mallat, 2015, Mallat, 2016, Koenderink, 1990, Koenderink and van Doorn, 2008, Petitot, 2003].

For mathematical background, we recommend Sharpe [1997], Marsh [2016], Folland [1995], Ceccherini-Silberstein et al. [2009], Gurarie [1992], Mackey [1951, 1952, 1953, 1968]. A preliminary version of this work appeared as Cohen et al. [2018b].

## 1.2 Limitations of the Theory

The theory presented here is quite general but still has several limitations. Firstly, we only cover fields over homogeneous spaces. Although fields can be defined over more general manifolds, and indeed there has been some effort aimed at defining convolutional networks on general (or Riemannian) manifolds [Bronstein et al., 2016], we restrict our attention to homogeneous spaces because they come naturally equipped with a group action to which the network can be made equivariant. A more general theory would not be able to make use of this additional structure, though our general framework provides a good starting point for further generalizations.

The theory only describes complete equivariance. For reasons of efficiency, one could also consider networks that are equivariant over a limited range of transformations, but this would make the theory significantly more complicated. We assume that both the group and its action on the input and hidden spaces is given as a hyperparameter. Learning the symmetry group and its representation is very interesting but beyond the scope of this paper. We also assume that the group acts linearly on the input and hidden spaces, though it should be noted that this does not imply linearity of the orbits. Assuming a linear action is not as limited as it sounds, since even very non-linear transformations such as diffeomorphisms act linearly on a function space or field (the feature space).

Finally, for reasons of mathematical elegance and simplicity, the theory idealizes feature maps as fields over a possibly continuous base space, but a computer implementation will usually involve discretizing this space. A similar approach is used in signal processing, where this step is justified by various sampling theorems and band-limit assumptions. It seems likely that a similar theory can be developed for deep networks, but this has not been done yet.

For readability and brevity, we will not aim to be fully rigorous; The purpose of this paper is to provide a map of the landscape, not a photorealistic picture.

In Sections 2 – 5 we present the mathematical framework. In Sec. 6 we present the main theorems on equivariance and convolution, Sec. 7 covers nonlinearities, and in Sec. 8 we provide examples.

# 2 Groups, Actions and Homogeneous Spaces

The set of transformations leaving some object invariant is called its *symmetry group*. Algebraically, such sets can be characterized by the axioms of *identity* (the identity $e$ is a symmetry), *closure* (the composition of two symmetries is a symmetry), and *inverses* (the inverse of a symmetry is a symmetry). In machine learning, we are interested in the symmetries of the learning problem. For instance, we may be able to transform an image without changing its label.

Groups can have additional structure, such as that of a topological space or differentiable manifold. Here we restrict our attention to locally compact topological groups, for which we can define a well-behaved integral, which is used to define the convolution [Folland, 1995].

Groups can *act* on other spaces. A group action is formally defined as a (well-behaved) map $\cdot : G \times X \to X$, where $G$ is a group and $X$ is the space acted on (which we assume to be locally compact). The map must satisfy $e \cdot x = x$ and $u(v \cdot x) = (uv) \cdot x$, where $u, v \in G$, $x \in X$, and $uv$ denotes the group operation. From here on, we will simply write $ux$ instead of $u \cdot x$.

The orbit of a point $x \in X$ is the set $O_x = \{gx \,|\, g \in G\}$. Each orbit is a homogeneous space, which means that for any $y, z \in O_x$, there exists $g \in G$ such that $gy = z$. It is easy to show that the orbits partition $X$, so any group action can be analyzed in terms of actions on homogeneous spaces. For this reason we assume the base space $X$ to be homogeneous.

Relative to an arbitrary origin $o \in X$, we can index other points in this homogeneous space using elements of $G$, as $x = go$. This addressing scheme is however not unique in general, since there can be multiple elements $g, g' \in G$ that both map $o$ to $x$. This ambiguity can be characterized by the stabilizer subgroup $H_x = \{g \in G \,|\, gx = x\}$. It can be shown that the stabilizers of all points in a homogeneous space are isomorphic, so we will simply denote it as $H$.

Another way to obtain a homogeneous space is by choosing a closed subgroup $H \leq G$, and then forming the quotient $G/H$, which is the set of *cosets* $gH = \{gh \,|h \in H\}$ for $g \in G$. The group acts on the quotient space via $u \cdot vH = (uv)H$, and this turns $G/H$ into a homogeneous space with stabilizer $H$. One can show that all homogeneous spaces arise in this way.

The other kind of action that we are interested in is a *linear group action* or *group representation* of $H$ on a vector space $V \simeq \mathbb{R}^n$ (called the canonical fiber). A representation of $H$ on $V$ is a homomorphism $\rho : H \to \mathrm{GL}(V)$, where $\mathrm{GL}(V)$ denotes the group of invertible linear maps from $V$ to $V$. Being a homomorphism means that for any $h, h' \in H$, we have $\rho(hh') = \rho(h)\rho(h')$. In our framework, $\rho$ describes the way that an individual fiber (e.g. a scalar, vector, tensor, or other geometric quantity) transforms under the action of the stabilizer $H$ on its receptive field.

# 3 Fiber Bundles

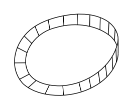

Figure 1: Möbius strip

Intuitively, a fiber bundle is a parameterization of a set of isomorphic spaces (the fibers) by another space (the base). The most familiar example is perhaps a convolutional feature space, which we can think of as a set of vector spaces $V_x \simeq \mathbb{R}^n$ (whose dimension is equal to the number of channels), one per position $x$ in the plane. This is an example of a *trivial bundle*, because it is simply the Cartesian product of the plane and a canonical fiber $\mathbb{R}^n$, but general fiber bundles are only *locally trivial*, meaning that they locally look like a product while having a different global structure. The classical example is the Mobius strip (Fig. 1), which locally looks like a product of the circle (the base) with a line segment (the fiber), but is globally distinct from a cylinder. A more practically relevant example is given by the tangent vector bundle of the sphere, see Section 8.

Formally, a bundle consists of topological spaces $E$ (total space), $B$ (base space), $F$ (fiber), and a projection map $p : E \to B$, satisfying the following local triviality condition: for every $a \in E$, there is an open neighbourhood $U$ of $p(a)$ and a trivializing homeomorphism $\varphi : p^{-1}(U) \to U \times F$ so that the map $p^{-1}(U) \xrightarrow{\varphi} U \times F \xrightarrow{\mathrm{proj}_1} U$ agrees with $p$ (where $\mathrm{proj}_1(u, f) = u$). Considering that the preimage ${\mathrm{proj}_1}^{-1}(x)$ for any $x \in U$ is $F$, and $\varphi$ is a homeomorphism, we see that the preimage $F_x = p^{-1}(x)$ for $x \in B$ is also homeomorphic to $F$. In other words, all fibers are homeomorphic. A bundle may be denoted as $E \xrightarrow{p} B$.

A *section* $s$ of a fiber bundle is an assignment to each $x \in B$ of an element $s(x) \in F_x$. Formally, it is a map $s : B \to E$ that satisfies $p \circ s = \mathrm{id}_B$. If the bundle is trivial, a section is essentially just a function $f : B \to F$, but for a non-trivial bundle we cannot align all the fibers in a canonical way, and so we must keep each $s(x)$ in its own fiber $F_x$. Nevertheless, on a trivializing neighbourhood $U \subseteq B$, we can describe the section as a function $f : U \to F$.

### 3.1 $G$ as a Principal $H$-Bundle

One example of a bundle, which we will use later on when constructing the feature space (Section 4) emerges naturally from the quotient $G/H$ that we discussed before. That is, the map $g \mapsto gH$ can serve as the projection $p : G \to G/H$ for a bundle that has $G$ as its total space, $G/H$ as its base, and $H$ as fiber. Intuitively, this allows us to think of $G$ as a base space $G/H$ with a twisted copy of $H$ attached at each point $x \in G/H$.

This bundle is called a *principal $H$-bundle*, because we have a transitive and fixed-point free group action $G \times H \to G$ that preserves the fibers. This action is given by right multiplication, $g \mapsto gh$, which preserves fibers because $ghH = gH$.

A section of the bundle $p : G \to G/H$ is a map $s : G/H \to G$ that satisfies $p \circ s = \mathrm{id}_{G/H}$. Since $p$ projects $g$ to its coset $gH$, the section chooses a representative $s(gH) \in gH$ for each coset $gH$. This bundle may not have continuous global sections, but we can always use a local section.

Aside from the right action of $H$, which turns $G$ into a principal $H$-bundle, we also have a left action of $G$ on itself, as well as an action of $G$ on the base space $G/H$. In general, the action of $G$ on $G/H$ does not agree with the action on $G$, in that $gs(x) \neq s(gx)$, because the action on $G$ includes a twist of the fiber. This twist is described by the function $\mathrm{h} : G/H \times G \to H$ defined by $gs(x) = s(gx)\mathrm{h}(x, g)$, which will play an important role in the quite general theory. We note that $\mathrm{h}$ satisfies the cocycle condition $\mathrm{h}(x, g_1 g_2) = \mathrm{h}(g_2 x, g_1)\mathrm{h}(x, g_2)$.

## 4 The Associated Vector Bundle

In the quite general theory, feature spaces are realized as spaces of sections of the associated vector bundle, which we will now define. In physics, a section of this bundle is simply called a field.

To define the associated vector bundle, we start with the principal $H$-bundle $G \xrightarrow{p} G/H$, and a group representation $\rho$ of $H$ on the vector space $V \simeq \mathbb{R}^n$, which will serve as the fiber. The representation $\rho$ describes the transformation behaviour of the features in $V$. These features could for instance transform as a scalar ($\rho(h) = 1$, the trivial representation), as a vector, a tensor, or some other geometrical quantity [Cohen and Welling, 2017].

First, we construct the product space $G \times V$. In the context of representation learning, we can think of an element $(g, v)$ of $G \times V$ as a feature vector $v \in V$ and an associated pose variable $g \in G$ that describes how the feature detector was steered to obtain $v$. If we apply a transformation $h \in H$ to $g$ and simultaneously apply its inverse to $v$, we get an equivalent element $(gh, \rho(h^{-1})v)$. So in order to create the associated bundle $P$, we take the quotient of $G \times V$ by this action: $P = G \times_\rho V = (G \times V)/H$. The projection $p_\rho : P \to G/H$ is defined as

$p_\rho([g, v]) = gH$, where $[g, v]$ denotes the equivalence class of $(g, v)$. Thus, the associated bundle has base $G/H$ and fiber $V$, meaning that locally it looks like $G/H \times V$. We note that this construction works for any principal $H$-bundle, not just $p : G \to G/H$, which suggests a direction for further generalization.

A field ("stack of feature maps") is a section of the associated bundle, meaning that it is a map $s : G/H \to P$ such that $\pi_\rho \circ s = \mathrm{id}_{G/H}$. We will refer to the space of sections of the associated vector bundle as $\mathcal{I}$. Concretely, we have two ways to encode a section: as functions $f : G \to V$ subject to a constraint, and as local functions from $U \subseteq G/H$ to $V$. We will now define both.

### 4.1 Sections as Mackey Functions

The construction of the associated bundle as a product $G \times V$ subject to an equivalence relation suggests a way to describe sections concretely: a section can be viewed as a function $f : G \to V$ subject to the equivariance condition

$$f(gh) = \rho(h^{-1})f(g). \tag{1}$$

Such functions are called Mackey functions. A linear combination of Mackey functions is a Mackey function, so they form a vector space, which we will refer to as $\mathcal{I}_G$. Mackey functions are easy to work with because they allow a concrete and global description of a field, but they give a redundant representation that is not suitable for computer implementation.

### 4.2 Local Sections as Functions on $G/H$

The associated bundle has base $G/H$ and fiber $V$, so locally, we can describe a section as an unconstrained function $f : U \to V$ where $U \subseteq G/H$ is a trivializing neighbourhood. We refer to the space of such sections as $\mathcal{I}_C$. Given such a local section, we can encode it as a Mackey function through the following lifting isomorphism $\Lambda : \mathcal{I}_C \to \mathcal{I}_G$:

$$\begin{aligned} [\Lambda f](g) &= \rho(\mathrm{h}(g)^{-1})f(gH), \\ [\Lambda^{-1} f'](x) &= f'(s(x)), \end{aligned} \tag{2}$$

where $\mathrm{h}(g) = \mathrm{h}(H, g) = s(gH)^{-1}g \in H$ and $s(x)$ is a coset representative for $x \in G/H$. This map is analogous to the lifting defined by Kondor and Trivedi [2018] for scalar fields (i.e. $\rho(h) = I$).

## 5 The Induced Representation

The induced representation $\pi = \mathrm{Ind}_H^G \rho$ describes the action of $G$ on fields. In $\mathcal{I}_G$, it is defined as:

$$[\pi_G(g)f](k) = f(g^{-1}k). \tag{3}$$

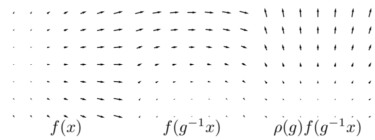


Figure 2: The rotation of a vector field in two steps: attaching each vector at the rotated position, and rotating the vectors themselves.

In $\mathcal{I}_C$, we can define the induced representation $\pi_C$ on a local neighbourhood $U$ as

$$[\pi_C(g)f](x) = \rho(\mathrm{h}(g^{-1}, x)^{-1})f(g^{-1}x). \tag{4}$$

One may verify, using the composition law for h, that this is indeed a representation of $G$. Moreover, one may verify that $\pi_G(g) \circ \Lambda = \Lambda \circ \pi_C(g)$, i.e. they define isomorphic representations.

Intuitively, we can interpret Eq. 4 as follows. To transform a field, we move the fiber at $g^{-1}x$ to $x$, and we apply a transformation to the fiber itself using $\rho$. This is visualized in Fig. 2. Some other examples include an RGB image ($\rho(h) = I_3$), a field of wind directions on earth ($\rho(h)$ a $2 \times 2$ rotation matrix), a diffusion tensor MRI image ($\rho(h)$ a representation of SO(3) acting on 2-tensors).

# 6 Equivariant Maps and Convolutions

Each feature space in a G-CNN is defined as the space of sections of some associated vector bundle, defined by a choice of base $G/H$ and representation $\rho$ of $H$ that describes how the fibers transform. A layer in a G-CNN is a map between these feature spaces that is equivariant to the induced representations acting on them. In this section we will show that such an equivariant linear map can always be written as a (twisted) convolution-like operation using an equivariant kernel. We will first derive this result for the induced representation realized in the space of Mackey functions, and then convert the result to local sections of the associated vector bundle in Section 6.2.

Consider adjacent feature spaces $i = 1, 2$ with a representation $(\rho_i, V_i)$ of $H_i \leq G$. Let $\pi_i = \mathrm{Ind}_{H_i}^G \rho_i$ be the representation acting on $\mathcal{I}_G^i$. A general linear map $\mathcal{I}_G^1 \to \mathcal{I}_G^2$ can be written as

$$[\kappa \cdot f](g) = \int_G \kappa(g,\, g')f(g')dg', \tag{5}$$

using a two-argument linear operator-valued kernel $\kappa : G \times G \to \mathrm{Hom}(V_1, V_2)$.

We are interested in the space of equivariant maps (intertwiners) between induced representations, $\mathcal{H} = \mathrm{Hom}_G(\mathcal{I}^1, \mathcal{I}^2) = \{\Phi \in \mathrm{Hom}(\mathcal{I}^1, \mathcal{I}^2) \,|\, \Phi\pi_1(g) = \pi_2(g)\Phi,\, \forall g \in G\}$. In order for Eq. 5 to define an equivariant map $\Phi \in \mathcal{H}$, the kernel $\kappa$ must satisfy a constraint. By (partially) resolving this constraint, we will show that Eq. 5 can always be written as a cross-correlation, and elucidate the structure of the space of equivariant kernels.

**Theorem 6.1.** *An equivariant map $\Phi \in \mathcal{H}$ can always be written as a cross-correlation.*

*Proof.* Since we are only interested in equivariant maps, we get a constraint on $\kappa$:

$$[\kappa \cdot [\pi_1(u) f]](g) = [\pi_2(u)[\kappa \cdot f]](g). \tag{6}$$

As shown in the supplementary material, this constraint is satisfied if and only if for all $g, g', u \in G$,

$$\kappa(ug, ug') = \kappa(g, g') \tag{7}$$

Hence, without loss of generality, we can define the two-argument kernel $\kappa(\cdot, \cdot)$ in terms of a one-argument kernel: $\kappa(g^{-1}g') \equiv \kappa(e, g^{-1}g') = \kappa(ge, gg^{-1}g') = \kappa(g, g')$.

The application of $\kappa$ to $f$ thus reduces to a cross-correlation:

$$[\kappa \cdot f](g) = \int_G \kappa(g, g') f(g') dg' = \int_G \kappa(g^{-1}g') f(g') dg' = [\kappa \star f](g). \tag{8}$$

$\square$

## 6.1 The Space of Equivariant Kernels

The constraint Eq. 7 implies a constraint on the one-argument kernel $\kappa$. The space of admissible kernels is in one-to-one correspondence with the space of equivariant maps. Here we give three different characterizations of this space of kernels. This knowledge can be used to construct parameterizations of the space of equivariant kernels for learning. Detailed proofs can be found in the supp. mat.

**Theorem 6.2.** *$\mathcal{H}$ is isomorphic to the space of bi-equivariant kernels on $G$, defined as:*

$$\begin{aligned} \mathcal{K}_G = \{\kappa : G \to \mathrm{Hom}(V_1, V_2) \,|\, \kappa(h_2 g h_1) = \rho_2(h_2)\kappa(g)\rho_1(h_1), \\ \forall g \in G, h_1 \in H_1, h_2 \in H_2\}. \end{aligned} \tag{9}$$

*Proof.* It is easily verified (see supp. mat.) that right equivariance follows from the fact that $f \in \mathcal{I}_G^1$ is a Mackey function, and left equivariance follows from the requirement that $\kappa \star f \in \mathcal{I}_G^2$ should be a Mackey function. The isomorphism is given by $\Gamma_G : \mathcal{K}_G \to \mathcal{H}$ defined as $[\Gamma_G \kappa] f = \kappa \star f$. $\square$

**Theorem 6.3.** *$\mathcal{H}$ is isomorphic to the space of left-equivariant kernels on $G/H_1$, defined as:*

$$\begin{aligned} \mathcal{K}_C = \{\overleftarrow{\kappa} : G/H_1 \to \mathrm{Hom}(V_1, V_2) \,|\, \overleftarrow{\kappa}(h_2 x) = \rho_2(h_2)\overleftarrow{\kappa}(x)\rho_1(\mathrm{h}_1(x, h_2)^{-1}), \\ \forall h_2 \in H_2, x \in G/H_1\} \end{aligned} \tag{10}$$

*Proof.* using the decomposition $g = s(gH_1)\mathrm{h}_1(g)$, we can define

$$\kappa(g) = \kappa(s(gH_1)\mathrm{h}_1(g)) = \kappa(s(gH_1))\,\rho_1(\mathrm{h}_1(g)) \equiv \overleftarrow{\kappa}(gH_1)\rho_1(\mathrm{h}_1(g)), \tag{11}$$

This defines the lifting isomorphism for kernels, $\Lambda_{\mathcal{K}} : \mathcal{K}_C \to \mathcal{K}_G$. It is easy to verify that when defined in this way, $\kappa$ satisfies right $H_1$-equivariance.

We still have the left $H_2$-equivariance constraint from Eq. 9, which translates to $\overleftarrow{\kappa}$ as follows (details in supp. mat.). For $g \in G$, $h_2 \in H_2$ and $x \in G/H_1$,

$$\kappa(h_2 g) = \rho_2(h_2)\kappa(g) \Leftrightarrow \overleftarrow{\kappa}(h_2 x) = \rho_2(h_2)\overleftarrow{\kappa}(x)\rho_1(\mathrm{h}_1(x, h_2)^{-1}). \tag{12}$$

□

**Theorem 6.4.** *$\mathcal{H}$ is isomorphic to the space of $H_2^{\gamma(x)H_1}$-equivariant kernels on $H_2\backslash G/H_1$:*

$$\begin{aligned}\mathcal{K}_D = \{\bar{\kappa} : H_2\backslash G/H_1 \to \mathrm{Hom}(V_1, V_2) \,|\, \bar{\kappa}(x) = \rho_2(h)\bar{\kappa}(x)\rho_1^x(h)^{-1},\\ \forall x \in H_2\backslash G/H_1, h \in H_2^{\gamma(x)H_1}\},\end{aligned} \tag{13}$$

*Where $\gamma : H_2\backslash G/H_1 \to G$ is a choice of double coset representatives, and $\rho_1^x$ is a representation of the stabilizer $H_2^{\gamma(x)H_1} = \{h \in H_2 \,|\, h\gamma(x)H_1 = \gamma(x)H_1\} \leq H_1$, defined as*

$$\rho_1^x(h) = \rho_1(\mathrm{h}_1(\gamma(x)H_1, h)) = \rho_1(\gamma(x)^{-1}h\gamma(x)), \tag{14}$$

*Proof.* In supplementary material. For examples, see Section 8. □

## 6.2 Local Sections on $G/H$

We have seen that an equivariant map between spaces of Mackey functions can always be realized as a cross-correlation on $G$, and we have studied the properties of the convolution kernel, which can be encoded as a kernel on $G$ or $G/H_1$ or $H_2\backslash G/H_1$, subject to the appropriate constraints. When implementing a G-CNN, it would be wasteful to use a Mackey function on $G$, so we need to understand what it means for fields realized by local functions $f : U \to V$ for $U \subseteq G/H_1$. This is done by conjugating the cross-correlation $\kappa\star : \mathcal{I}_G^1 \to \mathcal{I}_G^2$ by the lifting isomorphism $\Lambda_i : \mathcal{I}_C^i \to \mathcal{I}_G^i$.

$$\begin{aligned}[\Lambda_2^{-1}[\kappa \star [\Lambda_1 f]]](x) &= \int_G \kappa(s_2(x)^{-1}s_1(y))f(y)dy\\ &= \int_{G/H_1} \overleftarrow{\kappa}(s_2(x)^{-1}y)\rho_1(\mathrm{h}_1(s_2(x)^{-1}s_1(y)))f(y)dy\end{aligned} \tag{15}$$

Which we refer to as the $\rho_1$-twisted cross-correlation on $G/H_1$. We note that for semidirect product groups, the $\rho_1$ factor disappears and we are left with a standard cross-correlation on $G/H_1$ with an equivariant kernel $\overleftarrow{\kappa} \in \mathcal{K}_C$.

# 7 Equivariant Nonlinearities

The network as a whole is equivariant if all of its layers are equivariant. So our theory would not be complete without a discussion of equivariant nonlinearities and other kinds of layers. In a regular G-CNN [Cohen and Welling, 2016], $\rho$ is the regular representation of $H$, which means that it can be realized by permutation matrices. Since permutations and pointwise nonlinearities commute, any such nonlinearity can be used. For other kinds of representations $\rho$, special equivariant nonlinearities must be used. Some choices include norm nonlinearities [Worrall et al., 2017], tensor product nonlinearities [Kondor, 2018], or gated nonlinearities, where a scalar field gate is multiplied by an arbitrary field. Other constructions, such as batch norm and residual networks, can also be made equivariant [Cohen and Welling, 2016, 2017].

# 8 Concrete Examples

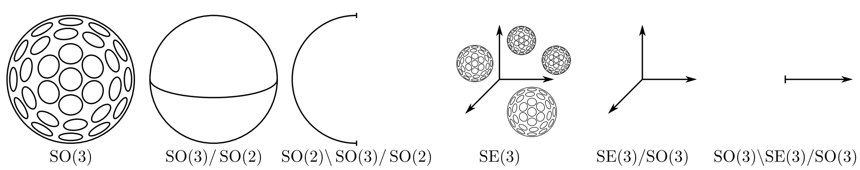


Figure 3: Quotients of SO(3) and SE(3).

## 8.1 The rotation group SO(3) and spherical CNNs

The group of 3D rotations SO(3) is a three-dimensional manifold that can be parameterized by ZYZ Euler angles $\alpha \in [0, 2\pi)$, $\beta \in [0, \pi]$ and $\gamma \in [0, 2\pi)$, i.e. $g = Z(\alpha)Y(\beta)Z(\gamma)$, (where $Z$ and $Y$ denote rotations around the Z and Y axes). We choose $H = H_1 = H_2 = \mathrm{SO}(2) = \{Z(\alpha) \,|\, \alpha \in [0, 2\pi)\}$ as the group of rotations around the Z-axis. A left $H$-coset is a set of the form $gH = \{Z(\alpha)Y(\beta)Z(\gamma)Z(\alpha') \,|\, \alpha' \in [0, 2\pi)\} = \{Z(\alpha)Y(\beta)Z(\alpha') \,|\, \alpha' \in [0, 2\pi)\}$. Thus, the coset space $G/H$ is the sphere $S^2$, parameterized by spherical coordinates $\alpha$ and $\beta$. As expected, the stabilizer $H_x$ of a point $x \in S^2$ is the set of rotations around the axis through $x$, which is isomorphic to $H = \mathrm{SO}(2)$ The orbit of a point $x = (\alpha, \beta) \in S^2$ under $H$ is a circle around the $Z$ axis at lattitude $\beta$, so the double coset space $H\backslash G/H$, which indexes these orbits, is $[0, \pi)$ (see Fig. 3).

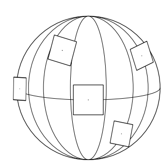

Figure 4: The tangent bundle of $S^2$

The section $s : G/H \to H$ may be defined (almost everywhere) as $s(\alpha, \beta) = Z(\alpha)Y(\beta)$, and $\gamma(\beta) = Y(\beta)$. Then the stabilizer $H_2^{\gamma(\beta)H_1}$ for $\beta \in H\backslash G/H$ is the set of Z-axis rotations that leave the point $\gamma(\beta)H_1 = (0, \beta) \in S^2$

invariant. For the north and south pole ($\beta = 0$ or $\beta = \pi$), this stabilizer is all of $H = \mathrm{SO}(2)$, but for other points it is the trivial subgroup $\{e\}$.

If we choose $\rho$ to be the standard representation of SO(2) in terms of $2 \times 2$ matrices, the associated bundle is the tangent vector bundle (Fig. 4). Equivariant kernels are functions on $S^2$ that depend only on lattitude $\beta$. These kernels are not constrained, because $H_2^{\gamma(x)H_1}$ is trivial (except at the poles).

### 8.2 The roto-translation group SE(3) and 3D Steerable CNNs

The group of rigid body motions SE(3) is a 6D manifold $\mathbb{R}^3 \rtimes \mathrm{SO}(3)$. We choose $H = H_1 = H_2 = \mathrm{SO}(3)$ (rotations around the origin). A left $H$-coset is a set of the form $gH = trH = \{trr' \,|\, r' \in \mathrm{SO}(3)\} = \{tr \,|\, r \in \mathrm{SO}(3)\}$. Thus, the coset space $G/H$ is $\mathbb{R}^3$. The stabilizer $H_x$ of a point $x \in \mathbb{R}^3$ is the set of rotations around $x$, which is isomorphic to SO(3). The orbit of a point $x \in \mathbb{R}^3$ is a spherical shell of radius $\|x\|$, so the double coset space $H\backslash G/H$ is the set of radii $[0, \infty)$.

Since SE(3) is a trivial bundle, we can choose a global section $s : G/H \to G$ by taking $s(x)$ to be translation by $x$, and $\gamma(\|x\|)$ to be the translation by $(0, 0, \|x\|)$. Then the stabilizer $H_2^{\gamma(\|x\|)H_1}$ for $\|x\| \in H\backslash G/H$ is the set of rotations around Z, i.e. SO(2), except for $\|x\| = 0$, where it is SO(3).

For any representations $\rho_1, \rho_2$, the equivaiant maps between sections of the associated vector bundle are given by convolutions with matrix-valued kernels on $\mathbb{R}^3$ that satisfy $\overleftarrow{\kappa}(rx) = \rho_2(r)\overleftarrow{\kappa}(x)\rho_1(r^{-1})$ for $r \in \mathrm{SO}(3)$ and $x \in \mathbb{R}^3$. Alternatively, we can define $\overleftarrow{\kappa}$ in terms of $\bar{\kappa}$, which is a kernel on $H\backslash G/H = [0, \infty)$ satisfying $\bar{\kappa}(x) = \rho_2(r)\bar{\kappa}(x)\rho_1(r)$ for $r \in \mathrm{SO}(2)$ and $x \in [0, \infty)$, [Weiler et al., 2018b].

# 9 Conclusion

In this paper we have developed the quite general theory of equivariant convolutional networks using the formalism of fiber bundles and fields. Field theories are the de facto standard formalism for modern physical theories, and this paper shows that the same formalism can elegantly describe the de facto standard learning machine: the convolutional network and its generalizations. By connecting this very successful class of networks to modern theories in mathematics and physics, our theory provides many opportunities for the development of new theoretical insights about deep learning, and the development of new equivariant network architectures.

## References


Taco S Cohen and Max Welling. Group equivariant convolutional networks. In *Proceedings of The 33rd International Conference on Machine Learning (ICML)*, volume 48, pages 2990–2999, 2016.

Taco S Cohen and Max Welling. Steerable CNNs. In *ICLR*, 2017.

Daniel E Worrall, Stephan J Garbin, Daniyar Turmukhambetov, and Gabriel J Brostow. Harmonic networks: Deep translation and rotation equivariance. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

Maurice Weiler, Fred A Hamprecht, and Martin Storath. Learning steerable filters for rotation equivariant CNNs. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018a.

Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. Tensor field networks: Rotation- and Translation-Equivariant neural networks for 3D point clouds. February 2018.

Risi Kondor. N-body networks: a covariant hierarchical neural network architecture for learning atomic potentials. March 2018.

Yann LeCun, Bernhard E Boser, John S Denker, Donnie Henderson, R E Howard, Wayne E Hubbard, and Lawrence D Jackel. Handwritten digit recognition with a Back-Propagation network. In D S Touretzky, editor, *Advances in Neural Information Processing Systems 2*, pages 396–404. Morgan-Kaufmann, 1990.

S Dieleman, J De Fauw, and K Kavukcuoglu. Exploiting cyclic symmetry in convolutional neural networks. In *International Conference on Machine Learning (ICML)*, 2016.

Emiel Hoogeboom, Jorn W T Peters, Taco S Cohen, and Max Welling. HexaConv. In *International Conference on Learning Representations (ICLR)*, 2018.

Marysia Winkels and Taco S Cohen. 3D G-CNNs for pulmonary nodule detection. April 2018.

Daniel Worrall and Gabriel Brostow. CubeNet: Equivariance to 3D rotation and translation. April 2018.

Yanzhao Zhou, Qixiang Ye, Qiang Qiu, and Jianbin Jiao. Oriented response networks. In *CVPR*, 2017.

Erik J Bekkers, Maxime W Lafarge, Mitko Veta, Koen A J Eppenhof, and Josien P W Pluim. Roto-Translation covariant convolutional networks for medical image analysis. April 2018.

Diego Marcos, Michele Volpi, Nikos Komodakis, and Devis Tuia. Rotation equivariant vector field networks. In *International Conference on Computer Vision (ICCV)*, 2017.

Taco S Cohen, Mario Geiger, Jonas Koehler, and Max Welling. Spherical CNNs. In *International Conference on Learning Representations (ICLR)*, 2018a.

Carlos Esteves, Christine Allen-Blanchette, Ameesh Makadia, and Kostas Daniilidis. 3D object classification and retrieval with spherical CNNs. 2018.

Risi Kondor and Shubhendu Trivedi. On the generalization of equivariance and convolution in neural networks to the action of compact groups. 2018.

Siamak Ravanbakhsh, Jeff Schneider, and Barnabas Poczos. Equivariance through Parameter-Sharing. February 2017.

Sara Sabour, Nicholas Frosst, and Geoffrey E Hinton. Dynamic routing between capsules. In I Guyon, U V Luxburg, S Bengio, H Wallach, R Fergus, S Vishwanathan, and R Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 3856–3866. Curran Associates, Inc., 2017.

Geoffrey Hinton, Nicholas Frosst, and Sara Sabour. Matrix capsules with EM routing. 2018.

Chris Olah. Groups and group convolutions. `https://colah.github.io/posts/2014-12-Groups-Convolution/`, 2014.

R Gens and P Domingos. Deep symmetry networks. In *Advances in Neural Information Processing Systems (NIPS)*, 2014.

Laurent Sifre and Stephane Mallat. Rotation, scaling and deformation invariant scattering for texture discrimination. *IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.

E Oyallon and S Mallat. Deep Roto-Translation scattering for object classification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2865–2873, 2015.

Stéphane Mallat. Understanding deep convolutional networks. *Philos. Trans. A Math. Phys. Eng. Sci.*, 374(2065):20150203, April 2016.

Jan J Koenderink. The brain a geometry engine. *Psychol. Res.*, 52(2-3):122–127, 1990.

Jan Koenderink and Andrea van Doorn. The structure of visual spaces. *J. Math. Imaging Vis.*, 31(2):171, April 2008.

Jean Petitot. The neurogeometry of pinwheels as a sub-riemannian contact structure. *J. Physiol. Paris*, 97(2-3):265–309, 2003.

R W Sharpe. *Differential Geometry: Cartan's Generalization of Klein's Erlangen Program*. 1997.

Adam Marsh. Gauge theories and fiber bundles: Definitions, pictures, and results. July 2016.

G B Folland. *A Course in Abstract Harmonic Analysis*. CRC Press, 1995.

T Ceccherini-Silberstein, A Machí, F Scarabotti, and F Tolli. Induced representations and mackey theory. *J. Math. Sci.*, 156(1):11–28, January 2009.

David Gurarie. *Symmetries and Laplacians: Introduction to Harmonic Analysis, Group Representations and Applications*. Elsevier B.V., 1992.

George W Mackey. On induced representations of groups. *Amer. J. Math.*, 73 (3):576–592, July 1951.

George W Mackey. Induced representations of locally compact groups I. *Ann. Math.*, 55(1):101–139, 1952.

George W Mackey. Induced representations of locally compact groups II. the frobenius reciprocity theorem. *Ann. Math.*, 58(2):193–221, 1953.

George W Mackey. *Induced Representations of Groups and Quantum Mechanics*. W.A. Benjamin Inc., New York-Amsterdam, 1968.

Taco S Cohen, Mario Geiger, and Maurice Weiler. Intertwiners between induced representations (with applications to the theory of equivariant neural networks). March 2018b.

Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. pages 1–22, November 2016.

Maurice Weiler, Mario Geiger, Max Welling, Wouter Boomsma, and Taco Cohen. 3D steerable CNNs: Learning rotationally equivariant features in volumetric data. July 2018b.

# Supplementary Material

## 1 General facts about Groups and Quotients

Let $G$ be a group and $H$ a subgroup of $G$. A left coset of $H$ in $G$ is a set $gH = \{gh \mid h \in H\}$ for $g \in G$. The cosets form a partition of $G$. The set of all cosets is called the quotient space or coset space, and is denoted $G/H$. There is a canonical projection $p : G \to G/H$ that assigns to each element $g$ the coset it is in. This can be written as $p(g) = gH$. Fig. 5 provides an illustration for the group of symmetries of a triangle, and the subgroup $H$ of reflections.

The quotient space carries a left action of $G$, which we denote with $ux$ for $u \in G$ and $x \in G/H$. This works fine because this action is associative with the group operation:

$$u(gH) = (ug)H. \tag{16}$$

for $u, g \in G$. One may verify that this action is well defined, i.e. does not depend on the particular coset representative $g$. Furthermore, the action is transitive, meaning that we can reach any coset from any other coset by transforming it with an appropriate $u \in G$. A space like $G/H$ on which $G$ acts transitively is called a homogeneous space for $G$. Indeed, any homogeneous space is isomorphic to some quotient space $G/H$.

A section of $p$ is a map $s : G/H \to G$ such that $p \circ s = \mathrm{id}_{G/H}$. We can think of $s$ as choosing a coset representative for each coset, i.e. $s(x) \in x$. In general, although $p$ is unique, $s$ is not; there can be many ways to choose coset representatives. However, the constructions we consider will always be independent of the particular choice of section.

Although it is not strictly necessary, we will assume that $s$ maps the coset $H = eH$ of the identity to the identity $e \in G$:

$$s(H) = e \tag{17}$$

We can always do this, for given a section $s'$ with $s'(H) = h \neq e$, we can define the section $s(x) = h^{-1}s'(hx)$ so that $s(H) = h^{-1}s'(hH) = h^{-1}s'(H) = h^{-1}h = e$. This is indeed a section, for $p(s(x)) = p(h^{-1}s'(hx)) = h^{-1}p(s'(hx)) = h^{-1}hx = x$ (where we used Eq. 16 which can be rewritten as $up(g) = p(ug)$).

One useful rule of calculation is

$$(gs(x))H = g(s(x)H) = gx = s(gx)H, \tag{18}$$

for $g \in G$ and $x \in G/H$. The projection onto $H$ is necessary, for in general $gs(x) \neq s(gx)$. These two terms are however related, through a function $\mathrm{h} : G/H \times G \to H$, defined as follows:

$$gs(x) = s(gx)\mathrm{h}(x, g) \tag{19}$$

That is,

$$\boxed{\mathrm{h}(x, g) = s(gx)^{-1}gs(x)}. \tag{20}$$

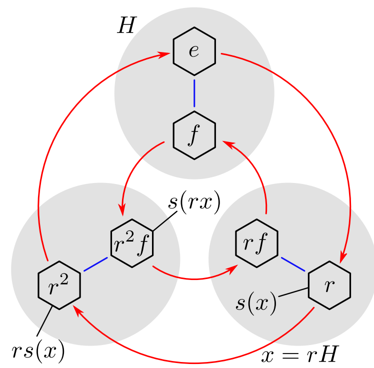


Figure 5: A Cayley diagram of the group $D3$ of symmetries of a triangle. The group is generated by rotations $r$ and flips $f$. The elements of the group are indicated by hexagons. The red arrows correspond to right multiplication by $r$, while the blue lines correspond to right multiplication by $f$. Cosets of the group of flips ($H = \{e, f\}$) are shaded in gray. As always, the cosets partition the group. As coset representatives, we choose $s(H) = e$, $s(rH) = r$, and $s(r^2H) = r^2f$. The difference between $s(rx)$ and $rs(x)$ is indicated. For this choice of section, we must set $\mathrm{h}(x, r) = \mathrm{h}(rH, r) = f$, so that $s(rx)\mathrm{h}(x, r) = (r^2f)(f) = r^2 = rs(x)$.

We can think of $\mathrm{h}(x, g)$ as the element of $H$ that we can apply to $s(gx)$ (on the right) to get $gs(x)$. The h function will play an important role in the definition of the induced representation, and is illustrated in Fig. 5.

From the fiber bundle perspective, we can interpret Eq. 20 as follows. The group $G$ can be viewed as a principal bundle with base space $G/H$ and fibers $gH$. If we apply $g$ to the coset representative $s(x)$, we move to a different coset, namely the one represented by $s(gx)$ (representing a different point in the base space). Additionally, the fiber is twisted by the right action of $\mathrm{h}(x, g)$. That is, $\mathrm{h}(x, g)$ moves $s(gx)$ to another element in its coset, namely to $gs(x)$.

The following composition rule for h is very useful in derivations:

$$\begin{aligned} \mathrm{h}(x, g_1g_2) &= s(g_1g_2x)^{-1}g_1g_2s(x) \\ &= [s(g_1g_2x)^{-1}g_1s(g_2x)][s(g_2x)^{-1}g_2s(x)] \\ &= \mathrm{h}(g_2x, g_1)\mathrm{h}(x, g_2) \end{aligned} \tag{21}$$

For elements $h \in H$, we find:

$$\mathrm{h}(H, h) = s(H)^{-1}hs(H) = h. \tag{22}$$

Also, for any coset $x$,

$$\mathrm{h}(H, s(x)) = s(s(x)H)^{-1}s(x)s(H) = s(H) = e. \tag{23}$$

Using Eq. 21 and 23, this yields,

$$\mathrm{h}(H, s(x)h) = \mathrm{h}(hH, s(x))\mathrm{h}(H, h) = h, \tag{24}$$

for any $h \in H$ and $x \in G/H$.

For $x = H$, Eq. 20 specializes to:

$$g = gs(H) = s(gH)\mathrm{h}(H, g) \equiv s(gH)\mathrm{h}(g), \tag{25}$$

where we defined

$$\boxed{\mathrm{h}(g) = \mathrm{h}(H, g) = s(gH)^{-1}g} \tag{26}$$

This shows that we can always factorize $g$ *uniquely* into a part $s(gH)$ that represents the coset of $g$, and a part $\mathrm{h}(g) \in H$ that tells us where $g$ is within the coset:

$$g = s(gH)\mathrm{h}(g) \tag{27}$$

A useful property of $\mathrm{h}(g)$ is that for any $h \in H$,

$$\mathrm{h}(gh) = s(ghH)^{-1}gh = s(gH)^{-1}gh = \mathrm{h}(g)h. \tag{28}$$

It is also easy to see that

$$\mathrm{h}(s(x)) = e. \tag{29}$$

When dealing with different subgroups $H_1$ and $H_2$ of $G$ (associated with the input and output space of an intertwiner), we will write $h_i$ for an element of $H_i$, $s_i : G/H_i \to G$, for the corresponding section, and $\mathrm{h}_i : G/H_i \times G \to H_i$ for the h-function (for $i = 1, 2$).

## 1.1 Double cosets

A $(H_2, H_1)$-double coset is a set of the form $H_2gH_1$ for $H_2, H_1$ subgroups of $G$. The space of $(H_2, H_1)$-double cosets is called $H_2\backslash G/H_1 \equiv \{H_2gH_1 \,|\, g \in G\}$. As with left cosets, we assume a section $\gamma : H_2\backslash G/H_1 \to G$ is given, satisfying $\gamma(H_2gH_1) \in H_2gH_1$.

The double coset space $H_2\backslash G/H_1$ can be understood as the space of $H_2$-orbits in $G/H_1$, that is, $H_2\backslash G/H_1 = \{H_2x | x \in G/H_1\}$. Note that although $G$ acts transitively on $G/H_1$ (meaning that there is only one $G$-orbit in $G/H_1$), the subgroup $H_2$ does not. Hence, the space $G/H_1$ splits into a number of disjoint orbits $H_2x$ (for $x = gH_1 \in G/H_1$), and these are precisely the double cosets $H_2gH_1$.

Of course, $H_2$ *does* act transitively within a single orbit $H_2x$, sending $x \mapsto h_2x$ (both of which are in $H_2x$, for $x \in G/H_1$). In general this action is not necessarily fixed point free which means that there may exist some $h_2 \in H_2$ which map the left cosets to themselves. These are exactly the elements in the stabilizer of $x = gH_1$, given by

$$\begin{aligned} H_2^x &= \{h \in H_2 \,|\, hx = x\} \\ &= \{h \in H_2 \,|\, hs_1(x)H_1 = s_1(x)H_1\} \\ &= \{h \in H_2 \,|\, hs_1(x) \in s_1(x)H_1\} \\ &= \{h \in H_2 \,|\, h \in s_1(x)H_1s_1(x)^{-1}\} \\ &= s_1(x)H_1s_1(x)^{-1} \cap H_2. \end{aligned} \tag{30}$$

Clearly, $H_2^x$ is a subgroup of $H_2$. Furthermore, $H_2^x$ is conjugate to (and hence isomorphic to) the subgroup $s_1(x)^{-1}H_2^x s_1(x) = H_1 \cap s_1(x)^{-1}H_2 s_1(x)$, which is a subgroup of $H_1$.

For double cosets $x \in H_2\backslash G/H_1$, we will overload the notation to $H_2^x \equiv H_2^{\gamma(x)H_1}$. Like the coset stabilizer, this double coset stabilizer can be expressed as

$$H_2^x = \gamma(x)H_1\gamma(x)^{-1} \cap H_2 \tag{31}$$

### 1.2 Semidirect products

For a semidirect product group $G$, such as $\mathrm{SE}(2) = \mathbb{R}^2 \rtimes \mathrm{SO}(2)$, some things simplify. Let $G = N \rtimes H$ where $H \leq G$ is a subgroup, $N \leq G$ is a normal subgroup and $N \cap H = \{e\}$. For every $g \in G$ there is a unique way of decomposing it into $nh$ where $n \in N$ and $h \in H$. Thus, the left $H$ coset of $g \in G$ depends only on the $N$ part of $g$:

$$gH = nhH = nH \tag{32}$$

It follows that for a semidirect product group, we can define the section so that it always outputs an element of $N \subseteq G$, instead of a general element of $G$. Specifically, we can set $s(gH) = s(nhH) = s(nH) = n$. It follows that $s(nx) = ns(x)\ \ \forall n \in N, x \in G/H$. This allow us to simplify expressions involving h:

$$\begin{aligned}
\mathrm{h}(x, g) &= s(gx)^{-1}gs(x) \\
&= s(gs(x)H)^{-1}gs(x) \\
&= s(\underbrace{gs(x)g^{-1}}_{\in N}gH)^{-1}gs(x) \\
&= \left(gs(x)g^{-1}\ s(gH)\right)^{-1} gs(x) \\
&= s(gH)^{-1}g \\
&= \mathrm{h}(g)
\end{aligned} \tag{33}$$

## 2 Haar measure

When we integrate over a group $G$, we will use the Haar measure, which is the essentially unique measure $dg$ that is invariant in the following sense:

$$\int_G f(g)dg = \int_G f(ug)dg \quad \forall u \in G. \tag{34}$$

Such measures always exist for locally compact groups, thus covering most cases of interest [Folland, 1995]. For discrete groups, the Haar measure is the counting measure, and integration can be understood as a discrete sum.

We can integrate over $G/H$ by using an integral over $G$,

$$\int_{G/H} f(x)dx = \int_G f(gH)dg. \tag{35}$$

# 3 Proofs

## 3.1 Equivariance ⇔ Convolution

Since we are only interested in equivariant maps, we get a constraint on $\kappa$:

$$\begin{aligned}
& & [\kappa \cdot [\pi_1(u)f]](g) &= \pi_2(u)[\kappa \cdot f](g) \\
&\Leftrightarrow & \int_G \kappa(g, g')f(u^{-1}g')dg' &= \int_G \kappa(u^{-1}g, g')f(g')dg' \\
&\Leftrightarrow & \int_G \kappa(g, ug')f(g')dg' &= \int_G \kappa(u^{-1}g, g')f(g')dg' \\
&\Leftrightarrow & \kappa(g, ug') &= \kappa(u^{-1}g, g') \\
&\Leftrightarrow & \kappa(ug, ug') &= \kappa(g, g')
\end{aligned} \tag{36}$$

Hence, without loss of generality, we can define the two-argument kernel $\kappa(\cdot, \cdot)$ in terms of a one-argument kernel:

$$\kappa(g^{-1}g') \equiv \kappa(e, g^{-1}g') = \kappa(ge, gg^{-1}g') = \kappa(g, g'). \tag{37}$$

The application of $\kappa$ to $f$ reduces to a cross-correlation:

$$[\kappa \star f](g) = \int_G \kappa(g^{-1}g')f(g')dg' = [\kappa \cdot f](g). \tag{38}$$

## 3.2 Bi-equivariance of one-argument kernels on $G$

### 3.2.1 Left equivariance of $\kappa$

We want the result $\kappa \star f$ (or $\kappa \cdot f$) to live in $\mathcal{I}_G^2$, which means that this function has to satisfy the Mackey condition,

$$\begin{aligned}
& & [\kappa \star f](gh_2) &= \rho_2(h_2^{-1})[\kappa \star f](g) \\
&\Leftrightarrow & \int_G \kappa((gh_2)^{-1}g')f(g')dg' &= \rho_2(h_2^{-1}) \int_G \kappa(g^{-1}g')f(g')dg' \\
&\Leftrightarrow & \kappa(h_2^{-1}g^{-1}g') &= \rho_2(h_2^{-1})\kappa(g^{-1}g') \\
&\Leftrightarrow & \kappa(h_2 g) &= \rho_2(h_2)\kappa(g)
\end{aligned} \tag{39}$$

for all $h_2 \in H_2$ and $g \in G$.

### 3.2.2 Right equivariance of $\kappa$

The fact that $f \in \mathcal{I}_G^1$ satisfies the Mackey condition ($f(gh) = \rho_1(h)f(g)$ for $h \in H_1$) implies a symmetry in the correlation $\kappa \star f$. That is, if we apply a right-$H_1$-shift to the kernel, i.e. $[R_h\kappa](g) = \kappa(gh)$, we find that

$$\begin{aligned} [[R_h\kappa] \star f](g) &= \int_G \kappa(g^{-1}uh)f(u)du \\ &= \int_G \kappa(g^{-1}u)f(uh^{-1})du \\ &= \int_G \kappa(g^{-1}u)\rho_1(h)f(u)du. \end{aligned} \tag{40}$$

It follows that we can take (for $h \in H_1$),

$$\kappa(gh) = \kappa(g)\rho_1(h). \tag{41}$$

## 3.3 Kernels on $H_2\backslash G/H_1$

We have seen the space $\mathcal{K}_C$ of $H_2$-equivariant kernels on $G/H_1$ appear in our analysis of both $\mathcal{I}_G$ and $\mathcal{I}_C$. Kernels in this space have to satisfy the constraint (for $h \in H_2$):

$$\overleftarrow{\kappa}(hy) = \rho_2(h)\overleftarrow{\kappa}(y)\rho_1(\mathrm{h}_1(y,h)^{-1}) \tag{42}$$

Here we will show that this space is equivalent to the space

$$\boxed{\begin{aligned} \mathcal{K}_D = \{\bar{\kappa} : H_2\backslash G/H_1 \to \mathrm{Hom}(V_1, V_2) \,|\, \bar{\kappa}(x) = \rho_2(h)\bar{\kappa}(x)\rho_1^x(h)^{-1}, \\ \forall x \in H_2\backslash G/H_1, h \in H_2^{\gamma(x)H_1}\}, \end{aligned}} \tag{43}$$

where we defined the representation $\rho_1^x$ of the stabilizer $H_2^{\gamma(x)H_1}$,

$$\begin{aligned} \rho_1^x(h) &= \rho_1(\mathrm{h}_1(\gamma(x)H_1, h)) \\ &= \rho_1(\gamma(x)^{-1}h\gamma(x)), \end{aligned} \tag{44}$$

with the section $\gamma : H_2\backslash G/H_1 \to G$ being defined as in section 1.1. To show the equivalence of $\mathcal{K}_C$ and $\mathcal{K}_D$, we define an ismorphism $\Omega_\mathcal{K} : \mathcal{K}_D \to \mathcal{K}_C$. We begin by defining $\Omega_\mathcal{K}^{-1}$:

$$\bar{\kappa}(x) = [\Omega_\mathcal{K}^{-1}\overleftarrow{\kappa}](x) = \overleftarrow{\kappa}(\gamma(x)H_1). \tag{45}$$

We verify that for $\overleftarrow{\kappa} \in \mathcal{K}_C$ we have $\bar{\kappa} \in \mathcal{K}_D$. Let $h \in H_2^{\gamma(x)H_1}$, then

$$\begin{aligned} \bar{\kappa}(x) &= \overleftarrow{\kappa}(\gamma(x)H_1) \\ &= \overleftarrow{\kappa}(h\gamma(x)H_1) \\ &= \rho_2(h)\overleftarrow{\kappa}(\gamma(x)H_1)\rho_1(\mathrm{h}_1(\gamma(x)H_1, h))^{-1} \\ &= \rho_2(h)\bar{\kappa}(x)\rho_1^x(h)^{-1} \end{aligned} \tag{46}$$

To define $\Omega_\mathcal{K}$, we use the decomposition $y = h\gamma(H_2y)H_1$ for $y \in G/H_1$ and $h \in H_2$. Note that $h$ may not be unique, because $H_2$ does not in general act freely on $G/H_1$.

$$\overleftarrow{\kappa}(y) = [\Omega_\mathcal{K}\bar{\kappa}](y) = [\Omega_\mathcal{K}\bar{\kappa}](h\gamma(H_2y)H_1) = \rho_2(h)\bar{\kappa}(H_2y)\rho_1(\mathrm{h}_1(\gamma(H_2y)H_1, h))^{-1}. \tag{47}$$

We verify that for $\bar{\kappa} \in \mathcal{K}_D$ we have $\overleftarrow{\kappa} \in \mathcal{K}_C$.

$$\begin{aligned}
\overleftarrow{\kappa}(h'y) &= \overleftarrow{\kappa}(h'h\gamma(H_2y)H_1) \\
&= \rho_2(h'h)\bar{\kappa}(H_2y)\rho_1(\mathrm{h}_1(\gamma(H_2y)H_1, h'h))^{-1} \\
&= \rho_2(h'h)\bar{\kappa}(H_2y)\rho_1(\mathrm{h}_1(h\gamma(H_2y)H_1, h')\mathrm{h}_1(\gamma(H_2y)H_1, h))^{-1} \\
&= \rho_2(h')\rho_2(h)\bar{\kappa}(H_2y)\rho_1(\mathrm{h}_1(\gamma(H_2y)H_1, h))^{-1}\rho_1(\mathrm{h}_1(h\gamma(H_2y)H_1, h'))^{-1} \\
&= \rho_2(h')\rho_2(h)\bar{\kappa}(H_2y)\rho_1(\mathrm{h}_1(\gamma(H_2y)H_1, h))^{-1}\rho_1(\mathrm{h}_1(y, h'))^{-1} \\
&= \rho_2(h')\overleftarrow{\kappa}(y)\rho_1(\mathrm{h}_1(y, h'))^{-1}
\end{aligned} \tag{48}$$

We verify that $\Omega_{\mathcal{K}}$ and $\Omega_{\mathcal{K}}^{-1}$ are indeed inverses:

$$\begin{aligned}
[\Omega_{\mathcal{K}}[\Omega_{\mathcal{K}}^{-1}\overleftarrow{\kappa}]](y) &= [\Omega_{\mathcal{K}}[\Omega_{\mathcal{K}}^{-1}\overleftarrow{\kappa}]](h\gamma(H_2y)H_1) \\
&= \rho_2(h)[\Omega_{\mathcal{K}}^{-1}\overleftarrow{\kappa}](H_2y)\rho_1(\mathrm{h}_1(\gamma(H_2y)H_1, h))^{-1} \\
&= \rho_2(h)\overleftarrow{\kappa}(\gamma(H_2y)H_1)\rho_1(\mathrm{h}_1(\gamma(H_2y)H_1, h))^{-1} \\
&= \overleftarrow{\kappa}(h\gamma(H_2y)H_1) \\
&= \overleftarrow{\kappa}(y).
\end{aligned} \tag{49}$$

In the other direction,

$$\begin{aligned}
[\Omega_{\mathcal{K}}^{-1}[\Omega_{\mathcal{K}}\bar{\kappa}]](x) &= [\Omega_{\mathcal{K}}\bar{\kappa}](\gamma(x)H_1) \\
&= [\Omega_{\mathcal{K}}\bar{\kappa}](\gamma(H_2\gamma(x)H_1)H_1) \\
&= \rho_2(e)\bar{\kappa}(H_2\gamma(x)H_1)\rho_1(\mathrm{h}_1(\gamma(H_2\gamma(x)H_1)H_1, e))^{-1} \\
&= \bar{\kappa}(x)
\end{aligned} \tag{50}$$