

HOMOGENEOUS VECTOR BUNDLES AND G -EQUIVARIANT CONVOLUTIONAL NEURAL NETWORKS

JIMMY ARONSSON

ABSTRACT. G -equivariant convolutional neural networks (GCNNs) is a geometric deep learning model for data defined on a homogeneous G -space \mathcal{M} . GCNNs are designed to respect the global symmetry in \mathcal{M} , thereby facilitating learning. In this paper, we analyze GCNNs on homogeneous spaces $\mathcal{M} = G/K$ in the case of unimodular Lie groups G and compact subgroups $K \leq G$. We demonstrate that homogeneous vector bundles is the natural setting for GCNNs. We also use reproducing kernel Hilbert spaces to obtain a precise criterion for expressing G -equivariant layers as convolutional layers. This criterion is then rephrased as a bandwidth criterion, leading to even stronger results for some groups.

CONTENTS

1. Introduction	1
2. Foundations of equivariant neural networks	3
2.1. Convolutional neural networks	3
2.2. Gauge theory and the equivariant framework	4
3. G -equivariant convolutional neural networks	9
3.1. Homogeneous spaces	9
3.2. Homogeneous vector bundles	9
3.3. Induced representations	10
3.4. G -equivariant and convolutional layers	12
3.5. RKHS and bandlimited functions	17
4. Discussion	20
References	21

1. INTRODUCTION

Developments in deep learning have increased dramatically in recent years. Even though multilayer perceptrons [2] and other general-architecture models work well for some tasks, achieving higher levels of performance often requires models that are more tailored to each application, and which incorporate some level of understanding of the data. *Geometric deep learning* [5, 6, 7, 13, 38] is the approach of using inherent geometric structure in data, and symmetry derived from geometry, to improve deep learning models.

Convolutional neural networks (CNNs) are among the simplest and most broadly applicable general-architecture models. They have been successfully applied to image classification and segmentation [41, 52, 53], text summarization [42], pose estimation [37], sign language recognition [27], and many other tasks. One reason why CNNs are so useful is that *convolutional layers*, the basic building blocks of CNNs, commute

Date: May 12, 2021.

with the translation operator in \mathbb{Z}^2 ; convolutional layers are *translation equivariant*. In image classification tasks, for instance, \mathbb{Z}^2 represents the underlying pixel lattice, and translation equivariance helps CNNs identify objects in images regardless of their exact pixel coordinates. As convolutional layers respect the global translation symmetry in \mathbb{Z}^2 , CNNs are examples of geometric deep learning models.

G-equivariant convolutional neural networks (GCNNs) [10, 12] are generalizations of CNNs to data points defined on homogeneous G -spaces \mathcal{M} . Convolutional layers that commute with the action $G \times \mathcal{M} \rightarrow \mathcal{M}$ of the global symmetry group G , remove the need for GCNNs to learn about the global symmetry. It is already built into the network. This enables GCNNs to focus on learning other relevant features in data, potentially improving performance. One example is the detection of tumors in digital pathology. Images of tumors can have any orientation, and GCNNs with both translation and rotation equivariant layers have higher accuracy than ordinary CNNs [46]. Rotation equivariance is also highly useful in 3D inference problems [49], in point cloud recognition [35], and in other tasks.

Gauge equivariant neural networks [9, 13, 18, 36] are instead designed to respect *local* symmetries. For example, computations involving vector fields - in meteorology or other areas - require vectors to be expressed in components. This requires a frame; a smooth assignment of a basis to each tangent space. However, the sphere and other non-parallelizable manifolds do not admit a global frame, so the computations must be performed locally, using different local frames for different regions on the manifold. It is then important that any numerical results obtained in one frame are compatible with those obtained in any other frame on overlapping regions. In other words, the computations should be equivariant with respect to the choice of local frame, which is viewed as a gauge degree of freedom; a local symmetry. Gauge equivariant neural networks have also been introduced for problems exhibiting other local symmetries, primarily in lattice gauge theory.

In this paper, we study the mathematical foundations of GCNNs and characterize convolutional layers in terms of more abstract layers. Our contributions are threefold:

- We analyze a general framework that include both gauge equivariant neural networks and GCNNs, that only differ in whether layers respect a local gauge symmetry or a global translation symmetry. Moreover, we show that GCNNs are naturally expressed in terms of homogeneous vector bundles.
- In general, not all G -equivariant layers can be written as convolutional layers. We investigate the relation between these types of layers for all homogeneous spaces $\mathcal{M} = G/K$ when G is a unimodular Lie group and $K \leq G$ is a compact subgroup. As a result of this investigation, we find a criterion for expressing G -equivariant layers as convolutional layers (Theorem 14).
- We highlight the close relationship between convolutional layers in GCNNs, reproducing kernel Hilbert spaces (RKHS), and bandwidth. We reformulate the criterion in Theorem 14 as a bandwidth criterion and prove that, when G is discrete abelian or finite,¹ all G -equivariant layers are indeed convolutional layers (Corollaries 19-21).

This work was inspired by a number of papers [9, 10, 11, 12, 13, 48]. The theoretical papers [9, 12] have been of particular importance, as our work grew from a desire to understand the mathematics of equivariant neural networks in even greater detail.

In the case of compact groups G , the Peter-Weyl theorem and other powerful tools have allowed researchers to study GCNNs using harmonic analysis. Among the most

¹Discrete Lie groups are countable, in this paper, as we assume smooth manifolds to be Hausdorff and second-countable.

well-known results in this direction is Theorem 1 in [28], which uses Fourier analysis on G to establish that the layers in a G -equivariant feed-forward neural network must be generalized convolutional layers, when G is compact. This result is similar to our second contribution above and we discuss the distinction in Section 3.4. Others have used the well-known representation theory of the compact group $G = SO(3)$ to study rotation equivariant GCNNs for spherical data [15, 16, 17].

The paper is structured as follows. We summarize the relevant machine learning background in Section 2.1, and discuss a framework for equivariant neural networks in Section 2.2. In Section 3, we restrict attention to homogeneous spaces G/K where G is a unimodular Lie group and $K \leq G$ is a compact subgroup. Section 3.2 explains the relation between GCNNs, homogeneous vector bundles, and induced representations. This relation is used to motivate the definition of G -equivariant layers in Section 3.4, where we also discuss convolutional layers and prove the aforementioned Theorem 14; this result characterizes when a G -equivariant layer is a convolutional layer, in terms of RKHS. Section 3.5 then relates RKHS to bandwidth, leading to a reformulation of Theorem 14 (Corollary 19) as well as a few stronger results. Finally, in Section 4, we summarize our work and end with a discussion.

2. FOUNDATIONS OF EQUIVARIANT NEURAL NETWORKS

In this section, we give an introduction to convolutional neural networks (CNNs) and discuss a simple framework for equivariant neural networks.

2.1. Convolutional neural networks. CNNs were first introduced in 1979 under the name of Neocognitrons, and were used to study visual pattern recognition [21]. In the 1990s, CNNs were successfully applied to problems such as automatic recognition of handwritten digits [33] and face recognition [32]. However, it was arguably not until 2012, when the GPU-based AlexNet CNN outperformed all competition on the ImageNet Large Scale Visual Recognition Challenge [30], that CNNs and other neural networks truly caught the public eye. Industrial work and academic research on deep learning has since soared, and current state-of-the-art deep learning architectures are significantly more powerful and more complex than AlexNet. Yet, convolutional layers remain important components.

In this introduction, we focus on data that can be represented by finitely supported functions

$$f : \mathbb{Z}^2 \rightarrow \mathbb{R}^m. \quad (2.1)$$

Digital images, for example, are of this form since each pixel $x \in \mathbb{Z}^2$ is associated with a color array $f(x) \in \mathbb{R}^m$, and finite support is analogous to finite image resolution. Note that $m = 1$ corresponds to grayscale images and $m = 3$ to RGB images, but we allow any number of channels m . In general, any data represented by a finite 2D ($m = 1$) or 3D ($m > 1$) array with real-valued entries is of the form (2.1).

Convolutional layers act on data points (2.1) by²

$$[\kappa \star f](x) = \sum_{y \in \mathbb{Z}^2} \kappa(y - x) f(y), \quad (2.2)$$

given a matrix-valued kernel $\kappa : \mathbb{Z}^2 \rightarrow \text{Hom}(\mathbb{R}^m, \mathbb{R}^n)$ for some $n \in \mathbb{N}$. The kernel is also finitely supported in practice, so the maps $\kappa \star f : \mathbb{Z}^2 \rightarrow \mathbb{R}^n$ are themselves data points (2.1) with n channels. Broadly speaking, CNNs consist of convolutional layers (2.2) combined with other transformations, such as non-linear activation functions and batch normalization layers. We are mainly interested in convolutional layers,

²The name *convolutional layer* is used even though (2.4) more closely resembles a cross-correlation. It can be expressed as a convolution if we replace the kernel with its involution $\kappa^*(y) = \kappa(-y)$.

so we do not go into detail about non-linear activation functions or other types of layers. For more extensive descriptions of CNNs, see [1, 24, 51].

In image classification tasks, for instance, CNNs categorize digital images into a predefined number k of distinct classes, based on what the images depict. The CNN maps each digital image $f : \mathbb{Z}^2 \rightarrow \mathbb{R}^m$ to a probability vector in \mathbb{R}^k estimating the probability that f belongs to any given class. During training, this probability vector is compared to the correct answer (which is known) and the discrepancy is computed using a *loss* norm or distance function. A gradient descent-based algorithm minimizes the loss function, thereby *learning* the kernel matrix elements and any other trainable network parameters. The result of this training procedure is a CNN that accurately classifies images in the training data set. Finally, the predictive power of the CNN is evaluated by using it to classify images from a test data set; images that were not used during training and which the CNN has not encountered before.

CNNs perform very well on image classification and similar machine learning tasks, and are important parts of many state-of-the-art network architectures on such tasks [4, 25, 44, 50]. One reason for their success is *translation equivariance*: Convolutional layers (2.2) commute with the translation operator in the image plane,

$$L_x : \mathbb{Z}^2 \rightarrow \mathbb{Z}^2, \quad L_x(y) = y + x, \quad x \in \mathbb{Z}^2. \quad (2.3)$$

Translation equivariance makes CNNs agnostic to the specific locations of individual pixels, while still taking into account the relative positions of different pixels; images are more easily classified based on relevant features of their subjects, and not based on technical artifacts such as specific pixel coordinates. This observation motivates the introduction of more general convolutional layers that act equivariantly on data points $f : \mathcal{M} \rightarrow V$, where the domain \mathcal{M} is homogeneous with respect to a locally compact group G [10, 12]. Given finite-dimensional vector spaces V, W , convolutional layers are defined as certain vector-valued integrals³

$$\kappa \star f : \mathcal{M} \rightarrow W, \quad (\kappa \star f)(g) = \int_G \kappa(g^{-1}g')f(g') \, dg', \quad (2.4)$$

with operator-valued kernels $\kappa : G \rightarrow \text{Hom}(V, W)$.

Remark 1. In (2.4), we integrate with respect to a Haar measure on the unimodular Lie group G .

Broadly speaking, G -equivariant convolutional neural networks (GCNNs) consist of sequences of convolutional layers (2.4) mixed with non-linear activation functions, and possibly other layers that are equivariant with respect to the global symmetry. This characterization is intentionally vague as we want to avoid making unnecessarily restrictive assumptions on the layers. For this reason, we will not study GCNNs from a holistic perspective, as a sequence of multiple layers, but instead focus on individual layers. We give a formal definition of abstract, G -equivariant layers in Definition 9, before defining the more specific convolutional layers in Definition 10.

2.2. Gauge theory and the equivariant framework. Before going into detail about GCNNs in Section 3, let us describe a mathematical framework for equivariant neural networks. The framework is based on gauge theoretic concepts but is equally suitable for GCNNs. Both gauge equivariant neural networks and GCNNs will thus be described by this framework, their main difference being the specific equivariance properties imposed on layers.

³For a summary on vector-valued integration on locally compact groups, see [19, Appendix 4].

Remark 2. This framework is already being used in GCNNs and gauge equivariant neural networks separately [9, 12]. We are simply presenting the unified theory that includes both types of equivariance as separate cases.

Gauge theory originated in physics as a way to model local symmetry. In quantum electrodynamics (QED), for example, the electron wave function can be locally phase shifted, $\psi \mapsto e^{i\alpha}\psi$, with no physically observable consequence, and so QED is said to possess a $U(1)$ gauge symmetry. Mathematicians have later adopted gauge theory in order to study other types of local symmetries. The introduction of gauge equivariant deep learning models has been suggested by deep learning practitioners and physicists alike. For example, [9] investigates the structure of gauge equivariant layers used for vector fields, tensor fields, and more general fields. Physicists have introduced gauge equivariant neural networks for applications in, e.g., lattice gauge theory [3, 18, 36].

We assume some familiarity with fiber bundles,⁴ but we still present a few relevant definitions and examples.

Definition 1. Let K be a Lie group. A smooth fiber bundle $\pi : P \rightarrow \mathcal{M}$ is called a *principal K -bundle* with *structure group K* if there is a free, smooth right K -action

$$P \times K \rightarrow P, \quad (p, k) \mapsto p \triangleleft k, \quad (2.5)$$

with the following properties for each $x \in \mathcal{M}$.

- (i) Let $P_x = \pi^{-1}(\{x\})$ be the fiber at x . Then

$$p \in P_x, k \in K \quad \Rightarrow \quad p \triangleleft k \in P_x. \quad (2.6)$$

That is, the K -action preserves fibers.

- (ii) For each $p \in P_x$, the mapping $k \mapsto p \triangleleft k$ is a diffeomorphism $K \rightarrow P_x$.

Principal bundles are natural tools for understanding local symmetries, i.e., gauge degrees of freedom. In theoretical physics, gauge degrees of freedom are redundancies in the mathematical theory with no physical relevance. This is both a blessing and a curse: Solving the Yang-Mills equations of motion as an initial value problem, for example, is an underdetermined problem that cannot be solved without taking the gauge degrees of freedom into account; without choosing a *gauge* [39]. This is similar to our example in the introduction, that computations involving vector fields may require a choice of basis in each tangent space, even if this choice is irrelevant for the underlying application. On the other hand, problems may also become easier to solve by choosing a gauge with some finesse.⁵

Definition 2. Let $\pi : P \rightarrow \mathcal{M}$ be a principal K -bundle and assume $U \subseteq \mathcal{M}$ is open.

- (i) A *gauge* is a local section $\omega : U \rightarrow P$.
(ii) A *gauge transformation* is an automorphism $\chi : P \rightarrow P$ that is equivariant,

$$\chi(p \triangleleft k) = \chi(p) \triangleleft k, \quad p \in P, k \in K, \quad (2.7)$$

and which preserves fibers: $\pi \circ \chi = \pi$.

We will go into more detail about the vector field example in Example 1. However, we first need to define associated bundles. To this end, let $\pi : P \rightarrow \mathcal{M}$ be a principal K -bundle and let $\rho : K \rightarrow GL(V_\rho)$ be a finite-dimensional representation. Define an equivalence relation \sim on $P \times V_\rho$ by

$$(p, v) \sim (p \triangleleft k, \rho(k)^{-1}v), \quad p \in P, v \in V_\rho, k \in K. \quad (2.8)$$

⁴Introductions to fiber bundles can be found in [26, 34, 40].

⁵See, for example, the temporal gauge in lattice gauge theory [22, §3.3.2].

Let $P \times_{\rho} V_{\rho} = (P \times V_{\rho}) / \sim$ denote the quotient space, whose elements are equivalence classes

$$[p, v] = [p \triangleleft k, \rho(k)^{-1}v], \quad p \in P, v \in V_{\rho}, k \in K, \quad (2.9)$$

and consider the projection $\pi_{\rho} : P \times_{\rho} V_{\rho} \rightarrow \mathcal{M}$ defined by $\pi_{\rho}([p, v]) = \pi(p)$. Observe that each fiber $\pi_{\rho}^{-1}(\{x\})$ has a natural vector space structure such that the mapping

$$V_{\rho} \rightarrow \pi_{\rho}^{-1}(\{x\}), \quad v \mapsto [p, v], \quad (2.10)$$

is a linear isomorphism for each fixed $p \in P_x$.

Lemma 1 ([26, §10.7]). *The associated bundle $\pi_{\rho} : P \times_{\rho} V_{\rho} \rightarrow \mathcal{M}$ is a smooth vector bundle.*

Example 1. Let $d = \dim \mathcal{M}$ and consider a coordinate chart $(u^1, \dots, u^d) : U \rightarrow \mathbb{R}^d$, for $U \subseteq \mathcal{M}$. Recall that for each $x \in U$, the induced coordinate basis in $T_x \mathcal{M}$,

$$\omega(x) := \left(\frac{\partial}{\partial u^1} \Big|_x, \dots, \frac{\partial}{\partial u^d} \Big|_x \right), \quad (2.11)$$

lets us express tangent vectors $X_x \in T_x \mathcal{M}$ in components $X_x^1, \dots, X_x^d \in \mathbb{R}$. Moreover, (2.11) defines a local frame $\omega : U \rightarrow F\mathcal{M}$ that sends each point $x \in U$ to its coordinate basis in $T_x \mathcal{M}$. Local frames are sections of the frame bundle $F\mathcal{M}$, which is a principal $GL(d, \mathbb{R})$ -bundle, hence local frames are examples of gauges.

Fix $x \in U$ and expand $X_x \in T_x \mathcal{M}$ in the coordinate basis:

$$X_x = \sum_{i=1}^d X_x^i \frac{\partial}{\partial u^i} \Big|_x. \quad (2.12)$$

By decomposing (2.12) into components and basis vectors, we can view the tangent vector X_x as one giant tuple

$$X_x = \left(\frac{\partial}{\partial u^1} \Big|_x, \dots, \frac{\partial}{\partial u^d} \Big|_x, X_x^1, \dots, X_x^d \right) = (\omega(x), X(x)) \in F\mathcal{M} \times \mathbb{R}^d. \quad (2.13)$$

Another choice of coordinate chart produces another local frame $\omega' : U' \rightarrow F\mathcal{M}$ and another decomposition (2.12), assuming that $x \in U'$. These decompositions are related by a change of basis

$$(\omega'(x), X'(x)) = (\omega(x)B(x), B(x)^{-1}X(x)), \quad (2.14)$$

for some $B(x) \in GL(d, \mathbb{R})$. Now observe that (2.14) is of the form

$$(p', v') = (p \triangleleft k, \rho(k)^{-1}v), \quad (2.15)$$

where $p, p' \in F\mathcal{M}$, $v, v' \in \mathbb{R}^d$, $k \in GL(d, \mathbb{R})$, the right-action \triangleleft is right-multiplication, and (ρ, \mathbb{R}^d) is the standard representation $\rho(k) = k$ of $GL(d, \mathbb{R})$. The basis-dependent description (2.12) of X_x thus resemble the pairs (p, v) in the construction of associated bundles. Passing to the quotient $(F\mathcal{M} \times_{\rho} \mathbb{R}^d) / \sim$ instead gives a basis-independent description of X_x , since it identifies all possible decompositions (2.12) in all possible bases. That is, the tangent bundle is isomorphic to $F\mathcal{M} \times_{\rho} \mathbb{R}^d$. ■

Equivariant neural networks use the language of principal and associated bundles. In the remainder of this subsection, let $E_{\rho} = P \times_{\rho} V_{\rho}$ and $E_{\sigma} = P \times_{\sigma} V_{\sigma}$ be associated bundles, given a principal bundle $\pi : P \rightarrow \mathcal{M}$ over a smooth manifold \mathcal{M} . Further let $\Gamma_c(E_{\rho})$ and $\Gamma_c(E_{\sigma})$ be the vector spaces of compactly supported continuous sections of E_{ρ} and E_{σ} , respectively.

Definition 3. A *data point* is a section $s \in \Gamma_c(E_{\rho})$.

Remark 3. Even though data is typically real-valued, we primarily consider complex representations (ρ, V_ρ) so to simplify the mathematical theory. The harmonic analysis in Section 3.5 especially benefits from this choice.

Our decision to restrict attention to compactly supported sections was also made for mathematical reasons: G -equivariant layers are defined in Section 3.4 in terms of an induced representation, which lives on the completion of $\Gamma_c(E_\rho)$ with respect to a certain inner product. This is not a serious restriction from an application viewpoint.

Definition 4. A *feature map* is a compactly supported continuous map $f : P \rightarrow V_\rho$ that satisfies the transformation property

$$f(p \triangleleft k) = \rho(k)^{-1} f(p), \quad (2.16)$$

for all $p \in P, k \in K$. The vector space of such feature maps is denoted $C_c(P; \rho)$

Data points and feature maps are, in a sense, dual to each other: Each data point in $\Gamma_c(E_\rho)$ is of the form

$$s_f(x) = [p, f(p)], \quad (2.17)$$

for a feature map $f \in C_c(P; \rho)$, where $p \in P_x$ is any element of the fiber at $x \in \mathcal{M}$. Note that (2.17) does not depend on the choice of p : Given another element $p' \in P_x$, there exists a unique $k \in K$ such that $p' = p \triangleleft k$ and

$$[p', f(p')] = [p \triangleleft k, f(p \triangleleft k)] = [p \triangleleft k, \rho(k^{-1})f(p)] = [p, f(p)]. \quad (2.18)$$

That is, the equivalence class $[p, f(p)]$ only depends on the basepoint x .

Lemma 2 ([26, §10.12]). *The linear map $C_c(P; \rho) \rightarrow \Gamma_c(E_\rho), f \mapsto s_f$ is a vector space isomorphism.*

We are almost ready to define general and gauge equivariant layers. Before doing so, however, we must say how gauge transformations $\chi : P \rightarrow P$ act on data points. Let $\theta_\chi : P \rightarrow K$ be the uniquely defined map satisfying $\chi(p) = p \triangleleft \theta_\chi(p)$ for all $p \in P$, and define the following action on the associated bundle E_ρ :

$$\chi \cdot [p, v] = [\chi(p), v] = [p, \rho(\theta_\chi(p))v], \quad [p, v] \in E_\rho. \quad (2.19)$$

The corresponding action on data points is given by

$$(\chi \cdot s_f)(p) = [p, \rho(\theta_\chi(p))f(p)] = s_{\rho(\theta_\chi)f}(x), \quad s_f \in \Gamma_c(E_\rho). \quad (2.20)$$

We distinguish between general layers and more specific gauge equivariant layers, as G -equivariant layers in GCNNs will only be a special case of the former.

Definition 5. Let $E_\rho = P \times_\rho V_\rho$ and $E_\sigma = P \times_\sigma V_\sigma$ be associated bundles.

- (i) A *(linear) layer* is a linear map $\Phi : \Gamma_c(E_\rho) \rightarrow \Gamma_c(E_\sigma)$.
- (ii) A layer Φ is *gauge equivariant* if, for all gauge transformations $\chi : P \rightarrow P$,

$$\Phi \circ \chi = \chi \circ \Phi. \quad (2.21)$$

In equivariant neural networks, data points are sent through a sequence of layers, which are mixed with non-linear activation functions. Again, we focus on individual layers in this paper, and leave the analysis of equivariant activation functions and multi-layer networks for an upcoming paper [23]. The fiber bundle-theoretic concepts discussed in this part describe two kinds of equivariant neural networks:

- (i) Gauge equivariant neural networks, which respect local gauge symmetry and whose layers are gauge equivariant.
- (ii) GCNNs, which respect global translation symmetry in homogeneous G -spaces \mathcal{M} , and whose layers are G -equivariant (Definition 9).

Remark 4. Our definition of layers is almost identical to the linear maps in [9]. The difference is that we focus on compactly supported sections, whereas [9] use sections that are supported on a single coordinate chart. Also, [9] investigates the structure of their linear maps under additional assumptions of so-called locality, covariance, and weight-sharing. Covariance is analogous to gauge equivariance in our setting.

A consequence of Lemma 2 is that (gauge equivariant) layers $\Phi : \Gamma_c(E_\rho) \rightarrow \Gamma_c(E_\sigma)$ induce a unique linear map $\phi : C_c(P; \rho) \rightarrow C_c(P; \sigma)$ such that $\Phi s_f = s_{\phi f}$. Writing data points as $s_f = [\cdot, f]$ allows us to also express this relation as $\Phi[\cdot, f] = [\cdot, \phi f]$. We think of Φ and ϕ as two sides of the same coin, and use the name (gauge equivariant) layer for both maps.

$$\begin{array}{ccc} \Gamma_c(E_\rho) & \xrightarrow{\Phi} & \Gamma_c(E_\sigma) \\ \updownarrow & & \updownarrow \\ C_c(P; \rho) & \xrightarrow{\phi} & C_c(P; \sigma) \end{array}$$

Example 2. Let $T : V_\rho \rightarrow V_\sigma$ be a linear transformation and consider the layer

$$\phi : C_c(P; \rho) \rightarrow C_c(P; \sigma), \quad (\phi f)(p) = T(f(p)), \quad (2.22)$$

for $p \in P, f \in C_c(P; \rho)$. Since f and ϕf are feature maps and thereby satisfy (2.16), the linear transformation T must satisfy

$$\sigma(k)T(f(p)) = T(f(p \triangleleft k^{-1})) = T(\rho(k)f(p)), \quad (2.23)$$

for all $k \in K, p \in P, f \in C_c(P; \rho)$. This can be seen to imply that $\sigma \circ T = T \circ \rho$, so T intertwines the representations ρ and σ . Another way to arrive at this conclusion is to analyze when the corresponding layer

$$\Phi : \Gamma_c(E_\rho) \rightarrow \Gamma_c(E_\sigma), \quad \Phi s_f = [\cdot, \phi f], \quad (2.24)$$

is well-defined.

Now consider a gauge transformation $\chi : P \rightarrow P$ and its induced map $\theta_\chi : P \rightarrow K$. Because T is an intertwiner,

$$(\Phi \circ \chi)s_f = s_{\phi\rho(\theta_\chi)f} = s_{\sigma(\theta_\chi)\phi f} = (\chi \circ \Phi)s_f, \quad (2.25)$$

hence the layer Φ is automatically gauge equivariant. \blacksquare

As this example illustrates, gauge equivariance is tightly connected to intertwining properties of ϕ . Rearranging (2.25) gives the following result.

Lemma 3. *A general layer $\Phi : \Gamma_c(E_\rho) \rightarrow \Gamma_c(E_\sigma)$ is gauge equivariant iff*

$$\phi \circ \rho(\theta_\chi)f = \sigma(\theta_\chi) \circ \phi f, \quad (2.26)$$

for all gauge transformations $\chi : P \rightarrow P$ and all feature maps $f \in C_c(P; \rho)$.

This concludes our discussion of gauge theory and of equivariant neural networks. The framework for the latter is evidently very general, consisting of layers and non-linear activation functions between data points. There are advantages of working at this level of generality: Ordinary (non-equivariant) neural networks have a multitude of different types of layers, many of them linear. Equivariant analogues of such layers are likely to satisfy either Definition 5(ii) or Definition 9, depending on the relevant type of equivariance. Any result that can be proven using this general framework, will thus be true for many different instances of equivariant neural networks. One example is Theorem 14 below, that characterizes the structure of abstract G -equivariant layers in any GCNN.

3. G -EQUIVARIANT CONVOLUTIONAL NEURAL NETWORKS

Recall that GCNNs generalize ordinary CNNs to data points $f : \mathcal{M} \rightarrow V$ defined on homogeneous G -spaces \mathcal{M} . Let us give a brief recap on homogeneous spaces and global symmetry, before moving on to discuss homogeneous vector bundles, sections, and induced representations. We will demonstrate that GCNNs and G -equivariant layers (originally defined in [12]) are most naturally understood from the perspective of homogeneous vector bundles. We then use reproducing kernel Hilbert spaces and bandwidth to understand which G -equivariant layers are expressible as convolutional layers.

3.1. Homogeneous spaces.

Definition 6. Let G be a Lie group. A smooth manifold \mathcal{M} is called a *homogeneous G -space* if there exists a smooth, transitive left G -action

$$G \times \mathcal{M} \rightarrow \mathcal{M}, \quad (g, x) \mapsto g \cdot x. \quad (3.1)$$

Since the action (3.1) is transitive, we may choose an arbitrary basepoint $x_0 \in \mathcal{M}$ and express any other point $x \in \mathcal{M}$ as $x = g \cdot x_0$ for some $g \in G$. This group element is typically not unique, but observe that

$$g \cdot x_0 = g' \cdot x_0 \iff g^{-1}g' \in H_{x_0}, \quad (3.2)$$

where $H_{x_0} = \{g \in G \mid g \cdot x_0 = x_0\}$ is the isotropy group of x_0 . In other words, there is a one-to-one correspondence between points $x \in \mathcal{M}$ and left cosets $gH_{x_0} \in G/H_{x_0}$.

Proposition 4 ([34, Theorem 21.18]). *Let \mathcal{M} be a homogeneous G -space and choose a basepoint $x_0 \in \mathcal{M}$. The isotropy group H_{x_0} is a closed subgroup of G , and the map*

$$F_{x_0} : G/H_{x_0} \rightarrow \mathcal{M}, \quad gH_{x_0} \mapsto g \cdot x_0, \quad (3.3)$$

is an equivariant diffeomorphism.

Homogeneous spaces are globally symmetric in the sense that any point $x_0 \in \mathcal{M}$ may be chosen as basepoint. Given another choice of basepoint $x'_0 \in \mathcal{M}$, the spaces $G/H_{x'_0} \simeq G/H_{x_0}$ are diffeomorphically related by a translation in G - more precisely, by the composition $F_{x_0}^{-1} \circ F_{x'_0}$. Euclidean space $\mathcal{M} = \mathbb{R}^d$, for example, possesses a global translation symmetry, allowing any point to be considered as origin. Similarly, the rotationally symmetric sphere $\mathcal{M} = S^2$ does not have a unique north pole.

We end this part with the following proposition, which is instrumental in relating homogeneous vector bundles to the equivariance framework in Section 2.2.

Proposition 5 ([43, §7.5]). *Let G be a Lie group and let $H \leq G$ be a closed subgroup. Then the quotient map*

$$q : G \rightarrow G/H, \quad g \mapsto gH, \quad (3.4)$$

defines a smooth principal H -bundle over the homogeneous G -space $\mathcal{M} = G/H$.

3.2. Homogeneous vector bundles. Vector bundles may inherit global symmetry from a homogeneous base space; the transitive action $(g, x) \mapsto g \cdot x$ may induce linear maps $E_x \mapsto E_{gx}$ between fibers. Such bundles are naturally called *homogeneous* and, because this symmetry is also encoded in its sections (data points), we will show that homogeneous vector bundles is the natural setting for studying GCNNs.

From this point on, we restrict attention to homogeneous spaces $\mathcal{M} = G/K$ where G is a unimodular Lie group and $K \leq G$ is a compact subgroup. Elements of the homogeneous space is interchangeably denoted as $x \in \mathcal{M}$ or $gK \in G/K$.

Remark 5. Examples of unimodular Lie groups include all finite, discrete, compact, or abelian Lie groups, the Euclidean groups, and many others. See [19, 20] for details.

Definition 7 ([47, 5.2.1]). Let \mathcal{M} be a homogeneous G -space and let $\pi : E \rightarrow \mathcal{M}$ be a smooth vector bundle with fibers E_x . We say that E is *homogeneous* if there is a smooth left G -action $G \times E \rightarrow E$ satisfying

$$g \cdot E_x = E_{gx}, \quad (3.5)$$

and such that the induced map $L_{g,x} : E_x \rightarrow E_{gx}$ is linear, for all $g \in G, x \in \mathcal{M}$.

Example 3. The frame bundle $F\mathcal{M}$ is a homogeneous vector bundle whenever \mathcal{M} is a homogeneous space, and the same is true of any associated bundle $F\mathcal{M} \times_\rho V_\rho$. In particular, the tangent bundle $T\mathcal{M}$ is a homogeneous vector bundle. ■

Example 4. If (ρ, V_ρ) is a finite-dimensional K -representation, then the associated bundle $E_\rho = G \times_\rho V_\rho$ is a homogeneous vector bundle with respect to the left action

$$g \cdot [g', v] = [gg', v]. \quad (3.6)$$

■

All homogeneous vector bundles E are of the form $G \times_\rho V_\rho$, up to isomorphism. To understand why, consider the fiber $E_K = E_{eK}$ and observe that the restriction of (3.5) to E_K and elements $k \in K$ yields invertible linear maps

$$L_k : E_K \rightarrow E_K. \quad (3.7)$$

The defining properties of group actions ensure that $\rho(k) = L_k$ is a finite-dimensional K -representation on E_K . Moreover, because the linear maps $L_{g,x}$ are isomorphisms, any element v' of any fiber E_x can be obtained as the image $v' = L_{g,K}(v) =: L_g(v)$ for some choices of $g \in q^{-1}(\{x\})$ and $v \in E_K$. The mapping

$$\xi : G \times E_K \rightarrow E, \quad (g, v) \mapsto L_g(v), \quad (3.8)$$

is thus surjective. It is not injective, though, since the relation

$$L_g = L_g \circ L_k \circ L_{k^{-1}} = L_{gk} \circ \rho(k^{-1}), \quad (3.9)$$

implies that $\xi(g, v) = L_g(v) = L_{gk}(\rho(k^{-1})v) = \xi(gk, \rho(k^{-1})v)$ for $k \in K$. However, the same argument shows that ξ is made injective by passing to the quotient $G \times_\rho E_K$.

Lemma 6 ([47, 5.2.3]). *The map*

$$G \times_\rho E_K \rightarrow E, \quad [g, v] \mapsto L_g(v), \quad (3.10)$$

is an isomorphism of homogeneous vector bundles.

We now have two perspectives on bundles $G \times_\rho V_\rho$: As bundles associated to the principal bundle $P = G$, and as homogeneous vector bundles (up to isomorphism). The former perspective offers a connection to the framework in Section 2.2, whereas the latter motivates the definition of G -equivariant layers in Section 3.4 below.

3.3. Induced representations. Let us show the relationship between homogeneous vector bundles and induced representations, which will be an essential ingredient in the definition of G -equivariant layers. To this end, let (ρ, V_ρ) be a finite-dimensional unitary K -representation and consider the homogeneous vector bundle $E_\rho = G \times_\rho V_\rho$.

We will need inner products on $\Gamma_c(E_\rho)$ and $C_c(G; \rho)$, the former of which is defined using the following unitary structure:

Lemma 7 ([47, 5.2.7]). *The unitary structure*

$$\langle [g, v], [g, w] \rangle_{gK} := \langle v, w \rangle_\rho, \quad (3.11)$$

defines a complete inner product on each fiber E_{gK} , making E_ρ into a Hilbert bundle with $L_{g,x}$ unitary. This unitary structure is unique in that, if we identify V_ρ with E_K in the canonical manner, then the inner product on V_ρ so induced agrees with $\langle \cdot, \cdot \rangle_\rho$.

We also need the following measure on G/K :

Theorem 8 (Quotient Integral Formula [14, §1.5]). *There is a unique G -invariant, nonzero Radon measure dx on G/K such that the following quotient integral formula holds for every $f \in C_c(G)$:*

$$\int_G f(g) \, dg = \int_{G/K} \int_K f(xk) \, dk \, dx. \quad (3.12)$$

Using these two ingredients, we make $\Gamma_c(E_\rho)$ into a pre-Hilbert space with respect to the inner product

$$\langle s, s' \rangle_{L^2(E_\rho)} := \int_{G/K} \langle s(x), s'(x) \rangle_x \, dx, \quad s, s' \in \Gamma_c(E_\rho), \quad (3.13)$$

and we denote its completion $L^2(E_\rho)$. Similarly, $C_c(G; \rho)$ is a pre-Hilbert space with respect to the inner product

$$\langle f, f' \rangle_{L^2(G; \rho)} = \int_G \langle f(g), f'(g) \rangle_\rho \, dg, \quad f, f' \in C_c(G; \rho), \quad (3.14)$$

the completion of which is denoted $L^2(G; \rho)$.

Definition 8. The G -representations

$$\text{ind}_K^G \rho(g) : L^2(E_\rho) \rightarrow L^2(E_\rho), \quad (\text{ind}_K^G \rho(g)s)(x) = g \cdot s(g^{-1}x), \quad (3.15)$$

$$\text{Ind}_K^G \rho(g) : L^2(G; \rho) \rightarrow L^2(G; \rho), \quad (\text{Ind}_K^G \rho(g)f)(g') = f(g^{-1}g'). \quad (3.16)$$

are called *induced representations*, or *representations induced by ρ* .

Both $\text{ind}_K^G \rho$ and $\text{Ind}_K^G \rho$ are unitary [47, 5.3.2] and may be identified:

Lemma 9. *The induced representations $\text{ind}_K^G(\rho)$, $\text{Ind}_K^G(\rho)$ are unitarily equivalent.*

Proof. This is [47, 5.3.4], but let us write down a proof for clarity. First observe that the isomorphism $C_c(G; \rho) \rightarrow \Gamma_c(E_\rho)$, $f \mapsto s_f$ is unitary, which follows by combining the quotient integral formula (3.12), the unitarity of ρ , and the compactness of K : For all $f, f' \in C_c(G; \rho)$, the map $g \mapsto \langle f(g), f'(g) \rangle_\rho$ lies in $C_c(G)$ and so

$$\begin{aligned} \langle f, f' \rangle_{L^2(G; \rho)} &= \int_{G/K} \int_K \langle f(xk), f'(xk) \rangle_\rho \, dk \, dx \\ &= \int_{G/K} \langle f(x), f'(x) \rangle_\rho \, dx \\ &= \int_{G/K} \langle [x, f(x)], [x, f'(x)] \rangle_x \, dx = \langle s_f, s_{f'} \rangle_{L^2(E_\rho)}. \end{aligned} \quad (3.17)$$

The same map $f \mapsto s_f$ satisfies

$$\left(\text{ind}_K^G(\rho)s_f \right)(x) = g \cdot s_f(g^{-1}x) = [x, f(g^{-1}x)] = s_{\text{Ind}_K^G(\rho)f}(x), \quad (3.18)$$

so it extends to a unitary isomorphism $L^2(G; \rho) \rightarrow L^2(E_\rho)$ intertwining the induced representations. \square

To gain a better understanding of the induced representations, consider the Bochner space $L^2(G, V)$, the space of square-integrable functions $f : G \rightarrow V$ that take values in a finite-dimensional Hilbert space V . It is itself a Hilbert space with inner product

$$\langle f, f' \rangle_{L^2(G, V)} = \int_G \langle f(g), f'(g) \rangle_V \, dg. \quad (3.19)$$

The induced representation $(\text{Ind}_K^G \rho, L^2(G; \rho))$ is nothing but the restriction of the left regular representation Λ on $L^2(G, V_\rho)$ to a closed, invariant subspace. Furthermore, Λ is intimately related to the left regular representation λ on $L^2(G)$, as the following lemma shows. The proof of this lemma is a short calculation.

Lemma 10. *Let V be a finite-dimensional Hilbert space and equip $L^2(G) \otimes V$ with the tensor product inner product. Then the natural unitary isomorphism*

$$\begin{aligned} A : L^2(G) \otimes V &\rightarrow L^2(G, V) \\ f \otimes v &\mapsto fv \end{aligned} \tag{3.20}$$

intertwines $\lambda \otimes \text{Id}_V$ with Λ .

This lemma also shows that, if we choose an orthonormal basis $e_1, \dots, e_{\dim V} \in V$, elements of $L^2(G, V)$ are simply linear combinations $f = \sum_i f^i e_i$ with component functions $f^i \in L^2(G)$. We use this fact in some calculations of vector-valued integrals, and the component functions will also be important in Section 3.5.

3.4. G -equivariant and convolutional layers. Given a homogeneous G -space \mathcal{M} , we observed that vector bundles $\pi : E \rightarrow \mathcal{M}$ may inherit the global symmetry of \mathcal{M} . We took a closer look at such homogeneous vector bundles and found that they are isomorphic to associated bundles $G \times_\rho V_\rho$, and therefore fit within the equivariance framework of Section 2.2. We also saw how the global symmetry of \mathcal{M} is encoded in data points and feature maps via induced representations, and we want G -equivariant layers to preserve this global symmetry.

Consider homogeneous vector bundles $E_\rho = G \times_\rho V_\rho$ and $E_\sigma = G \times_\sigma V_\sigma$, and recall Definition 5 of layers as general linear maps $\Phi : \Gamma_c(E_\rho) \rightarrow \Gamma_c(E_\sigma)$. We are mainly interested in *bounded* layers from an application point of view, and we can make this restriction now that the domain and codomain are normed spaces. Furthermore, any bounded layer can be uniquely extended to a bounded linear map

$$\Phi : L^2(E_\rho) \rightarrow L^2(E_\sigma), \tag{3.21}$$

and we assume this extension has already been made.

Definition 9. A bounded linear map $\Phi : L^2(E_\rho) \rightarrow L^2(E_\sigma)$ is called a *G -equivariant layer* if it intertwines the induced representations:

$$\Phi \circ \text{ind}_K^G \rho = \text{ind}_K^G \sigma \circ \Phi. \tag{3.22}$$

That is, G -equivariant layers are elements $\Phi \in \text{Hom}_G(L^2(E_\rho), L^2(E_\sigma))$.

Remark 6. We could also have defined G -equivariant layers as bounded linear maps $\phi : L^2(G; \rho) \rightarrow L^2(G; \sigma)$ that intertwine the induced representations:

$$\phi \circ \text{Ind}_K^G \rho = \text{Ind}_K^G \sigma \circ \phi, \tag{3.23}$$

i.e., elements $\phi \in \text{Hom}_G(L^2(G; \rho), L^2(G; \sigma))$. These definitions are clearly equivalent.

Apart from minor technical differences, Definition 9 coincides with the definition of equivariant maps in [12]. We have thus obtained GCNNs almost directly from the definition of homogeneous vector bundles and a desire for layers to respect the global symmetry. This shows that homogeneous vector bundles is the natural setting for GCNNs.

Let us now define convolutional layers.

Definition 10. A *convolutional layer* $L^2(G; \rho) \rightarrow L^2(G; \sigma)$ is a bounded operator

$$[\kappa \star f](g) = \int_G \kappa(g^{-1}g')f(g') \, dg', \quad f \in L^2(G; \rho), \tag{3.24}$$

with an operator-valued kernel $\kappa : G \rightarrow \text{Hom}(V_\rho, V_\sigma)$.

Of course, not any function $\kappa : G \rightarrow \text{Hom}(V_\rho, V_\sigma)$ can be chosen as the kernel of a convolutional layer. The kernel must ensure both that (3.24) is bounded and that $\phi f \in L^2(G; \sigma)$ for each $f \in L^2(G; \rho)$. We give a sufficient condition for boundedness in Lemma 12 and the other requirement has been studied in detail in [12, 31].

The next result is an almost immediate consequence of the Fubini-Tonelli theorem.

Proposition 11. *The adjoint of (3.24) is the integral operator*

$$[f * \kappa^*](g) = \int_G \kappa^*(g'^{-1}g) f(g') \, dg', \quad f \in L^2(G; \sigma), \quad (3.25)$$

where κ^* is the pointwise adjoint of κ . That is, $(\kappa \star \cdot)^* = \cdot * \kappa^*$.

One way to ensure that the operators (3.24)-(3.25) are bounded, is to put a bound on the kernel matrix elements $\kappa_{ij} : G \rightarrow \mathbb{C}$ for any given choice of bases in V_ρ, V_σ .

Lemma 12. *The operators (3.24) and (3.25) are bounded if $\kappa_{ij} \in L^1(G)$ for all i, j .*

Proof. We need only prove that (3.25) is bounded, its adjoint (3.24) will be bounded as well. Choose bases $e_1, \dots, e_{\dim V_\rho} \in V_\rho$ and $\tilde{e}_1, \dots, \tilde{e}_{\dim V_\sigma} \in V_\sigma$ and observe that, because $L^2(G; \sigma) \subset L^2(G, V_\sigma)$, Lemma 10 enables the decomposition of $f \in L^2(G; \sigma)$ into component functions $f^i \in L^2(G)$:

$$f = \sum_{i=1}^{\dim V_\sigma} f^i \tilde{e}_i. \quad (3.26)$$

To be clear, the kernel κ is similarly decomposed into matrix elements $\kappa_{ij} = \langle \tilde{e}_j, \kappa e_i \rangle_\sigma$ and we have $\kappa_{ji}^* = \overline{\kappa_{ij}}$. The integral (3.25) now takes the form

$$[f * \kappa^*](g) = \sum_{j=1}^{\dim V_\sigma} \left(\int_G \sum_{i=1}^{\dim V_\sigma} \kappa_{ji}^*(g'^{-1}g) f^i(g') \, dg' \right) e_j, \quad (3.27)$$

so by Young's convolution inequality,

$$\begin{aligned} \|f * \kappa^*\|_{L^2(G; \rho)}^2 &\leq \sum_{i,j} \int_G \left| \int_G \kappa_{ji}^*(g'^{-1}g) f^i(g') \, dg' \right|^2 dg = \sum_{i,j} \|f^i * \kappa_{ji}^*\|_2^2 \\ &\leq \sum_{i,j} \|\kappa_{ij}\|_1^2 \|f^i\|_2^2 \leq M \sum_i \|f^i\|_2^2 = M \|f\|_{L^2(G; \sigma)}^2, \end{aligned} \quad (3.28)$$

where $M = \sum_{i,j} \|\kappa_{ij}\|_1^2 < \infty$ if $\kappa_{ij} \in L^1(G)$ for all i, j . \square

We are interested in convolutional layers partly because they are concrete examples of G -equivariant layers, which we show next.

Proposition 13. *Convolutional layers are G -equivariant layers.*

Proof. Convolutional layers $\kappa \star \cdot : L^2(G; \rho) \rightarrow L^2(G; \sigma)$ are bounded linear operators by definition, so the only thing we need to prove is that $\kappa \star \cdot$ intertwines the induced representations. This follows immediately from left-invariance of the Haar measure: For each $f \in L^2(G; \rho)$ and all $g, h \in G$,

$$\begin{aligned} [\kappa \star (\text{Ind}_K^G \rho(g)f)](h) &= \int_G \kappa(h^{-1}g') f(g^{-1}g') \, dg' \quad (g' \mapsto gg') \\ &= \int_G \kappa((g^{-1}h)^{-1}g') f(g') \, dg' = [\kappa \star f](g^{-1}h), \end{aligned} \quad (3.29)$$

hence $[\kappa \star \text{Ind}_K^G \rho(g)f] = \text{Ind}_K^G \sigma(g)[\kappa \star f]$. \square

Example 5. Let us describe where ordinary CNNs fit in the present context. CNNs represent the case $G = \mathbb{Z}^2$ when $K = \{0\}$ is the trivial subgroup. The corresponding homogeneous space is $G/K = \mathbb{Z}^2/\{0\} = \mathbb{Z}^2$ and the quotient map $q : G \rightarrow G/K$ is thus the identity map on \mathbb{Z}^2 . Its inverse, the identity map $\omega : G/K \rightarrow G$, is a globally defined gauge that eliminates the need for gauge equivariance, as we may choose to work exclusively in this one gauge. This is just a reflection of the fact that

$$G = \mathbb{Z}^2 = \mathbb{Z}/\{0\} \times \{0\} = G/K \times K, \quad (3.30)$$

is (obviously) trivial as a principal bundle. Its associated bundles $E_\rho = \mathbb{Z}^2 \times_\rho V_\rho$ are also trivial: partly because the finite-dimensional K -representation σ must be trivial, and partly because each equivalence class $[g, v]$ only contains a single representative. These reasons are, of course, due to the triviality of K .

This is not to say that the equivariant framework of Section 2.2 is uninteresting when dealing with CNNs, or with GCNNs for other homogeneous spaces $\mathcal{M} = G/K$ with K trivial. We saw in Sections 3.2-3.3 how the homogeneity give rise to induced representations, which encode the global symmetry in both data points and feature maps. This is a useful perspective to have, and G -equivariant layers are interesting even when the bundles are trivial.

Triviality of the associated bundles, $E_\rho \simeq \mathbb{Z}^2 \times \mathbb{C}^m$ where $m = \dim V_\rho$,⁶ implies that data points and feature maps are general square-integrable functions,

$$L^2(E_\rho) \simeq L^2(\mathbb{Z}^2; \rho) \simeq L^2(\mathbb{Z}^2, \mathbb{C}^m), \quad (3.31)$$

and are thereby extensions of compactly supported functions $f : \mathbb{Z}^2 \rightarrow \mathbb{C}^m$. This ties well into the discussion in Section 2.1. Convolutional layers (3.24) reduce to bounded linear operators $L^2(\mathbb{Z}^2, \mathbb{C}^m) \rightarrow L^2(\mathbb{Z}^2, \mathbb{C}^n)$ and take the form

$$(\kappa \star f)(x) = \sum_{y \in \mathbb{Z}^2} \kappa(y - x) f(y), \quad (3.32)$$

as the Haar measure on \mathbb{Z}^2 is the counting measure. The kernel $\kappa : \mathbb{Z}^2 \rightarrow \text{Hom}(\mathbb{C}^m, \mathbb{C}^n)$ is finitely supported in practice, so boundeness of (3.32) is ensured by Lemma 12.

Interestingly, all \mathbb{Z}^2 -equivariant layers are convolutional layers; there are no other types of \mathbb{Z}^2 -equivariant layers than (3.32). This is a consequence of Theorem 14 and is proven in Corollary 20 below. ■

For more general groups G , it is no longer true that all G -equivariant layers are convolutional layers; we give an example of this fact in Example 6. Implementations of GCNNs, however, are usually based on convolutional layers, or on analogous layers in the Fourier domain. What consequences does the restriction to convolutional layers have for the expressivity of GCNNs? Can we tell whether a given G -equivariant layer is expressible as a convolutional layer? The answer to this last question, it turns out, requires the following notion of reproducing kernel Hilbert spaces.

Definition 11. Let G be a group, let V be a finite-dimensional normed vector space, and let \mathcal{H} be a Hilbert space of functions $G \rightarrow V$. Then \mathcal{H} is a *reproducing kernel Hilbert space (RKHS)* if the evaluation operator

$$\mathcal{E}_g : \mathcal{H} \rightarrow V, \quad f \mapsto f(g), \quad (3.33)$$

is bounded for all $g \in G$. Moreover, by *left-invariant RKH subspace* $\mathcal{H} \subseteq L^2(G, V)$ we mean a closed subspace that is both a RKHS and an invariant subspace for the left regular representation Λ on $L^2(G, V)$.

⁶Recall that we focus on complex vector bundles, hence the use of \mathbb{C}^m instead of \mathbb{R}^m .

Remark 7. The term RKHS is typically reserved for the scalar case $V = \mathbb{C}$, when the evaluation operator is a linear functional. Our version would instead be dubbed *vector-valued* RKHS. We see little benefit from distinguishing between these cases, however, so we use the term RKHS all-encompassingly.

The name RKHS is due to the existence of a kernel-type function that reproduces all elements of \mathcal{H} . To see how, choose an orthonormal basis $e_1, \dots, e_{\dim V} \in V$ and write elements $v \in V$ as linear combinations $v = \sum_i v^i e_i$. The projection $P_i(v) = v^i$ onto the i 'th component is always continuous, so the composition $\mathcal{E}_{g,i} := P_i \circ \mathcal{E}_g$ is a continuous linear functional

$$\mathcal{E}_g^i : \mathcal{H} \rightarrow \mathbb{C}, \quad f \mapsto f^i(g), \quad (3.34)$$

for all $g \in G$ and $i = 1, \dots, \dim V$. By the Riesz representation theorem, there are elements $\varphi_{g,i} \in \mathcal{H}$ such that $f^i(g) = \mathcal{E}_{g,i}(f) = \langle f, \varphi_{g,i} \rangle$, hence

$$f(g) = \sum_{i=1}^{\dim V} \langle f, \varphi_{g,i} \rangle e_i. \quad (3.35)$$

Now, if $\mathcal{H} \subseteq L^2(G, V)$ is a left-invariant RKH subspace, expanding the functions $\varphi_{g,i}$ in the orthonormal basis, $\varphi_{g,i} = \sum_j \varphi_{g,i}^j e_j$, yields the formula

$$f(g) = \sum_i \left(\int_G \sum_j \overline{\varphi_{g,i}^j(g')} f^j(g') \, dg' \right) e_i = \int_G \varphi_g^*(g') f(g') \, dg', \quad (3.36)$$

where φ_g^* is the conjugate transpose of the matrix $(\varphi_g)_i^j = \varphi_{g,i}^j$. By left-invariance,

$$f(g) = (\Lambda(g^{-1})f)(e) = \int_G \varphi_e^*(g^{-1}g') f(g') \, dg', \quad (3.37)$$

hence $f \in \mathcal{H}$ is *reproduced* by the operator-valued *kernel* $\varphi_e : G \rightarrow \text{Hom}(V)$.

Remark 8. The reproducing kernel φ_e is unique and thus independent of the choice of basis in V . This follows from uniqueness in the Riesz representation theorem.

It is now clear why left-invariant RKH subspaces of $L^2(G, V)$ are relevant when discussing convolutional layers, as the latter are given by integral operators similar to (3.37). In order to show that an abstract G -equivariant layer $\phi : L^2(G; \rho) \rightarrow L^2(G; \sigma)$ can be written as a convolutional layer, it is almost necessary for it to act in a RKHS:

Example 6. The identity operator $\phi : L^2(G; \sigma) \rightarrow L^2(G; \sigma)$ is clearly a G -equivariant layer regardless of G, K, σ , but it is only a convolutional layer if $L^2(G; \sigma)$ is a RKHS. This is because when ϕ is the identity, (3.24) becomes the reproducing property

$$f(g) = \int_G \kappa(g^{-1}g') f(g') \, dg', \quad f \in L^2(G; \sigma). \quad (3.38)$$

It follows that not every G -equivariant layer is a convolutional layer, because $L^2(G; \sigma)$ is not always a RKHS. When σ is the trivial representation, for instance, $L^2(G; \sigma)$ reduces to $L^2(G)$ which is not a RKHS when G is nondiscrete [20, Theorem 2.42]. ■

At this point, we know that global symmetry manifests itself in feature maps and data points through the induced representation, and we used this knowledge to define G -equivariant layers. We also defined convolutional layers and showed that these are special cases of G -equivariant layers, but the converse problem is much more subtle: When can a G -equivariant layer be expressed as a convolutional layer? The answer, as we have just seen, is directly related to the concept of RKHS and our next result makes this relation precise. It can be considered our main theorem.

Theorem 14. *Let G be a unimodular Lie group, let $K \leq G$ be a compact subgroup, and consider homogeneous vector bundles E_ρ, E_σ over $\mathcal{M} = G/K$. Suppose that*

$$\phi : L^2(G; \rho) \rightarrow L^2(G; \sigma), \quad (3.39)$$

is a G -equivariant layer. If ϕ maps into a left-invariant RKH subspace $\mathcal{H} \subseteq L^2(G; \sigma)$, then ϕ is a convolutional layer.

Proof. Fix orthonormal bases in V_ρ, V_σ . For $i = 1, \dots, \dim \sigma$, consider the functionals

$$\mathcal{E}^i : L^2(G; \rho) \rightarrow \mathbb{C}, \quad \mathcal{E}^i(f) = (\phi f)^i(e), \quad (3.40)$$

composing ϕ with evaluation at the identity element $e \in G$ and projection onto the i 'th component. As ϕ maps into a left-invariant RKH subspace $\mathcal{H} \subseteq L^2(G; \sigma)$, (3.40) is a bounded linear functional: $|\mathcal{E}^i(f)| \leq \|(\phi f)(e)\|_\sigma \leq \|\phi f\|_{L^2(G; \sigma)} \leq \|\phi\| \|f\|_{L^2(G; \rho)}$. By the Riesz representation theorem, there is a unique $\varphi_i \in L^2(G; \rho)$ such that

$$\mathcal{E}^i(f) = \int_G \langle f(g), \varphi_i(g) \rangle_\rho dg = \int_G \sum_{j=1}^{\dim \rho} f^j(g) \overline{\varphi_i^j(g)} dg, \quad (3.41)$$

and proceeding as in (3.36)-(3.37) with $\kappa := \varphi_e^*$ yields the desired relation

$$(\phi f)(g) = \int_G \kappa(g^{-1}g') f(g') dg'. \quad (3.42)$$

□

Remark 9. Theorem 14 is a generalization of [12, Theorem 6.1], which was proven under the assumption that ϕ is an integral operator $(\phi f)(g) = \int_G \kappa(g, g') f(g') dg'$.

Remark 10. While Theorem 14 is similar in spirit to [28, Theorem 1], there are also some clear differences. For example, we work with unimodular Lie groups whereas [28] use compact groups, but [28, Theorem 1] is also stronger in this case as there is no criterion on the layer. Another difference is that [28] analyzes the whole network structure while we focus on individual layers. We also assume that the homogeneous space G/K is the same before and after each layer, in contrast to [28].

In the special case of single-layer networks with compact G , [28, Theorem 1] states that any G -equivariant layer is a convolutional layer. Example 6 seems to contradict this statement when G is non-discrete compact. This conflict is possibly due to minor technical differences in the assumptions on layers and data points, but we have not identified the precise cause.

We end this section with a result that could simplify the numerical computations of convolutional layers, as integrals over G/K are sometimes easier to compute than integrals over G . For example when $G = SO(3)$, $K = SO(2)$, and $G/K \simeq S^2$. This result is similar to the generalized convolutions described in [28, Section 4.1]

Corollary 15. *Let $\phi : L^2(G; \rho) \rightarrow L^2(G; \sigma)$ be as in Theorem 14 and let κ be the kernel of the resulting convolutional layer (3.42). Then*

$$(\phi f)(g) = \int_{G/K} \kappa(g^{-1}x) f(x) dx. \quad (3.43)$$

Proof. In the proof of Theorem 14, we constructed the kernel κ from the components of $\varphi_i \in L^2(G; \rho)$, and unitarity of ρ clearly implies that the expression $\langle f(x), \varphi_i(x) \rangle_\rho$

is well-defined. We may therefore use the unitary structure (3.11) to get the following relation for all component functions $(\phi f)^i$ and all $g \in G$:

$$(\phi f)^i(g) = \langle f, \text{Ind}_K^G \rho(g) \varphi_i \rangle_{L^2(G; \rho)} = \langle s_f, s_{\text{Ind}_K^G \rho(g) \varphi_i} \rangle_{L^2(E_\rho)} \quad (3.44)$$

$$= \int_{G/K} \langle s_f(x), s_{\text{Ind}_K^G \rho(g) \varphi_i}(x) \rangle_x dx \quad (3.45)$$

$$= \int_{G/K} \langle f(x), \text{Ind}_K^G \rho(g) \varphi_i(x) \rangle_\rho dx \quad (3.46)$$

$$= \int_{G/K} \sum_{j=1}^{\dim \rho} \overline{\varphi_i^j(g^{-1}x)} f^j(x) dx. \quad (3.47)$$

We now obtain (3.43) by reconstructing κ from its components $\kappa_{ij} = \overline{\varphi_i^j}$. \square

3.5. RKHS and bandlimited functions. The strength of Theorem 14 naturally depends on how common left-invariant RKH subspaces of $L^2(G; \sigma)$ are. Our analysis of G -equivariant layers would not be complete without a discussion on this topic.

Let us proceed by investigating when the component functions f^i of $f \in L^2(G; \sigma)$ are contained in a left-invariant RKH subspace $\mathcal{H} \subset L^2(G)$; these subspaces have been fully characterized when the unimodular Lie group G is of *type I* [8, 20]. The unitary equivalence (3.20) then ensures that $A(\mathcal{H} \otimes V_\sigma) \subset L^2(G, V_\sigma)$ is a left-invariant RKH subspace, and so is the closed subspace

$$A(\mathcal{H} \otimes V_\sigma) \cap L^2(G; \sigma) \subset L^2(G; \sigma). \quad (3.48)$$

Remark 11. Groups of type I are, in a sense, groups with manageable representation theory. They include the most common groups, such as all finite, discrete, compact, or abelian groups, the Euclidean groups and many other groups. In particular, there is a considerable overlap between type I groups and the unimodular Lie groups that we already consider. See [19, 20] for more details.

Remark 12. While ρ, σ still denote finite-dimensional unitary representations of K , we reserve the letter γ for elements of the unitary dual \widehat{G} , i.e., the space of equivalence classes of unitary representations. Specific representatives of γ are written as (π_γ, V_γ) , and note that V_γ need not be finite-dimensional unless G is compact. The unimodular Lie group G is assumed to be of type I throughout this section.

Proposition 16 ([20, Proposition 2.40]). *Let $\mathcal{H} \subseteq L^2(G)$ be a left-invariant RKH subspace. The kernel $\varphi \in \mathcal{H}$ is then a self-adjoint convolution idempotent,⁷ and*

$$\mathcal{H} = L^2(G) * \varphi = \{f * \varphi \mid f \in L^2(G)\} \subset C(G). \quad (3.49)$$

*Conversely, if $\varphi \in L^2(G)$ is a self-adjoint convolution idempotent, then $\mathcal{H} = L^2(G) * \varphi$ is a left-invariant RKH subspace of $L^2(G)$.*

Example 7. Consider the real line $G = \mathbb{R}$ and suppose $\mathcal{H} \subseteq L^2(\mathbb{R})$ is a left-invariant RKH subspace with kernel $\varphi \in \mathcal{H}$. The calculation in (3.37) with $V = \mathbb{C}$ shows that, for all $f \in \mathcal{H}$,

$$f(x) = \int_{-\infty}^{\infty} \overline{\varphi(y-x)} f(y) dy = \langle f, \lambda(x) \varphi \rangle = (f * \varphi^*)(x). \quad (3.50)$$

Since the regular representation λ is continuous, f must be continuous, so $\mathcal{H} \subset C(\mathbb{R})$. Setting $f = \varphi$ shows that the kernel is a self-adjoint convolution idempotent:

$$\varphi = \varphi * \varphi^* = (\varphi * \varphi^*)^* = \varphi^*. \quad (3.51)$$

⁷That is, $\varphi = \varphi * \varphi^* = \varphi^*$ where $\varphi^*(g) := \overline{\varphi(g^{-1})}$ denotes involution.

Combining the Plancherel transform on $L^2(\mathbb{R})$ (see Theorem 17 and Section 3.5.1) with the convolution theorem in Fourier analysis, we observe that, for all $f \in \mathcal{H}$,

$$\hat{f} = \widehat{f * \varphi} = \hat{f} \hat{\varphi}. \quad (3.52)$$

In particular, $\hat{\varphi} = \hat{\varphi}^2$, so $\hat{\varphi}$ is the characteristic function 1_E on a subset $E \subset \widehat{\mathbb{R}} \simeq \mathbb{R}$. Inserting $\hat{\varphi} = 1_E$ in (3.52) immediately tells us that $\text{supp}(\hat{f}) \subset E$, so \mathcal{H} is a space of bandlimited functions. Moreover, the set E has finite Lebesgue measure according to the Plancherel theorem: $\text{vol}(E) = \|1_E\|_2^2 = \|\varphi\|_2^2 < \infty$. ■

This example illustrates that any measurable subset $E \subset \mathbb{R}$ with finite Lebesgue measure induces a left-invariant closed RKH subspace

$$\mathcal{H}_E = \left\{ f \in L^2(\mathbb{R}) \mid \text{supp}(\hat{f}) \subset E \right\} = L^2(\mathbb{R}) * \varphi_E, \quad (3.53)$$

φ_E being the inverse Plancherel transform of 1_E [20, 2.63-2.65]. This relation between left-invariant RKH subspaces $\mathcal{H} \subseteq L^2(G)$ and bandlimited functions generalizes to unimodular Lie groups G of type I, although the necessary harmonic analysis becomes significantly more advanced. Going into detail on this rather technical subject would distract from the topic at hand, so we refer curious readers to the relevant literature instead [20]. Let us take the short route of stating a theorem on the direct integral decomposition of the left regular representation λ and its commutant

$$\lambda(G)' = \left\{ T \in \mathcal{B}(L^2(G)) \mid T \circ \lambda(g) = \lambda(g) \circ T \text{ for all } g \in G \right\}, \quad (3.54)$$

and discuss a few consequences of this decomposition, before restricting attention to two important cases where we can be more explicit: Abelian and compact groups.

Definition 12 ([20, §3.5]). The *operator-valued Fourier transform* on G maps each $f \in L^1(G)$ to the family $\mathcal{F}(f) = (\hat{f}(\gamma))_{\gamma \in \widehat{G}}$, where each $\hat{f}(\gamma) \in \mathcal{B}(V_\gamma)$ is a bounded operator given by the Bochner integral

$$\hat{f}(\gamma) = \int_G f(g) \pi_\gamma(g) \, dg. \quad (3.55)$$

Theorem 17 ([20, Theorem 3.48]). *There is a canonical Plancherel measure ν for the unitary dual \widehat{G} with the following properties:*

- (a) \mathcal{F} extends to a unitary operator

$$\mathcal{P} : L^2(G) \rightarrow \int_{\widehat{G}}^{\oplus} V_\gamma \otimes V_\gamma \, d\nu(\gamma), \quad (3.56)$$

called the Plancherel transform of G .

- (b) \mathcal{P} implements the following unitary equivalences:

$$\lambda \simeq \int_{\widehat{G}}^{\oplus} \pi_\gamma \otimes \text{Id} \, d\nu(\gamma) \quad (3.57)$$

$$\lambda(G)' \simeq \int_{\widehat{G}}^{\oplus} \text{Id} \otimes \mathcal{B}(V_\gamma) \, d\nu(\gamma) \quad (3.58)$$

Observe that if $\mathcal{H} \subseteq L^2(G)$ is a left-invariant closed subspace, then the projection $P : L^2(G) \rightarrow \mathcal{H}$ commutes with the left-regular representation and is thus an element of the commutant $\lambda(G)'$. It therefore has a direct integral decomposition

$$P = \int_{\widehat{G}}^{\oplus} \text{Id} \otimes \hat{P}_\gamma \, d\nu(\gamma), \quad (3.59)$$

where $\hat{P}_\gamma \in \mathcal{B}(V_\gamma)$ for each $\gamma \in \widehat{G}$.

Theorem 18 ([20, Theorem 4.22, Proposition 2.40]). *Suppose that $\mathcal{H} \subseteq L^2(G)$ is a left-invariant closed subspace with equation (3.59) denoting the projection onto \mathcal{H} . Then \mathcal{H} is a RKHS iff*

$$\int_{\widehat{G}} \text{rank}(\hat{P}_\gamma) \, d\nu(\gamma) < \infty. \quad (3.60)$$

We interpret this theorem as a bandwidth restriction, similar to Example 7. The integrand in (3.60) is an integer-valued function on \widehat{G} , so the integral is finite only if the projection (3.59) is supported on a set $E \subseteq \widehat{G}$ of finite Plancherel measure,

$$\nu\left(\underbrace{\left\{\gamma \in \widehat{G} : \hat{P}_\gamma \neq 0\right\}}_E\right) = \int_{\widehat{G}} 1_E(\gamma) \, d\nu(\gamma) \leq \int_{\widehat{G}} \text{rank}(\hat{P}_\gamma) \, d\nu(\gamma). \quad (3.61)$$

That is, the left-invariant RKH subspaces $\mathcal{H} \subseteq L^2(G)$ are precisely those subspaces whose elements are bandlimited on a set $E \subseteq \widehat{G}$, in the sense that, for all $f \in \mathcal{H}$ and each equivalence class $\gamma \notin E$,

$$\hat{f}(\gamma) = \widehat{P}f(\gamma) = \hat{f}(\gamma) \circ \hat{P}_\gamma = 0. \quad (3.62)$$

Remark 13. The second equality in (3.62) is [20, Corollary 4.17].

Bandlimited functions are thus central to the theory of RKHS and, by extension, to the mathematical theory of GCNNs. Indeed, by extending the concept of bandwidth to feature maps, through (3.48), we obtain the following rephrasing of Theorem 14.

Corollary 19. *Let G be a unimodular Lie group of type I, let $K \leq G$ be a compact subgroup, and consider homogeneous vector bundles E_ρ, E_σ over G/K . Suppose that*

$$\phi : L^2(G; \rho) \rightarrow L^2(G; \sigma), \quad (3.63)$$

is a G -equivariant layer. If ϕ maps into a space of bandlimited functions, then ϕ is a convolutional layer.

Remark 14. The relevance of bandwidth for convolutional layers has already been recognized in the case of azimuthally equivariant linear operators on $L^2(S^2)$ [45]. In our setting, these operators translate to certain G -equivariant layers

$$\Phi : L^2(E_\rho) \rightarrow L^2(E_\sigma), \quad (3.64)$$

when $G = SO(3)$, $K = SO(2)$, and ρ, σ are the trivial representation.

Remark 15. Some implementations of GCNNs use Fourier transforms and a variant of the convolution theorem $\widehat{f_1 * f_2} = \hat{f}_1 \hat{f}_2$ to compute convolutional layers [16, 29, 45, 48]. Feature maps f are then represented by their Fourier transform \hat{f} which, for numerical reasons, is only approximated up to a finite bandlimit. That is, bandwidth is already being used in implementations.

3.5.1. Abelian groups. The irreducible representations $\gamma \in \widehat{G}$ of any abelian group G are 1-dimensional, and may thus be identified with their character $\chi_\gamma = \text{tr } \pi_\gamma$. There are several useful consequences of this fact.

First, the unitary dual \widehat{G} is now the set of continuous homomorphisms $\chi : G \rightarrow \mathbb{T}$, where \mathbb{T} is the circle group. This is a locally compact group with respect to pointwise multiplication and, as $\gamma \in \widehat{G}$ is unitary, we may write $\chi_\gamma = e^{i\xi_\gamma}$ where $\xi_\gamma : G \rightarrow \mathbb{R}$. The Fourier transform then takes the more familiar form

$$\hat{f}(\gamma) = \int_G f(g) e^{-i\xi_\gamma(g)} \, dg, \quad (3.65)$$

for $f \in L^1(G) \cap L^2(G)$. Moreover, the Haar measure on \widehat{G} can be made to coincide with the Plancherel measure such that (3.56) becomes a unitary equivalence

$$\mathcal{P} : L^2(G) \rightarrow L^2(\widehat{G}). \quad (3.66)$$

Another consequence of the fact that irreducible representations are 1-dimensional, is that the integrand in (3.60) takes values in $\{0, 1\}$ and (3.61) becomes an equality. By the same arguments as in Example 7, we see that the left-invariant RKH subspaces $\mathcal{H} \subset L^2(G)$ are the spaces of bandlimited functions, $\text{supp}(\widehat{f}) \subset E$, for subsets $E \subset \widehat{G}$ of finite Haar/Plancherel measure. Also, the kernel $\varphi_E \in \mathcal{H}$ is the inverse Plancherel transform of the characteristic function 1_E .

Corollary 20. *If G is a discrete and abelian group, then any G -equivariant layer is a convolutional layer.*

Proof. Discrete groups are unimodular Lie group of type I, so we may use results from the current section. We note that the integral (3.60) converges for any left-invariant, closed subspace $\mathcal{H} \subseteq L^2(G)$, as the unitary dual \widehat{G} is compact when G is discrete and abelian [14, Proposition 3.1.5], and because the integrand is bounded. Consequently, $\mathcal{H} = L^2(G)$ is itself a RKHS,⁸ and the same is true for both $L^2(G, V_\sigma) \simeq L^2(G) \otimes V_\sigma$ and its closed, left-invariant subspace $L^2(G; \sigma)$, independently of $K \leq G$ and (σ, V_σ) . The result now follows from Theorem 14. \square

By setting $G = \mathbb{Z}^2$, Corollary 20 establishes that convolutional layers are the only possible translation equivariant layers in the ordinary CNN setting.

3.5.2. Compact groups. When the group G is compact, all irreducible representations are finite-dimensional. Furthermore, the unitary dual \widehat{G} is discrete and the Plancherel measure on \widehat{G} is simply the counting measure. For these reasons, the integral (3.60) reduces to a discrete sum with finite summands, and converges iff $\widehat{P}_\gamma = 0$ for all but finitely many $\gamma \in \widehat{G}$.

Corollary 21. *If G is finite, then any G -equivariant layer is a convolutional layer.*

Proof. When G is a finite group, \widehat{G} is also finite [19, Proposition 5.27] and the integral (3.60) reduces to a finite sum. That is, $L^2(G)$ is a RKHS and the result now follows in the same way as Corollary 20. \square

4. DISCUSSION

In this paper, we have investigated the mathematical foundations of G -equivariant convolutional neural networks (GCNNs), which are designed for deep learning tasks exhibiting global symmetry. We presented a basic framework for equivariant neural networks that include both gauge equivariant neural networks and GCNNs as special cases. We also demonstrated how GCNNs can be obtained from homogeneous vector bundles, when G is a unimodular Lie group and $K \leq G$ is a compact subgroup.

In Theorem 14, we gave a precise criterion for when a given G -equivariant layer is, in fact, a convolutional layer. This criterion uses reproducing kernel Hilbert spaces (RKHS) and cannot be circumvented, as shown in Example 6. After discussing the relation between RKHS and bandwidth, we were able to reformulate Theorem 14 to get an analogous bandlimit-criterion in Corollary 19. In Corollaries 20-21, we showed that the criterion is automatically satisfied when G is discrete abelian or finite, hence all G -equivariant layers are convolutional layers for these groups.

⁸In fact, the kernel is simply the Kronecker delta $\varphi(g) = \delta_{g,e}$.

One limitation of the current paper, compared to [12, 28], is that the homogeneous space G/K does not change between layers. This restriction was made in order to limit the scope of our analysis, and the same goes for our restriction to unimodular Lie groups. It would be interesting to go beyond these restrictions in the future.

I am grateful for the support from my research group: Oscar Carlsson, Jan Gerken, Hampus Linander, Fredrik Ohlsson, Christoffer Petersson, and last but not least, my advisor Daniel Persson. Thanks also to David Müller for the interesting discussions on gauge equivariance in physics. This work was supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

REFERENCES

- [1] S. Albawi, T. A. Mohammed, and S. Al-Zawi. Understanding of a convolutional neural network. In *International Conference on Engineering and Technology (ICET)*. IEEE, 2017.
- [2] E. Alpaydin. *Introduction to machine learning*. MIT Press, 2014.
- [3] D. L. Boyda, M. N. Chernodub, N. V. Gerasimeniuk, V. A. Goy, S. D. Liubimov, and A. V. Molochkov. Finding the deconfinement temperature in lattice yang-mills theories from outside the scaling window with machine learning. *Physical Review D*, 2021.
- [4] A. Brock, S. De, S. L. Smith, and K. Simonyan. High-performance large-scale image recognition without normalization. *arXiv preprint arXiv:2102.06171*, 2021.
- [5] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 2017.
- [6] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021.
- [7] W. Cao, Z. Yan, Z. He, and Z. He. A comprehensive survey on geometric deep learning. *IEEE Access*, 2020.
- [8] A. L. Carey. Group representations in reproducing kernel hilbert spaces. *Reports on Mathematical Physics*, 1978.
- [9] M. Cheng, V. Anagiannis, M. Weiler, P. de Haan, T. Cohen, and M. Welling. Covariance in physics and convolutional neural networks. *arXiv preprint arXiv:1906.02481*, 2019.
- [10] T. Cohen and M. Welling. Group equivariant convolutional networks. In *International conference on machine learning*. PMLR, 2016.
- [11] T. Cohen, M. Geiger, J. Köhler, and M. Welling. Spherical cnns. *arXiv preprint arXiv:1801.10130*, 2018.
- [12] T. Cohen, M. Geiger, and M. Weiler. A general theory of equivariant cnns on homogeneous spaces. *arXiv preprint arXiv:1906.02481*, 2018.
- [13] T. Cohen, M. Weiler, B. Kicanaoglu, and M. Welling. Gauge equivariant convolutional networks and the icosahedral cnn. In *International Conference on Machine Learning*. PMLR, 2019.
- [14] A. Deitmar and S. Echterhoff. *Principles of harmonic analysis*. Springer International Publishing, 2014.
- [15] C. Esteves. Theoretical aspects of group equivariant neural networks. *arXiv preprint arXiv:2004.05154*, 2020.

- [16] C. Esteves, C.e Allen-Blanchette, A. Makadia, and K. Daniilidis. 3D object classification and retrieval with spherical cnns. *arXiv preprint arXiv:1711.06721*, 2017.
- [17] C. Esteves, C. Allen-Blanchette, A. Makadia, and K. Daniilidis. Learning $so(3)$ equivariant representations with spherical cnns. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.
- [18] M. Favoni, A. Ipp, D. Müller, and D. Schuh. Lattice gauge equivariant convolutional neural networks. *arXiv preprint arXiv:2012.12901*, 2020.
- [19] G. B. Folland. *A course in abstract harmonic analysis*. CRC press, 2016.
- [20] H. Führ. *Abstract harmonic analysis of continuous wavelet transforms*. Springer Science & Business Media, 2005.
- [21] K. Fukushima and S. Miyake. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In *Competition and cooperation in neural nets*. Springer, 1982.
- [22] C. Gattringer and C. Lang. *Quantum chromodynamics on the lattice: an introductory presentation*. Springer Science & Business Media, 2009.
- [23] J. E. Gerken, J. Aronsson, O. Carlsson, H. Linander, F. Ohlsson, C. Petersson, and D. Persson. Geometric deep learning and equivariant neural networks. Forthcoming.
- [24] I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. MIT Press, 2016. URL <http://www.deeplearningbook.org>.
- [25] C. Jia, Y. Yang, Y. Xia, Y-T. Chen, Z. Parekh, H. Pham, Q.V. Le, Y. Sung, Z. Li, and T. Duerig. Scaling up visual and vision-language representation learning with noisy text supervision. *arXiv preprint arXiv:2102.05918*, 2021.
- [26] I. Kolár, P. W. Michor, and J. Slovák. *Natural operations in differential geometry*. Springer Science & Business Media, 2013.
- [27] O. Koller, S. Zargaran, H. Ney, and R. Bowden. Deep sign: Enabling robust statistical continuous sign language recognition via hybrid cnn-hmms. *International Journal of Computer Vision*, 2018.
- [28] R. Kondor and S. Trivedi. On the generalization of equivariance and convolution in neural networks to the action of compact groups. In *International Conference on Machine Learning*. PMLR, 2018.
- [29] R. Kondor, Z. Lin, and S. Trivedi. Clebsch-gordan nets: a fully fourier space spherical convolutional neural network. *arXiv preprint arXiv:1806.09231*, 2018.
- [30] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 2012.
- [31] L. Lang and M. Weiler. A wigner-eckart theorem for group equivariant convolution kernels. *arXiv preprint arXiv:2010.10952*, 2020.
- [32] S. Lawrence, C.L. Giles, A.C. Tsoi, and A.D. Back. Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks*, 1997.
- [33] Y. LeCun, L.D. Jackel, L. Bottou, C. Cortes, J.S. Denker, H. Drucker, I. Guyon, U.A. Muller, E. Sackinger, P. Simard, et al. Learning algorithms for classification: A comparison on handwritten digit recognition. *Neural networks: the statistical mechanics perspective*, 1995.
- [34] J. M. Lee. *Introduction to smooth manifolds*. Springer, 2013.
- [35] J. Li, Y. Bi, and G.H. Lee. Discrete rotation equivariance for point cloud recognition. In *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019.
- [36] D. Luo, G. Carleo, B. Clark, and J. Stokes. Gauge equivariant neural networks for quantum lattice gauge theories. *Bulletin of the American Physical Society*,

- 2021.
- [37] D. Mehta, S. Sridhar, O. Sotnychenko, H. Rhodin, M. Shafiei, H-P. Seidel, W. Xu, D. Casas, and C. Theobalt. Vnect: Real-time 3d human pose estimation with a single rgb camera. *ACM Transactions on Graphics (TOG)*, 2017.
 - [38] F. Monti, D. Boscaini, J. Masci, E. Rodola, J. Svoboda, and M. M. Bronstein. Geometric deep learning on graphs and manifolds using mixture model cnns. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017.
 - [39] D. Müller. Yang-mills theory, lattice gauge theory and simulations, 2019. URL <https://www.jku.at/en/institut-of-analysis/research/guest-lecture-series/>.
 - [40] M. Nakahara. *Geometry, topology and physics*. CRC press, 2003.
 - [41] C. Peng, X. Zhang, G. Yu, G. Luo, and J. Sun. Large kernel matters - improve semantic segmentation by global convolutional network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017.
 - [42] S. Song, H. Huang, and T. Ruan. Abstractive text summarization using lstm-cnn based deep learning. *Multimedia Tools and Applications*, 2019.
 - [43] N. Steenrod. *The topology of fibre bundles*. Princeton University Press, 1960.
 - [44] M. Tan and Q.V. Le. Efficientnetv2: Smaller models and faster training. *arXiv preprint arXiv:2104.00298*, 2021.
 - [45] C. Toft, G. Bökman, and F. Kahl. Azimuthal rotational equivariance in spherical cnns. 2021. URL <https://openreview.net/forum?id=sp3Z1jiS2vn>.
 - [46] B.S. Veeling, J. Linmans, J. Winkens, T. Cohen, and M. Welling. Rotation equivariant cnns for digital pathology. In *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2018.
 - [47] N. R. Wallach. *Harmonic analysis on homogeneous spaces*. M. Dekker, 2018.
 - [48] M. Weiler, M. Geiger, M. Welling, W. Boomsma, and T. Cohen. 3D Steerable CNNs: Learning Rotationally Equivariant Features in Volumetric Data. *arXiv preprint arXiv:1807.02547*, 2018.
 - [49] D. Worrall and G. Brostow. Cubenet: Equivariance to 3d rotation and translation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.
 - [50] H. Wu, B. Xiao, N. Codella, M. Liu, X. Dai, L. Yuan, and L. Zhang. Cvt: Introducing convolutions to vision transformers. *arXiv preprint arXiv:2103.15808*, 2021.
 - [51] J. Wu. Introduction to convolutional neural networks. *National Key Lab for Novel Software Technology. Nanjing University. China*, 2017.
 - [52] Y. Xu, Z. Jia, L-B. Wang, Y. Ai, F. Zhang, M. Lai, I Eric, and C. Chang. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC bioinformatics*, 2017.
 - [53] Y. Zhang, Z. Qiu, T. Yao, D. Liu, and T. Mei. Fully convolutional adaptation networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

CHALMERS UNIVERSITY OF TECHNOLOGY, DEPARTMENT OF MATHEMATICAL SCIENCES, SE-412 96 GOTHENBURG, SWEDEN

Email address: jimmyar@chalmers.se