Computing equivariant matrices on homogeneous spaces for geometric deep learning and automorphic Lie algebras

Vincent Knibbeler[1,2,3]

Abstract

We develop an elementary method to compute spaces of equivariant maps from a homogeneous space G/H of a Lie group G to a module of this group. The Lie group is not required to be compact. More generally, we study spaces of invariant sections in homogeneous vector bundles, and take a special interest in the case where the fibres are algebras. These latter cases have a natural global algebra structure. We classify these automorphic algebras for the case where the homogeneous space has compact stabilisers. This work has applications in the theoretical development of geometric deep learning and also in the theory of automorphic Lie algebras.

1 Introduction

Let G be a Lie group and X a homogeneous space of G, that is, a smooth manifold with transitive G-action: for any two points $x_0, x \in X$, there exists $g \in G$ with $gx_0 = x$. Let V be a representation of G and denote the space of G-equivariant maps $X \to V$ by

$$M_G(X, V) = \{K : X \to V \mid K(gx) = gK(x) \ \forall g \in G, \ \forall x \in X\}.$$

Communicated by: Gitta Kutyniok

✉ Vincent Knibbeler
  V.Knibbeler@gmail.com

1   Department of Mathematical Sciences, Loughborough University, Loughborough LE11 3TU, UK

2   Maxwell Institute for Mathematical Sciences, The Bayes Centre, Edinburgh EH8 9BT, UK

3   Department of Mathematics, Heriot-Watt University, Edinburgh EH14 4AS, UK

We develop an elementary method to compute this space. More generally, for any closed subgroup H of G and representation W of H, we compute the G-invariant sections of the homogeneous vector bundle $G \times_H W$ over G/H, as we explain in detail in Section 2. In Section 3, we replace V by an algebra $\mathfrak{a}$ with compatible G-module structure, and study the algebra $\mathfrak{M}_G(X, \mathfrak{a})$ of G-equivariant maps $X \to \mathfrak{a}$.

This investigation is motivated on one hand by recent developments of geometric deep learning, and on the other hand by the expanding horizons of research on automorphic Lie algebras. The connection between these two topics has not been studied so far, but they are dealing with closely related problems.

This paper can be read without knowledge of deep learning and automorphic Lie algebras, because we only tackle the differential geometric problem described above. Nonetheless, we give a brief description of the two areas in this introduction and explain why they motivate our objective. One can skip to Section 2 from here.

Artificial intelligence is currently developing at a rapid pace, and there is a great need for mathematics in this field, as is explained in the recent survey article by Kutyniok [1]. A particular challenge where researchers are looking to mathematics for answers is the search for good architectures of deep neural networks. One approach to this problem is the highly active field of geometric deep learning.

Two survey articles on geometric deep learning appeared in May 2021, by Bronstein et al. [2] and Gerken et al. [3] and another by Weiler et al. [4] appeared only a month later. We will not attempt to provide an overview of the literature but simply refer to the combined literature reviews in these three works, which we would not be able to improve. We do mention however the paper of Cohen et al. [5] and Cheng et al. [6], which highlight the importance of convolution with equivariant kernels. These papers inspired this research: we compute spaces of equivariant kernels.

We start our slightly longer summary of the geometric deep learning that motivated this paper with Cohen et al.'s description of equivariant convolutional networks on homogeneous spaces [7]. The features of such convolutional networks are sections of homogeneous vector bundles $G \times_H V$, which can be presented as maps $f : G \to V$ such that $f(gh) = h^{-1} f(g)$ for all $h \in H$ (see Section 2 for more details). The authors use the fact that cross-correlation on the group

$$(\kappa \star f)(g) = \int_G \kappa(g^{-1}h) f(h) \mathrm{d}h \tag{1}$$

is equivariant under the group G in the sense that

$$\kappa \star (u \cdot f) = u \cdot (\kappa \star f), \quad \forall u \in G,$$

granted that $\mathrm{d}h$ is a left-invariant measure on G.[1] The map $\kappa : G \to \mathrm{Hom}(V, V_{\mathrm{out}})$ is called the kernel, or the filter, and the group acts as $u \cdot f = f \circ u^{-1}$. The main result of [7] shows that $\kappa \star$ takes sections of $G \times_H V_{\mathrm{in}}$ to sections of $G \times_H V_{\mathrm{out}}$ if and only if κ descends to a section of the homogeneous vector bundle $G \times_{H \times H} \mathrm{Hom}(V_{\mathrm{in}}, V_{\mathrm{out}})$

---

[1] Cross-correlation is sometimes called convolution in the literature, although the term convolution is classically reserved for the operation $(\kappa * f)(g) = \int_G \kappa(gh^{-1}) f(h) \mathrm{d}h$, which is not G-equivariant.

defined by the right action of $H \times H$ on $G \times \mathrm{Hom}(V_{\mathrm{in}}, V_{\mathrm{out}})$ given by $(g, m) \cdot (h_1, h_2) = (h_2^{-1} g h_1, \rho_{\mathrm{out}}(h_2)^{-1} m \rho_{\mathrm{in}}(h_1))$.

Geometric deep learning developed further in the papers by Cohen et al. [5] and Cheng et al. [6] where one considers convolutional networks on manifolds M in general, rather than homogeneous spaces specifically. This adds another level of abstraction that we briefly describe. For a comprehensive treatment, we refer to [4].

Lacking a transitive group action on a manifold M, one has to find another way to move a filter over the manifold. The solution proposed in [5, 6] is to use parallel transport.

The i-th feature space is the space of sections Γ_i of a vector bundle $\pi_i : E_i \to M$. Features f_i are elements of Γ_i. The kernel is defined by a map

$$K : T_{x^*}M \to \mathrm{Hom}(\pi_{\mathrm{in}}^{-1}(x^*), \pi_{\mathrm{out}}^{-1}(x^*))$$

on the tangent space of M at the point x^*, and the convolution is then defined by an integral of the form

$$(K * f_{\mathrm{in}})(x) = \int_{B \subset T_{x^*}M} K(v) f_{\mathrm{in}}|_{\exp_x v}(x) \, \mathrm{d}^d v \qquad (2)$$

where $f_{\mathrm{in}}|_{\exp_x v}(x)$ is the parallel transport of $f_{\mathrm{in}}(\exp_x v)$ along the geodesic from $\exp_x v$ to x, and d the dimension of M.

The tangent space $T_{x^*}M$ is isomorphic to $\mathbb{R}^d$. A family of such isomorphisms parametrised by M is called a gauge. A choice of gauge allows us to consider the kernel as a map

$$K : \mathbb{R}^d \to \mathrm{Hom}(V_{\mathrm{in}}, V_{\mathrm{out}})$$

where we keep notation light by using the same letter K, and renaming the fibres.

Because the choice of gauge is arbitrary, convolution has to be invariant under any gauge transformation in the structure group $G \subset \mathrm{GL}(d, \mathbb{R})$. The action of the structure group on the feature spaces is denoted $\rho_{\mathrm{in}} : G \to \mathrm{GL}(V_{\mathrm{in}})$ and $\rho_{\mathrm{out}} : G \to \mathrm{GL}(V_{\mathrm{out}})$. Weiler et al. [4] show that the convolution is invariant under gauge transformations if an only if

$$K(gv) = \det(g)^{-1} \rho_{\mathrm{out}}(g) K(v) \rho_{\mathrm{in}}(g)^{-1}, \quad g \in G, \quad v \in \mathbb{R}^d. \qquad (3)$$

Cheng et al. [6] show that if the maps $\Gamma_{\mathrm{in}} \to \Gamma_{\mathrm{out}}$ between feature spaces satisfy some reasonable conditions, then they are of the form (2).

The kernel constraint (3) is solved for subgroups G of $\mathrm{O}(2)$ by Weiler and Cesa [8] and for $G = \mathrm{SO}(3)$ by Weiler et al. [9]. For compact groups in general, the solution can be expressed in terms of harmonic basis functions, as Lang and Weiler explain [10]. This leaves the constraint still to be solved for noncompact groups and for those compact groups for which we do not have harmonic basis functions readily available.

We propose to compute the space of G-equivariant kernels K on $\mathbb{R}^d$ by considering the G-orbits in $\mathbb{R}^d$ separately. Although it is not always a trivial task to combine the results on these orbits in order to describe the kernels on $\mathbb{R}^d$ (see Appendix B for

worked out examples), the advantage is that the computation of equivariant kernels on the orbits, homogeneous spaces X, can be done with simple methods that work equally well for noncompact groups as for compact ones.

Therefore, we will be looking for the G-equivariant maps $X \to V$ where $V = \det^* \otimes V_{\text{out}} \otimes V_{\text{in}}^*$. The special case $\det \otimes V_{\text{in}} = V_{\text{out}}$ comes with an associative algebra structure (matrix multiplication) and is therefore most interesting to us (the case $V_{\text{in}} = V_{\text{out}}$ also allows the use of biases in neural networks, and this case is highlighted in the discussion of [6] due to interesting analogies in physics).

To ensure the integrals (1) and (2) are finite, the kernels are often required to have compact support. We want to study equivariant kernels for noncompact groups G, which do not have compact support, so we do not impose this condition. The reader interested in convolution is advised to keep this in mind.

Automorphic Lie algebras were introduced in relation to integrable systems in mathematical physics by Lombardo and Mikhailov [11, 12]. They are Lie algebras of meromorphic maps from a Riemann surface X into a finite-dimensional Lie algebra $\mathfrak{g}$ which are equivariant with respect to a group Γ acting discretely on X and on $\mathfrak{g}$ by automorphisms, and have some restriction on the poles. The best-known examples are the twisted loop algebras

$$\mathcal{L}(\mathfrak{g}, \sigma, m) = \left\{ f : \mathbb{C}^* \to \mathfrak{g} \text{ Laurent polynomial} \mid f\left(e^{\frac{2\pi i}{m}} z\right) = \sigma f(z) \right\}$$

where σ is an automorphism of order m of the Lie algebra $\mathfrak{g}$. Twisted loop algebras are used in concrete realisations of affine Kac-Moody algebras. These are however somewhat misleading examples because they are the only simple automorphic Lie algebras due to the landmark papers by Kac and Mathieu [13, 14]: loop algebras are highly exceptional.

The algebra of Onsager [15] gives a better impression of the nature of the topic. In the original work of Onsager, it is not at all obvious that this Lie algebra $\mathfrak{O}$ fits in the framework of automorphic Lie algebras, but Roan showed that this is indeed the case [16], with the construction of an isomorphism

$$\mathfrak{O} \cong \left\{ f : \mathbb{C}^* \to \mathfrak{sl}(2, \mathbb{C}) \text{ Laurent polynomial} \mid f\left(z^{-1}\right) = \sigma f(z) \right\}$$

where $\mathfrak{sl}(2, \mathbb{C})$ is the Lie algebra of traceless complex 2×2 matrices and σ is an order 2 automorphism thereof. The Onsager algebra is not simple, the commutator subalgebra $[\mathfrak{O}, \mathfrak{O}]$ is a nontrivial ideal. The involved action of the 2 element group on $\mathbb{C}^*$ given by $z \mapsto z^{-1}$ has nontrivial stabilisers at $z = 1$ and $z = -1$. The automorphic Lie algebra evaluates to a lower dimensional algebra at these points than elsewhere. This is the more typical situation. In fact, many automorphic Lie algebras with $\mathfrak{sl}(2, \mathbb{C})$ as target Lie algebra are isomorphic to the Onsager algebra, even if the involved group actions are more complicated. Examples abound in [17–20].

Before the recent work of Lombardo, Veselov and the author [19], the groups Γ used in the construction of automorphic Lie algebras had been finite. The successful study of algebras of $SL(2, \mathbb{Z})$-equivariant matrices shows that the infinite discrete groups are in reach, and there is an optimism that we can start to understand automorphic Lie

algebras in various geometries. A first step in this direction, studied in this paper, is to simplify the situation to the case of continuous (Lie) groups (which contain the desired discrete groups) and their homogeneous spaces. This enables us to obtain subalgebras of the desired automorphic Lie algebras.

There are methods to compute invariants of discrete groups from other invariants, such as transvection [21], Rankin-Cohen brackets and modular derivatives [22]. These methods have been proven effective in the development of automorphic Lie algebras by Lombardo, Sanders and the author [23] and Lombardo, Veselov and the author [19], but they do have a weakness: you need an invariant to start with, before you can compute more of them. The subalgebras of automorphic Lie algebras that can be computed with the method of this paper provide a starting point for discrete methods and further study of automorphic Lie algebras in new geometries.

Various examples have been selected to illustrate the theory and inspire further development of geometric deep learning and automorphic Lie algebras.

2 Invariant sections of homogeneous vector bundles

Throughout this section, let X be a homogeneous space of a Lie group G, and V a representation of G. Our main practical contribution is this:

Elementary method to compute all G-equivariant maps $X \to V$.

1: Choose a base point $x_0 \in X$.
2: Determine $H := G_{x_0} = \{g \in G \mid g\, x_0 = x_0\}$.
3: Find a map $f : X \to G$ such that $f(x)x_0 = x$.
4: Compute a basis $\{v_1, \ldots, v_n\}$ for the invariants V^H.

The maps $x \mapsto f(x)v_1, \ldots, x \mapsto f(x)v_n$ constitute a basis for $M_G(X, V)$.

A few comments are in order. If $f : X \to G$ is such that $f(x)x_0 = x$, then $x \mapsto f(x)H$ is an isomorphism $X \to G/H$. The existence of this map is equivalent to the transitivity of the action of G on X. It is well known that such a map can be chosen to be smooth near any point, but in general, there is no such map that is globally smooth.[2]

It is not well known how to find a map like f and we are in fact not aware of a general method to tackle this problem. Luckily it is easy to do for many of the cases that appear in geometric deep learning applications. It is for instance trivial for Euclidean spaces $E^n \cong SE(n)/SO(n)$, and for spherical spaces $S^n \cong SO(n + 1)/SO(n)$, it can be solved with the Gram-Schmidt process (cf. Appendix A). For hyperbolic spaces, we did not find a general approach. The two-dimensional case is simple enough, and for the three-dimensional case, we present an ad hoc solution in Example 3.8.

The computation of the invariants V^H is simplified by moving to the Lie algebra level. That is, one determines a basis B for the Lie algebra $\mathfrak{h}$ of the Lie group H and

---

[2] There is a smooth map $f : X \to G$ such that $f(x)x_0 = x$ if and only if the H-bundle $G \to X$ is isomorphic to the trivial bundle $X \times H \to X$. Indeed, given such f, one can construct a map $G \to X \times H$ by $g \mapsto (gx_0, f(gx_0)^{-1}g)$. This is a bijection with inverse $(x, h) \mapsto f(x)h$, and hence an isomorphism of H-bundles. For the converse, suppose that $\phi : X \times H \to G$ is an isomorphism of H bundles, i.e. a smooth bijection such that $\phi(x, h)x_0 = x$. Then, we can define a smooth map $f : X \to G$ is such that $f(x)x_0 = x$ by $f(x) = \phi(x, 1)$.

solves the system of linear equations $hv = 0$ for $v \in V$ and $h \in B$. The complicating factor of this step is that the dimension of V can be very large in deep learning applications. Finzi et al. [24] approached this problem using singular value decompositions, and they made their software library publicly available.

Arriving at the solution, we can also see why we do not impose conditions on the maps we consider. It is the action of G that determines the properties of equivariant maps on homogeneous spaces. For example, if the action is smooth, so are the maps.

Before we prove that the method works, we give an example.

Example 2.1 (The hyperbolic plane H^2). *The hyperbolic plane H^2 can be realised as the space of complex numbers $\tau = x + iy$ with a positive imaginary part, $y > 0$. The group $\mathrm{SL}(2, \mathbb{R})$ acts transitively on H^2 by Möbius transformations*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau = \frac{a\tau + b}{c\tau + d}.$$

If we let $\mathrm{SL}(2, \mathbb{R})$ act on $\mathrm{End}(\mathbb{R}^2)$ by conjugation, then we can define the space of $\mathrm{SL}(2, \mathbb{R})$-equivariant maps $\mathrm{H}^2 \to \mathrm{End}(\mathbb{R}^2)$. This example fits in the setting above with $X = \mathrm{H}^2$, $G = \mathrm{SL}(2, \mathbb{R})$ and $V = \mathrm{End}(\mathbb{R}^2)$. To see this, we must check that $\mathrm{SL}(2, \mathbb{R})$ can take one point in H^2, say i, to an arbitrary point $\tau = x + iy \in \mathrm{H}^2$. A group element that does this is

$$f(\tau) = \begin{pmatrix} \sqrt{y} & x/\sqrt{y} \\ 0 & 1/\sqrt{y} \end{pmatrix}.$$

To carry out our method, we pick a point $i \in \mathrm{H}^2$ to take on the role of x_0, and determine that the subgroup H of $\mathrm{SL}(2, \mathbb{R})$ fixing i is $\mathrm{SO}(2)$. The map f defined above satisfies $f(\tau) \cdot i = \tau$, so the only part that is possibly challenging is already taken care of. It remains to compute a basis for $\mathrm{End}(\mathbb{R}^2)^{\mathrm{SO}(2)}$. For instance, we can take

$$v_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

We then arrive at a basis of the space of equivariant matrix-valued maps given by $\tau \mapsto f(\tau) v_i f(\tau)^{-1}$, $i = 1, 2$. This first basis element is given by the identity matrix, and the second reads

$$\tau = x + iy \mapsto \begin{pmatrix} \frac{x}{y} & -\frac{x^2+y^2}{y} \\ \frac{1}{y} & -\frac{x}{y} \end{pmatrix}.$$

As a bonus, we have in this case a real algebra structure, where the product is the pointwise matrix multiplication. Identifying the above matrix with the imaginary unit provides an isomorphism to the algebra of complex numbers.

We will now prove that this computational method gives the correct answers. Maps $X \to V$ can be identified with sections of the trivial vector bundle $X \times V$. We will generalise this setup to homogeneous vector bundles, following Cohen et al. [7] and

Aronsson [25], which are nontrivial bundles over X defined by a representation of H rather than G, known for their role in the celebrated Borel-Weil theorem. The invariant sections of homogeneous vector bundles are easily determined. These are then mapped onto the desired space of equivariant maps $X \to V$.

The following is a standard construction of a homogeneous vector bundle (see for instance [26, Section 7.4.1]). Let W be a H-module and define a right action of H on $G \times W$ by $(g, w)h = (gh, h^{-1}w)$. The orbit space $(G \times W)/H$ is denoted

$$G \times_H W.$$

An element of this space is of the form $[g, w] = \{(gh, h^{-1}w) \mid h \in H\}$. Together with the projection map

$$\pi : G \times_H W \to G/H, \quad [g, w] \mapsto gH$$

this is a W-bundle over G/H. It carries a G-action given by $g_1[g, w] = [g_1 g, w]$. This action restricts to linear maps between fibres and descends to a transitive action on G/H. The vector space $\Gamma(G/H, G \times_H W)$ of sections s becomes a G-module by the action

$$(g_1 s)(gH) = g_1 s(g_1^{-1} gH).$$

Notice that we have started with a representation W of H and ended up with a representation of G.

These representations are commonly reformulated in the language of maps. Let $M(X)$ denote the maps $X \to \mathbb{F}$, where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, and $M(X, V)$ the maps $X \to V$. The sections of the homogeneous vector bundle $G \times_H W$ can be identified with the maps

$$M(G, H, W) = \{f : G \to W \mid f(gh) = h^{-1} f(g), \ \forall h \in H, \forall g \in G\}$$

with G-action $(g_1 \cdot f)(g) = f(g_1^{-1} g)$. This identification is realised by sending a map $f \in M(G, H, W)$ to the section $s_f : gH \mapsto [g, f(g)]$. To write down the inverse of this map, we must first relate the fibre at eH of the homogeneous vector bundle to W. This is done with the H-intertwining bijection $\phi : \pi^{-1}(eH) \to W$ sending $[h, w]$ to hw. The inverse mapping from $\Gamma(G/H, G \times_H W)$ to $M(G, H, W)$ is then given by sending s to $f_s : g \mapsto \phi(g^{-1} s(gH))$. This identification intertwines the G-action.

The sections of the trivial vector bundle $G/H \times V$ can be identified with the maps

$$M(G, H, \mathbb{F}) \otimes V \cong \{f : G \to V \mid f(gh) = f(g), \ \forall h \in H, \forall g \in G\}$$

with G-action $(g_1 \cdot f)(g) = g_1 f(g_1^{-1} g)$ (where the field $\mathbb{F}$ is considered the trivial representation of H). This group action corresponds to the standard action on the tensor product $M \otimes V$ of two representations, given by $g(m \otimes v) = (gm) \otimes (gv)$, on the left-hand side of the isomorphism.

Remark 2.2 *If G is compact, then $M(G, H, W)$ is known as the representation of G induced by the representation W of H. For locally compact groups G such as the Lie*

groups we work with, induced representations have been defined as various subrepresentations of $M(G, H, W)$. Mackey required functions to be L^2 (square integrable) with respect to a well-chosen measure [27], and Moore worked with L^1 functions [28]. For our simpler objective, it is neither desirable nor necessary to make any of these restrictions.

The next result identifies the homogeneous vector bundles to the representation spaces where we look for automorphic Lie algebras and convolutional kernels for geometric deep learning.

Lemma 2.3 *Let V be a representation of G and W a representation of H. Let $\mathbb{F}$ denote the trivial representation of H. Then,*

$$\Phi : M(G, H, W) \to M(G, H, \mathbb{F}) \otimes V, \quad \Phi(f)(g) = gf(g)$$

is an isomorphism of G-representations if and only if W is the restriction of V to H.

Proof Let us denote the restriction of V to H by $\mathrm{Res}_H^G(V)$. Suppose that $W = \mathrm{Res}_H^G(V)$. Then $\Phi(f) \in M(G, H, \mathbb{F}) \otimes V$ whenever $f \in M(G, H, W)$ since $\Phi(f)(gh) = ghf(gh) = gf(g) = \Phi(f)(g)$, so that Φ is well defined. It has an inverse given by $\Phi^{-1}(\tilde{f})(g) = g^{-1}\tilde{f}(g)$, and the G-linearity follows from the computation

$$
\begin{aligned}
\Phi(g_1 \cdot f)(g) &= \Phi(f \circ g_1^{-1})(g) \\
&= gf(g_1^{-1}g) \\
&= g_1 g_1^{-1} g f(g_1^{-1} g) \\
&= g_1 \Phi(f)(g_1^{-1} g) \\
&= (g_1 \cdot \Phi(f))(g),
\end{aligned}
$$

which completes the proof of one direction.

Suppose now that Φ is an isomorphism of G-modules. We show first that $W = V$ as vector spaces. Let $G = \sqcup_{x \in X} g_x H$, and for each $w \in W$, define the map $g_x h \mapsto h^{-1} w$, which is an element of $M(G, H, W)$. Therefore, Φ sends it to an element $g_x h \mapsto g_x w \in V$. Taking $g_x = 1$, we find that $W \subset V$. In the other direction, we have an easier construction because for each $v \in V$, the constant map $g \mapsto v$ is in $M(G, H, \mathbb{F}) \otimes V$. The inverse of Φ sends this map to the map $g \mapsto g^{-1}v$ in $M(G, H, W)$. With $g = 1$, we see that $V \subset W$.

To see that the action of H on W is such that $W = \mathrm{Res}_H^G(V)$, we write out the equivariance condition for $g \mapsto f(g) = g^{-1}v$ in $M(G, H, W)$ while we keep track of the actions we are using with subscripts:

$$h \cdot_W (g^{-1} \cdot_V v) = h \cdot_W f(g) = f(gh^{-1}) = (gh^{-1})^{-1} \cdot_V v = h \cdot_V (g^{-1} \cdot_V v).$$

Putting $g = 1$, we get $h \cdot_W v = h \cdot_V v$ and we conclude that $W = \mathrm{Res}_H^G(V)$. $\square$

Lemma 2.4 to follow is a special instance of geometric Frobenius reciprocity. It can be found in the excellent but challenging book by Čap and Slovák, [29, Theorem 1.4.4],[3] where they use it to understand G-invariant geometric structures on G/H.

Lemma 2.4 *The space $M(G, H, W)^G$ corresponds to the space of constant maps $G \to W^H$. This provides an isomorphism $M(G, H, W)^G \cong W^H$.*

Proof If $f \in M(G, H, W)^G$, then $f(g) = g \cdot f(g) = f(1)$; hence, f is constant. If $f(g) = w$, then $hw = hf(g) = f(gh^{-1}) = w$; hence, $w \in W^H$ and f is a constant map $G \to W^H$. Vice versa, any constant map $G \to W^H$ is clearly contained in $M(G, H, W)^G$. □

We will use the isomorphisms $M(G, H, \mathbb{F}) \cong M(G/H)$ and $M(G, H, \mathbb{F}) \otimes V \cong M(G/H, V)$ and the resulting isomorphism

$$(M(G, H, \mathbb{F}) \otimes V)^G \cong M_G(X, V),$$

which is nothing more than a change of perspective. To see this, notice that $M(G, H, \mathbb{F})$ are the functions on G which are constant on H-cosets. These correspond to the function on G/H and are denoted $M(G/H)$. Likewise, elements of $M(G, H, \mathbb{F}) \otimes V$ can be considered maps $G \to V$ invariant on H-cosets, corresponding to maps $G/H \to V$ and denoted $M(G/H, V)$. When we recall how G acts on $M(G, H, \mathbb{F}) \otimes V$, we see that G-invariant elements of this space correspond to G-equivariant elements of $M(G/H, V)$.

Using this reformulation, we can combine Lemmas 2.3 and 2.4 into the following theorem.

Theorem 2.5 *Let V be a representation of G. The space of equivariant maps $M_G(G/H, V)$ corresponds to the algebra of maps $gH \mapsto gv$ where $v \in V^H$. This provides an isomorphism $M_G(G/H, V) \cong V^H$.*

Proof We map the constant maps $G \to V^H$ that constitute $M(G, H, W)^G$ by Lemma 2.4 onto $M_G(G/H, V)$ using the isomorphism Φ of Lemma 2.3. □

Theorem 2.5 reduces the infinite-dimensional problem of determining the invariants in $M(G/H, V)$ to the finite-dimensional problem to determine V^H. It finishes the proof of our method to compute equivariant convolutional kernels.

We apply the method to a second common example.

Example 2.6 (The sphere S^2). *The sphere is compact, and the equivariant maps on the sphere are therefore described in the existing literature of geometric deep learning (cf. [9], [10, Appendix E.4] and [3, Section 6]). This makes it a good example to illustrate our alternative approach to this problem.*

For convenience, we will work with the homogeneous space

$$\overline{\mathbb{C}} \cong \mathrm{PSU}(2)/\mathrm{PSU}(1)$$

---

[3] In [29, Theorem 1.4.4], the notation of sections $\Gamma(E)$ of a homogeneous bundle $E \to G/H$ is used in place of $M(G, H, W)$, and the fibre E_o at the 'base point' $o = H$ is used in place of W.

where $\mathrm{PSU}(2) = \mathrm{SU}(2)/\pm \mathrm{Id}$ *acts on the Riemann sphere* $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ *by Möbius transformations, rather than the more obvious* $\mathrm{S}^2 \cong \mathrm{SO}(3)/\mathrm{SO}(2)$. *Stereographic projection* $\mathrm{S}^2 \to \overline{\mathbb{C}}$ *takes one description to the other.*

The complex irreducible representations (irreps) of $\mathrm{SU}(2)$ *can be realised as*

$$V_n = \mathbb{C}\langle x^n, x^{n-1}y, x^{n-2}y^2, \ldots, y^n \rangle$$

where an element of $g \in \mathrm{SU}(2)$ *acts on a polynomial* $p \in V_n$ *by* $(g \cdot p)(x, y) = p(g^{-1}(x, y))$. *If* $p \in V_{2m}$, *then* $p(-x, -y) = p(x, y)$ *so that* V_{2m} *is a representation of* $\mathrm{SU}(2)/\pm \mathrm{Id} \cong \mathrm{SO}(3)$. *All irreps of* $SO(3)$ *are obtained this way. Any complex representation of a compact group such as* $\mathrm{SU}(2)$ *and* $\mathrm{SO}(3)$ *is a direct sum of irreps. See for instance* [26] *or another textbook on this topic.*

We carry out our elementary method to compute $M_{\mathrm{PSU}(2)}(\overline{\mathbb{C}}, V_{2m})$. *As base point, we choose* $x_0 := 0 \in \overline{\mathbb{C}}$. *The subgroup of* $\mathrm{PSU}(2)$ *stabilising* 0 *is*

$$H = \left\{ \pm \begin{pmatrix} a & 0 \\ 0 & \overline{a} \end{pmatrix} \mid a \in \mathbb{C}, |a| = 1 \right\}.$$

We need a map $f : \overline{\mathbb{C}} \to \mathrm{PSU}(2)$ *such that* $f(Z)$ *sends* 0 *to* Z. *For instance, we can define*[4]

$$f(Z) = \begin{cases} \pm \dfrac{1}{\sqrt{1+|Z|^2}} \begin{pmatrix} 1 & Z \\ -\overline{Z} & 1 \end{pmatrix} & \text{if } Z \neq \infty, \\[2ex] \pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \text{if } Z = \infty. \end{cases}$$

Our next step is to determine V^H. *In this case, we quickly find that* $V_{2m}^H = \mathbb{C}x^m y^m$. *Thus, we find the solution:* $M_{\mathrm{SU}(2)}(\overline{\mathbb{C}}, V_{2m})$ *is the one-dimensional space spanned by the map* $Z \mapsto f(Z)x^m y^m$ *where*

$$f(Z)x^m y^m = \begin{cases} \dfrac{(x-Zy)^m (\overline{Z}x+y)^m}{1+|Z|^2} & \text{if } Z \neq \infty, \\[2ex] (-1)^m x^m y^m & \text{if } Z = \infty. \end{cases}$$

It was explained in [9, 10] *that this problem can be solved using knowledge of spherical harmonics (see, e.g.* [30]*). With our approach, we do not need to know bases of spherical harmonics a priori. Such bases appear automatically in the computation, because the coefficients of an equivariant vector form a basis for the spherical*

---

[4] The map f is not continuous at $Z = \infty$, but this poses no problem. A continuous solution does not exist because $\mathrm{PSU}(2)$ and $\overline{\mathbb{C}} \times \mathrm{PSU}(1)$ are topologically inequivalent.

harmonics. To find these coefficients, we rewrite

$$\frac{(x-Zy)^m(\overline{Z}x+y)^m}{1+|Z|^2} = \frac{1}{1+|Z|^2}\left(\sum_{i=0}^{m}\binom{m}{i}x^{m-i}(-Zy)^i\right)\left(\sum_{j=0}^{m}\binom{m}{j}(\overline{Z}x)^{m-j}y^j\right)$$

$$= \frac{1}{1+|Z|^2}\sum_{i=0}^{m}\sum_{j=0}^{m}\binom{m}{i}\binom{m}{j}x^{2m-i-j}y^{i+j}(-Z)^i\overline{Z}^{m-j}$$

$$= \sum_{\ell=0}^{2m}x^{2m-\ell}y^{\ell}\left[\frac{1}{1+|Z|^2}\sum_{i+j=\ell}\binom{m}{i}\binom{m}{j}(-1)^iZ^i\overline{Z}^{m-j}\right]$$

and find in square brackets a basis for the spherical harmonics of degree m.

We end this section with the important specialisation to tangent vector fields. The tangent bundle of a homogeneous space can be studied in the framework of homogeneous vector bundles, using the isomorphism

$$T(G/H) \cong G \times_H \mathfrak{g}/\mathfrak{h} \tag{4}$$

of vector bundles over G/H described by Sharpe in [31, Chapter 4, Proposition 5.1]. The symbols $\mathfrak{g}$ and $\mathfrak{h}$ denote the Lie algebras of the groups G and H, respectively. Combining Lemma 2.4 and (4), we obtain

Corollary 2.7 *The space of G-invariant tangent vector fields on X is isomorphic to* $(\mathfrak{g}/\mathfrak{h})^H$. *If H is connected, then one can write* $(\mathfrak{g}/\mathfrak{h})^H$ *in purely Lie algebraic terms as* $N_{\mathfrak{g}}(\mathfrak{h})/\mathfrak{h}$ *where* $N_{\mathfrak{g}}(\mathfrak{h})$ *is the normaliser* $\{A \in \mathfrak{g} \mid [A, B] \in \mathfrak{h} \,\forall B \in \mathfrak{h}\}$ *of* $\mathfrak{h}$ *in* $\mathfrak{g}$.

When H is the trivial group, Corollary 2.7 reduces to the well-known fact that the space of left-invariant vector fields on a Lie group is isomorphic to its Lie algebra.

3 Fixed point subalgebras of homogeneous algebra bundles

In this section, we investigate what happens when the fibres of the homogeneous vector bundles have the structure of an algebra.

Throughout this section, we assume that G is a Lie group and H a closed subgroup, and that $\mathfrak{a}$ is an $\mathbb{F}$-algebra and H-representation and these two structures are compatible in the sense that $h(aa_1) = (ha)(ha_1)$ for all $h \in H$ and $a, a_1 \in \mathfrak{a}$.

Lemma 3.1 $M(G, H, \mathfrak{a})$ *has an algebra structure defined by* $(ff_1)(g) = f(g)f_1(g)$, *where* $f, f_1 \in M(G, H, \mathfrak{a})$ *and* $g \in G$. *The invariants* $M(G, H, \mathfrak{a})^G$ *form a subalgebra.*

We will take the algebra structure of $M(G, H, \mathfrak{a})$ into account and write $\mathfrak{M}(G, H, \mathfrak{a})$ to emphasise this. A subalgebra of invariants will be called an automorphic algebra. The automorphic algebra $M(G, H, \mathfrak{a})^G$ will be denoted $\mathfrak{A}(G, H, \mathfrak{a})$.

The linear isomorphisms obtained in the previous section are algebra isomorphisms in the setting of the current section. We state the results but omit the elementary proofs.

Theorem 3.2 *In the context of invariant sections of homogeneous bundles and equivariant maps, we have respectively the following results.*

(a) The automorphic algebra $\mathfrak{A}(G, H, \mathfrak{a})$ corresponds to the algebra of constant maps $G \to \mathfrak{a}^H$.

(b) If G acts on $\mathfrak{a}$ by algebra isomorphisms, then the automorphic algebra $\mathfrak{M}_G(G/H, \mathfrak{a})$ corresponds to the algebra of maps $gH \mapsto g\mathfrak{a}$ where $a \in \mathfrak{a}^H$.

This provides isomorphisms $\mathfrak{A}(G, H, \mathfrak{a}) \cong \mathfrak{M}_G(G/H, \mathfrak{a}) \cong \mathfrak{a}^H$ of algebras.

Remark 3.3 *The isomorphism class of $\mathfrak{A}(G, H, \mathfrak{a})$ does not depend on G. The full group only plays a role in the realisation of the algebra.*

Example 3.4 (Automorphic Lie algebras on the Riemann sphere). *Lombardo and Sanders classified automorphic Lie algebras on the Riemann sphere with target algebra $\mathfrak{sl}(2, \mathbb{C})$ [17] (elements of these Lie algebras are invariants of a finite group). Their analysis starts with the observation that the map*

$$\mathbb{C}^2 \to \mathfrak{sl}(2, \mathbb{C}), \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} xy & -x^2 \\ y^2 & -xy \end{pmatrix}$$

is $SL(2, \mathbb{C})$-equivariant with respect to left multiplication on $\mathbb{C}^2$ and the adjoint action (conjugation) on $\mathfrak{sl}(2, \mathbb{C})$.

We recover this observation with the approach of this paper, where we consider $\mathbb{C}^2 \setminus \{0\}$ as $SL(2, \mathbb{C})$-space. The stabiliser of $x_0 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is

$$A = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \mid b \in \mathbb{C} \right\} = e^{\mathbb{C}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}.$$

A map $f : \mathbb{C}^2 \setminus \{0\} \cong SL(2, \mathbb{C})$ with $f(x)x_0 = x$ is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x & -y/r^2 \\ y & x/r^2 \end{pmatrix}, \quad r^2 = x^2 + y^2.$$

and provides an isomorphism $\mathbb{C}^2 \setminus \{0\} \cong SL(2, \mathbb{C})/A$.

The space of A-invariants in $\mathfrak{sl}(2, \mathbb{C})$ is spanned by $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Hence, by Theorem 3.2, *the space of $SL(2, \mathbb{C})$-equivariant maps $\mathbb{C}^2 \setminus \{0\} \to \mathfrak{sl}(2, \mathbb{C})$ is spanned by*

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \mathrm{Ad}\left(\begin{pmatrix} x & -y/r^2 \\ y & x/r^2 \end{pmatrix} \right) \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -xy & x^2 \\ -y^2 & xy \end{pmatrix}.$$

Theorem 3.2 *is not necessary to compute this example. For example, if V_n is the irreducible representation of $SL(2, \mathbb{C})$ with the highest weight n, then we know from representation theory that we have the isomorphisms of $SL(2, \mathbb{C})$-modules*

$$\mathrm{End}(V_n) \cong V_n \otimes V_n^* \cong V_n \otimes V_n \cong V_{2n} \oplus V_{2n-2} \oplus \ldots \oplus V_0.$$

With Schur's lemma, we can then see that there is precisely one equivariant map $\mathbb{C}^2 \setminus \{0\} \to \mathfrak{gl}(V_n)$ of homogeneous order $0, 2, 4, \ldots, 2n$, up to a multiplicative constant, and these can be computed by brute force. It is however difficult to say anything about the algebra structure of this space based only on the representation theory of $\mathrm{SL}(2, \mathbb{C})$. Here, we find the value of Theorem 3.2 for this example, as it shows that the solution space is conjugate to the space of A-invariants.

If we define the nilpotent operator E for a representation $\rho_n : \mathrm{SL}(2, \mathbb{C}) \to \mathrm{GL}(V_n)$ by

$$E = \mathrm{d}\rho_n \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

then any power of E is fixed by A (because $\rho_n(A) = e^{\mathbb{C}E}$ and $\mathrm{Ad}(e^{cE})E^n = e^{c\,\mathrm{ad}(E)}E^n = E^n$). Moreover, the minimal polynomial of E is λ^{n+1}, thus

$$\mathfrak{gl}(V_n)^A = \mathbb{C}[E] \cong \mathbb{C}[T]/(T^{n+1}),$$

and we have an isomorphism

$$\mathbb{C}[T]/(T^{n+1}) \to \mathfrak{M}_{\mathrm{SL}(2,\mathbb{C})}(\mathbb{C}^2 \setminus \{0\}, \mathfrak{gl}(V_n))$$

of algebras given by

$$T \mapsto \left[\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \mathrm{Ad}\left(\rho_n \begin{pmatrix} x & -y/r^2 \\ y & x/r^2 \end{pmatrix} \right) E \right].$$

This cyclic algebra is of course a subalgebra of any automorphic Lie algebra defined by restricting the above actions of $\mathrm{SL}(2, \mathbb{C})$ to one of its subgroups.

If we replace V_n by a reducible representation U, then the algebra $\mathfrak{gl}(U)^A$ is no longer abelian.

In the context of convolutional networks, the target representations are spaces of real matrices. In the context of automorphic Lie algebras, we are interested in target spaces with algebraic structures. For this reason, we will emphasise the intersection of these special cases, where $W = \mathfrak{a} = \mathrm{End}(U)$. When we involve the algebra structure, we will write $\mathfrak{gl}(U)$ instead of $\mathrm{End}(U)$.

Due to the Skolem-Noether theorem, we know that any automorphism of $\mathfrak{gl}(U)$ is inner, i.e. of the form $m \mapsto rmr^{-1}$ for some $r \in \mathrm{GL}(U)$. Hence, any homomorphism $\rho : H \to \mathrm{Aut}(\mathfrak{gl}(U))$ is of the form $\rho(h)m = \sigma(h)m\sigma(h)^{-1}$ where σ is a homomorphism $H \to \mathrm{PGL}(U)$, that is, a projective representation of H.

Recall that the endomorphism ring of an irrep is a division ring due to Schur's lemma, and that a real division ring is isomorphic to either the real numbers $\mathbb{R}$, the complex numbers $\mathbb{C}$ or the quaternions $\mathbb{H}$ (this statement is known as the Frobenius theorem). Real irreps are said to be of a real type, complex type or quaternionic type accordingly. Combining these facts with the complete reducibility of representations of compact groups, we obtain the following corollary of Lemma 2.4.

Theorem 3.5 *Let K be a compact subgroup of G and let U be a real projective representation of K. Then, there is an isomorphism of algebras*

$$\mathfrak{A}(G, K, \mathfrak{gl}(U)) \cong \bigoplus_r \mathfrak{gl}(r, \mathbb{R}) \oplus \bigoplus_c \mathfrak{gl}(c, \mathbb{C}) \oplus \bigoplus_h \mathfrak{gl}(h, \mathbb{H})$$

where r ranges over the multiplicities of projective K-irreps in $U_\mathbb{C}$ of real type, c ranges over the multiplicities of conjugate pairs of projective K-irreps in $U_\mathbb{C}$ of complex type, and h ranges over the multiplicities of pairs of projective K-irreps in $U_\mathbb{C}$ of quaternionic (or symplectic) type.

We refer to Bourbaki [32, Ch. VIII, §7, Proposition 12] for a general method to determine the type of an irrep of a simple Lie algebra.

Example 3.6 (Common compact stabilisers). *All two- and three-dimensional model geometries in the sense of Thurston [33, §3.8] have compact stabilisers K with connected component isomorphic to the trivial group or* SO(2) *or* SO(3). *The hyperbolic, spherical and Euclidean geometries* H^n, S^n *and* E^n *have stabilisers* O(n) *in their isometry groups and stabilisers* SO(n) *in their orientation preserving isometry groups. These cases are therefore expected to be most common in applications and deserve special attention.*

The group SO(2) *has only one irrep of real type: the trivial representation. All nontrivial irreducible projective representations of* SO(2) *are of complex type. Hence,*

$$\mathfrak{A}(G, \mathrm{SO}(2), \mathfrak{gl}(U)) \cong \mathfrak{gl}(r, \mathbb{R}) \oplus \bigoplus_c \mathfrak{gl}(c, \mathbb{C})$$

where $r = \dim U^{\mathrm{SO}(2)}$, *cf.* Example 2.1.

The irreducible projective representations of SO(3) *correspond to the irreps of its double cover* SU(2). *There is precisely one irrep of* SU(2) *of any finite dimension. It is of real type when the dimension is odd (and the representation descends to a representation of* SO(3)*) and of quaternionic type if the dimension is even, so that*

$$\mathfrak{A}(G, \mathrm{SO}(3), \mathfrak{gl}(U)) \cong \bigoplus_r \mathfrak{gl}(r, \mathbb{R}) \oplus \bigoplus_h \mathfrak{gl}(h, \mathbb{H}).$$

Example 3.7 (Quaternions). *Let V_1 be the natural representation of* SU(2) *of 2 complex-dimensions. It is of quaternionic type. Hence, $V_1 \oplus V_1$ has a basis in which* SU(2) *is represented by matrices with real entries. Let U be the corresponding (4 real-dimensional) representation. Then, $\mathfrak{gl}(U)$ is a real representation of* $\mathrm{SU}(2)/\pm \mathrm{Id} \cong \mathrm{SO}(3)$ *(and of its extensions). We write down a basis for $\mathfrak{gl}(U)^{\mathrm{SO}(3)}$ to confirm it is isomorphic to the quaternions.*

With the embedding $\mathbb{C}^* \mapsto \mathrm{GL}(2, \mathbb{C})$ *given by* $a + bi \mapsto \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, *we construct real matrices*

$$\rho_U \left(\begin{pmatrix} a + bi & -c + di \\ c + di & a - bi \end{pmatrix} \right) = \begin{pmatrix} a & b & -c & d \\ -b & a & -d & -c \\ c & d & a & -b \\ -d & c & b & a \end{pmatrix}$$

for the representation $\rho_U : \mathrm{SU}(2) \to \mathrm{GL}(U)$. *The matrices that commute with the image of this representation are spanned by the identity matrix and*

$$I = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad K = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

These matrices satisfy $I^2 = J^2 = K^2 = IJK = -\mathrm{Id}$ *and provide the isomorphism* $\mathfrak{gl}(U)^{\mathrm{SO}(3)} \cong \mathbb{H}$.

We end this section with a detailed description of an interesting automorphic algebra illustrating Theorem 3.2. This example has a second goal of illustrating the fact that we can hardwire Lorentz group-equivariance into a convolutional network even if the group is not compact. This can be compared to the approach of Bogatskiy et al. [34] where a Lorentz group-equivariant network is used for the study of particle physics. The equivariance is incorporated as another learning task in this network, rather than being hardwired.

Example 3.8 (Hyperbolic space). *In this example, we will compute the algebra of maps from hyperbolic 3-space to the algebra of matrices* $\mathfrak{gl}(U)$ *from Example 3.7, which are equivariant under the Lorentz group, and exhibit its isomorphism to the quaternion algebra.*

Hyperbolic 3-space H^3 *will be identified with the hyperboloid model;*

$$\mathrm{H}^3 = \{(t, x, y, z) \in \mathbb{R}_{\geq 1} \times \mathbb{R}^3 \mid -t^2 + x^2 + y^2 + z^2 = -1\},$$

a component of a level set of the quadratic form

$$Q(t, x, y, z) = -t^2 + x^2 + y^2 + z^2.$$

The group of linear transformations preserving Q *is denoted* $\mathrm{O}(1, 3)$ *and corresponds to the group of automorphisms of* H^3. *The connected component* $G := \mathrm{SO}^+(1, 3)$ *containing the identity acts transitively on* H^3 *(as we will find below), and the stabiliser subgroups are isomorphic to* $\mathrm{SO}(3)$.

There is an isomorphism of real Lie groups between $\mathrm{SO}^+(1, 3)$ *and* $\mathrm{PSL}(2, \mathbb{C})$, *which can be understood using the Weyl representation that we now describe. The space of* 2×2 *hermitian matrices of determinant 1 has two connected components*

distinguished by the sign of the trace. Its component of matrices with positive trace is in one-to-one correspondence with Hyperbolic space by the map

$$\phi : (t, x, y, z) \mapsto A = \begin{pmatrix} t + z & x - iy \\ x + iy & t - z \end{pmatrix}.$$

The group $\mathrm{PSL}(2, \mathbb{C})$ *acts on this space by*

$$[g]A = gAg^*$$

where g^* is the conjugate transpose of g (it is easy to see that gAg^* is hermitian and has determinant 1. *To argue that its trace is positive, we can use the fact that* $\mathrm{SL}(2, \mathbb{C})$ *is connected, so that gAg^* is in the same connected component as* Id A Id* $=$ A). *When this action of* $\mathrm{PSL}(2, \mathbb{C})$ *is translated to* H^3 *using ϕ, one finds an isomorphism* $\mathrm{PSL}(2, \mathbb{C}) \to \mathrm{SO}^+(1, 3)$. *We will continue this example using the Weyl representation.*

We choose a base point $x_0 := (1, 0, 0, 0) \in \mathrm{H}^3$. Its stabiliser is $H := \mathrm{PSU}(2)$. To find a map $f : \mathrm{H}^3 \to \mathrm{PSL}(2, \mathbb{C})$ such that $f(x)x_0 = x$, we need to solve

$$gg^* = A \tag{5}$$

for $g \in \mathrm{SL}(2, \mathbb{C})$. Since A is hermitian, we can write it as

$$A = PDP^*$$

with $D = \mathrm{diag}(\lambda^+, \lambda^-)$ where

$$\lambda^{\pm} = t \pm \sqrt{t^2 - 1}$$

are the eigenvalues of A and $P \in \mathrm{SU}(2)$ (uniquely defined up to right multiplication by $\mathrm{diag}(\zeta, \zeta^*)$, $\zeta \in \mathrm{U}(1)$). *We then find a solution to* (5) *given by*

$$g = P\sqrt{D}P^*$$

where $\sqrt{D}$ is defined as $\mathrm{diag}(\sqrt{\lambda^+}, \sqrt{\lambda^-})$, *and hence an explicit isomorphism*

$$\mathrm{H}^3 \to \mathrm{PSL}(2, \mathbb{C})/\mathrm{PSU}(2), \quad (t, x, y, z) \mapsto \pm g\mathrm{PSU}(2).$$

The matrix g thus constructed reads

$$g = \frac{1}{\sqrt{2}\sqrt{t+1}} \begin{pmatrix} t + 1 + z & x - iy \\ x + iy & t + 1 - z \end{pmatrix}. \tag{6}$$

Let us write down a basis for the automorphic algebra $\mathfrak{M}_{\mathrm{SO}(1,3)^+}(\mathrm{H}^3, \mathfrak{gl}(U))$ *where U is the representation defined in* Example 3.7 *in which we computed a basis* $\{1, I, J, K\}$ *for* $\mathfrak{gl}(U)^{\mathrm{SO}(3)}$.

The group element (6) *acts on U with the matrix*

$$\rho_U(g) = \frac{1}{\sqrt{2}\sqrt{t+1}} \begin{pmatrix} t+1+z & 0 & x & -y \\ 0 & t+1+z & y & x \\ x & y & t+1-z & 0 \\ -y & x & 0 & t+1-z \end{pmatrix}$$

and on $\mathfrak{gl}(U)$ *by conjugation with this matrix. Applying g to I, J and K respectively yields*

$$\mathcal{I} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\mathcal{J} = \begin{pmatrix} x & y & -t-z & 0 \\ y & -x & 0 & t+z \\ t-z & 0 & -x & y \\ 0 & -t+z & y & x \end{pmatrix}$$

$$\mathcal{K} = \begin{pmatrix} -y & x & 0 & -t-z \\ x & y & -t-z & 0 \\ 0 & t-z & -y & -x \\ t-z & 0 & -x & y \end{pmatrix}$$

Together with the identity matrix, these form a basis for $\mathfrak{M}_{SO(1,3)^+}(H^3, \mathfrak{gl}(U))$ *due to Theorem* 3.2. *The quaternion algebra can thus be presented by matrix-valued functions on hyperbolic space equivariant under the Lorentz group.*

4 Discussion

We have described how the problem of computing all vector-valued equivariant maps on a homogeneous space can be reduced to finite-dimensional linear algebra with a geometric version of Frobenius reciprocity.

If G is compact, then one can take the alternative approach of Fourier analysis to find the invariants. Representations of compact groups are completely reducible, so that $L^2(X)$ and V decompose into irreps, and each pair of isomorphic irreps in these respective modules gives rise to one invariant in $L^2(X) \otimes V \cong L^2(X, V)$. Such an approach is presented by Lang and Weiler [10] and Lombardo, Sanders and the author in [23]. The hard part of this approach is the determination of the decomposition of $L^2(X)$. For important cases, the solution is readily available, e.g. the decomposition of $L^2(SO(2))$ is known from classical Fourier analysis and $L^2(SO(3)/SO(2))$ from harmonic analysis on the two-dimensional sphere.

The approach presented in this paper has the advantage that G can be noncompact, Fourier analysis is not needed, irreps of G need not be determined and neither do isotypical components of $L^2(X)$. This simple approach is possible because the domains

X considered here are G-orbits; hence, any equivariant map is determined by one value, contrary to the situation of automorphic Lie algebras where X/G is a Riemann surface. This also shows the disadvantage of our approach. If X is not a homogeneous space, there is more work to be done (see Appendix B).

For the study of algebra structures, our construction has a second advantage: it preserves the algebra structure. For this reason, there is no need to compute structure constants and find normal forms after the invariants are determined, which is one of the major challenges in the research on automorphic Lie algebras to date.

Another natural approach to solve a Lie group-invariance constraint is to start with differentiation and reduce it to a linear problem defined by the generators of the Lie algebra. Finzi et al. [24] used this approach in the context of geometric deep learning. The authors show how the resulting linear system can be effectively solved using the singular value decomposition in case the representation is finite-dimensional, and they made their software library publicly available. For the infinite-dimensional representation $M(G, H, W)$ considered in this paper, and commonly in geometric deep learning, the linearised problem is a system of partial differential equations. Solving this system is much harder than the method presented in this paper, and is therefore no longer practical. However, our results reduce the problem of finding G-invariants in the infinite-dimensional function space to finding H-invariants in the finite-dimensional representation W of H. For high-dimensional representations W, the approach and software of [24] can be used to compute W^H and complete the computation of $M(G, H, W)^G$.

We hope that the simplicity of the presented method, and its applicability to non-compact groups, will make it useful to the designers of convolutional neural networks with symmetries. Furthermore, we hope that this material shall serve as a starting point for the study of automorphic Lie algebras in unexplored geometries such as hyperbolic 3-space or Siegel upper-half space, and uncover new Lie structures.

Appendix A: The Gram-Schmidt process and $S^n \cong SO(n+1)/SO(n)$

In order to compute the space of G-equivariant maps from a homogeneous space $X \cong G/H$ to a representation V of G, we propose to find an explicit isomorphism $X \cong G/H$ first. That is, obtain a map $f : X \to G$ such that $f(x)x_0 = x$ for any $x \in X$, where x_0 is a 'base point' of our choice. As we mentioned in Section 2, we are not aware of a general method to construct such a map (in the mathematical literature, one is usually satisfied knowing that it exists, and there is no need for a construction), but for specific families of homogeneous spaces, there are constructions available. In this section, we solve the question for the spheres

$$S_r^n = \{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \mid x_1^2 + \ldots + x_{n+1}^2 = r^2\} \cong SO(n+1)/SO(n), \quad r > 0,$$

using the Gram-Schmidt process.

To warm up, we consider $S_r^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2\}$. When we pick the base point $x_0 = (r, 0)$, we need to find $f_1 : S_r^1 \to SO(2)$ (the subscript of f refers to

the dimension of the sphere) such that $f_1(x, y)(r, 0)^T = (x, y)^T$. That is,

$$f_1(x, y) = \begin{pmatrix} x/r & * \\ y/r & * \end{pmatrix}.$$

It remains to replace the $*$ by entries which ensure that $f_1(x, y) \in SO(2)$. In this case, we quickly find the solution

$$f_1(x, y) = \begin{pmatrix} x/r & -y/r \\ y/r & x/r \end{pmatrix}.$$

Going one dimension higher, we consider $S_r^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = r^2\}$ and pick the base point $x_0 = (r, 0, 0)$. Then, we need to construct an orthogonal matrix $f_2(x, y, z)$ with first column $(x/r, y/r, z/r)^T$. This time, it is harder to find a solution. We will return to this problem shortly.

The Gram-Schmidt process takes a set of linearly independent vectors $\{v_1, \ldots, v_k\}$ in a vector space with inner product $\langle \cdot, \cdot \rangle$ and returns a set of orthogonal vectors $\{u_1, \ldots, u_k\}$ with the same linear span. If we write the projection of v on a nonzero vector u as

$$P_u(v) = \frac{\langle u, v \rangle}{\langle u, u \rangle} u,$$

then the orthogonal set of vectors is given by

$$u_1 = v_1$$
$$u_2 = v_2 - P_{u_1}(v_2)$$
$$u_3 = v_3 - P_{u_1}(v_3) - P_{u_2}(v_3)$$
$$\vdots$$
$$u_k = v_k - \sum_{i=1}^{k-1} P_{u_i}(v_k).$$

To transform this orthogonal set $\{u_1, \ldots, u_k\}$ into an orthonormal set $\{e_1, \ldots, e_k\}$, one computes $e_i = u_i / \sqrt{\langle u_i, u_i \rangle}$.

Let us return to the problem of constructing an orthogonal matrix $f_2(x, y, z)$ with the first column $(x/r, y/r, z/r)^T$. A solution to this problem is not unique. Indeed, any solution can be multiplied on the right by a matrix of the form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

because this multiplication preserves the first column, orthogonality and the determinant. Using this wiggle room, we can ensure that f_2 is of the form

$$f_2(x, y, z) = \begin{pmatrix} x/r & * & 0 \\ y/r & * & * \\ z/r & * & * \end{pmatrix}.$$

Next, we construct orthogonal columns with the Gram-Schmidt process. Due to the zero on the top right of the matrix $f_2(x, y, x)$, we can save ourselves some work by computing columns from right to left. At the right, we start with a vector $(0, 1, 0)^T$. Instead of computing its component orthogonal to $(x/r, y/r, z/r)$, we use $(0, y/r, z/r)$ in order to preserve the first zero. That is, we compute $(0, 1, 0) - P_{(0, y/r, z/r)}(0, 1, 0) = \left(0, \frac{z^2}{y^2+z^2}, \frac{-yz}{y^2+z^2}\right)$ (which is only defined when $y^2 + z^2 > 0$). This vector normalises to $\left(0, \frac{z}{\sqrt{y^2+z^2}}, \frac{-y}{\sqrt{y^2+z^2}}\right)$ and is used for the right column of f_2.

For the middle column, we start with a vector $(1, 0, 0)^T$ which is already orthogonal to the right column, leaving us only to compute its component orthonormal to the left column: $(1, 0, 0) - P_{(x/r, y/r, z/r)}(1, 0, 0) = \left(\frac{y^2+z^2}{r^2}, -\frac{xy}{r^2}, -\frac{xz}{r^2}\right)$, which normalises to $\left(\frac{\sqrt{y^2+z^2}}{r}, \frac{-xy}{r\sqrt{y^2+z^2}}, \frac{-xz}{r\sqrt{y^2+z^2}}\right)$, resulting in the matrix

$$f_2(x, y, z) = \begin{pmatrix} \frac{x}{r} & \frac{\sqrt{y^2+z^2}}{r} & 0 \\ \frac{y}{r} & \frac{-xy}{r\sqrt{y^2+z^2}} & \frac{z}{\sqrt{y^2+z^2}} \\ \frac{z}{r} & \frac{-xz}{r\sqrt{y^2+z^2}} & \frac{-y}{\sqrt{y^2+z^2}} \end{pmatrix}, \quad y^2 + z^2 > 0.$$

By construction, we have $f_2(x, y, z) \in O(3)$, and we can check that the determinant of f is 1 hence $f_2(x, y, z) \in SO(3)$. It remains to define f_2 for the case $y^2 + z^2 = 0$, i.e. $x = \pm r$. Here, we can simply put $f_2(r, 0, 0) = \mathrm{Id}$ and $f_2(-r, 0, 0) = \mathrm{diag}(-1, -1, 1)$.

In much the same way, we can construct a map $f_n : S_r^n \to SO(n + 1)$ solving $f(x)x_0 = x$, where $x = (x_1, x_2, \ldots, x_{n+1})$. Generally, the equation leaves the freedom to multiply a solution $f(x)$ by an element of the stabiliser subgroup $G_{x_0} = H$ on the right. In our particular case we choose $x_0 = (r, 0, 0, \ldots, 0)^T$ and the stabiliser subgroup consists of block-diagonal matrices of the form $\mathrm{diag}(1, A)$ where $A \in SO(n)$. It can be used to ensure that the first row of $f_n(x_1, \ldots, x_{n+1})$ has the form $(x_1/r, *, 0, 0, \ldots, 0)$. Likewise, an element $\mathrm{diag}(1, 1, A) \in H$ where $A \in SO(n - 1)$ can be used to ensure the second row of $f_n(x_1, \ldots, x_{n+1})$ has the form $(x_2/r, *, *, 0, 0, \ldots, 0)$, without changing the first row. By continuing this argument for the whole sequence of subgroups $SO(n) \supset SO(n - 1) \supset \ldots \supset SO(2)$ of the

stabiliser, we can assume without loss of generality that f_n takes the form

$$f_n(x_1, \ldots, x_{n+1}) = \begin{pmatrix} x_1/r & * & 0 & 0 & 0 & \ldots & \ldots & 0 \\ x_2/r & * & * & 0 & 0 & \ldots & \ldots & 0 \\ x_3/r & * & * & * & 0 & \ldots & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \ddots & \ddots & \vdots \\ x_{n-1}/r & * & * & * & \ldots & \ldots & * & 0 \\ x_n/r & * & * & * & \ldots & \ldots & * & * \\ x_{n+1}/r & * & * & * & \ldots & \ldots & * & * \end{pmatrix}.$$

We will compute orthonormal columns for this matrix from right to left. For the right-most column, we start with the vector $(0, \ldots, 0, 1, 0)$ and obtain a norm 1 vector orthogonal to $(0, \ldots, 0, x_n/r, x_{n+1}/r)$ by computing $(0, \ldots, 0, 1, 0) - P_{(0,\ldots,0,x_n/r,x_n/r)}(0, \ldots, 0, 1, 0)$ and subsequently normalising to get

$$\left(0, \ldots, 0, \frac{x_{n+1}}{\sqrt{x_n^2+x_{n+1}^2}}, -\frac{x_n}{\sqrt{x_n^2+x_{n+1}^2}}\right).$$

Consider now the column of $f_n(x)$ at position $k + 1$, and assume all columns to its right fit the form above, and constitute an orthogonal set of vectors together with the vector $(0, \ldots, 0, x_k/r, \ldots, x_{n+1}/r)$. We can apply the Gram-Schmidt process for this set of vectors together with the vector $(0, \ldots, 0, 1, 0, \ldots, 0)^T$ where the nonzero entry is at position k, and see that all but one of the projections to be computed are zero, leaving us only to compute $(0, \ldots, 0, 1, 0, \ldots, 0) - P_{(0,\ldots,0,x_k/r,\ldots,x_{n+1}/r)}((0, \ldots, 0, 1, 0, \ldots, 0))$ which equals

$$\frac{1}{\sum_{i=k}^{n+1} x_i^2}(0, \ldots, 0, \sum_{i=k+1}^{n+1} x_i^2, -x_k x_{k+1}, -x_k x_{k+2}, \ldots, -x_k x_{n+1}).$$

It is convenient to introduce the notation

$$r_k = \sqrt{\sum_{i=k}^{n+1} x_i^2}, \quad k = 1, \ldots, n,$$

because the norm of the last vector is r_{k+1}/r_k, and its normalisation reads

$$\left(0, \ldots, 0, \frac{r_{k+1}}{r_k}, -\frac{x_k x_{k+1}}{r_k r_{k+1}}, \ldots, -\frac{x_k x_{n+1}}{r_k r_{k+1}}\right).$$

Notice that $r_n = \sqrt{x_n^2 + x_{n+1}^2} > 0$ implies $r_i > 0$ for $i = 1, \ldots, n$; hence, all these vectors are defined if and only if $x_n^2 + x_{n+1}^2 > 0$.

Using these vectors, for $k = 1, \ldots, n$, as columns for $f_n(x)$, we obtain for $x_n^2 + x_{n+1}^2 > 0$

$$f_n(x) = \begin{pmatrix} \frac{x_1}{r_1} & \frac{r_2}{r_1} & 0 & 0 & 0 & \cdots & \cdots & 0 \\ \frac{x_2}{r_1} & -\frac{x_1 x_2}{r_1 r_2} & \frac{r_3}{r_2} & 0 & 0 & \cdots & \cdots & 0 \\ \frac{x_3}{r_1} & -\frac{x_1 x_3}{r_1 r_2} & -\frac{x_2 x_3}{r_2 r_3} & \frac{r_4}{r_3} & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \ddots & \ddots & \vdots \\ \frac{x_{n-1}}{r_1} & -\frac{x_1 x_{n-1}}{r_1 r_2} & -\frac{x_2 x_{n-1}}{r_2 r_3} & -\frac{x_3 x_{n-1}}{r_3 r_4} & \cdots & \cdots & \frac{r_n}{r_{n-1}} & 0 \\ \frac{x_n}{r_1} & -\frac{x_1 x_n}{r_1 r_2} & -\frac{x_2 x_n}{r_2 r_3} & -\frac{x_3 x_n}{r_3 r_4} & \cdots & \cdots & -\frac{x_{n-1} x_n}{r_{n-1} r_n} & (-1)^{n+1} \frac{x_{n+1}}{r_n} \\ \frac{x_{n+1}}{r_1} & -\frac{x_1 x_{n+1}}{r_1 r_2} & -\frac{x_2 x_{n+1}}{r_2 r_3} & -\frac{x_3 x_{n+1}}{r_3 r_4} & \cdots & \cdots & -\frac{x_{n-1} x_{n+1}}{r_{n-1} r_n} & (-1)^n \frac{x_n}{r_n} \end{pmatrix}.$$

In the right column of this matrix, we added a factor $(-1)^{n+1}$ in order to have determinant 1. To see this, notice that the function $\det \circ f_n$ is continuous and takes values in the discrete set $\{1, -1\}$ due to orthonormality; hence, it is constant on the connected domain $\{x_n^2 + x_{n+1}^2 > 0\} = S^n \setminus S^{n-2}$, and we can compute the value of $\det \circ f_n$ by evaluating in one point. For instance, $\det f_n(0, \ldots, 0, r) = 1$. Thus, we have $f_n(x) \in SO(n+1)$ when $x_n^2 + x_{n+1}^2 > 0$.

We conclude using induction. If $x_n^2 + x_{n+1}^2 = 0$ then $(x_1, \ldots, x_{n-1}) \in S^{n-2}$ and we define $f_n(x)$ as the block-diagonal matrix $\text{diag}(f_{n-2}(x), 1, 1)$. Given that we already defined f_1 and f_2, this provides f_n for all $n \geq 1$.

It is worth noting that f_n is constant on lines emanating from the origin. That is, $f_n(\lambda x) = f_n(x)$ for any $\lambda \geq 0$ and any $x \neq 0$. All values of f_n are therefore achieved on S_1^n.

Appendix B: From homogeneous space to $\mathbb{R}^d$

This appendix concerns the application of this paper to geometric deep learning and is intended as a starting point for future research.

In geometric deep learning, one encounters a group $G \in GL(\mathbb{R}^d)$ and is interested in G equivariant maps from $\mathbb{R}^d$ to a representation V of G, as we explained in the introduction. In the body of this paper, we restrict the problem to G-orbits in $\mathbb{R}^d$. Whether this can be extended to the whole space depends on the particular group and particularly the quotient space $\mathbb{R}^d / G$.

We solve the problem for the spheres in $\mathbb{R}^d$ with Proposition B.1, where the quotient space is a manifold (appearing as $[0, \infty)$ in the proposition). A similar result for the case $d = 3$ can be found in [9]. Afterwards, we discuss the situation for hyperbolic spaces in $\mathbb{R}^d$, where the quotient space is not Hausdorff.

This illustrates that future research should not aim for a general solution but rather focus on the application. In geometric deep learning, the equivariant maps are used as

convolution kernels in an integral. Hence, it is harmless to exclude a set of measure zero, and one could restrict to a full measure subset of the quotient space $\mathbb{R}^d/G$ which is better behaved, allowing also solutions for hyperbolic spaces in $\mathbb{R}^d$.

In the setting of this appendix, we can choose what class of functions we want to study, unlike the situation for homogeneous spaces, where the function class is determined by the group actions. This appendix is written for continuous maps, but the proofs can be modified to work for other classes of maps.

Proposition B.1 *Let V be a continuous representation of* $\mathrm{SO}(d)$*,* $H \subset \mathrm{SO}(d)$ *the stabiliser of* $(1, 0, \ldots, 0) \in \mathbb{R}^d$*, and* $\{v_1, \ldots, v_m\}$ *a basis of* V^H*, extending a basis* $\{v_1, \ldots, v_{m'}\}$ *of* $V^{\mathrm{SO}(d)}$*. Then, the space of continuous* $\mathrm{SO}(d)$*-equivariant maps from* $\mathbb{R}^d$ *to V is given by*

$$C_{\mathrm{SO}(d)}\left(\mathbb{R}^d, V\right) = \left\{ x \mapsto \sum_{i=1}^{m} c_i(r) f_n(x) v_i \mid c_i \in C\left([0, \infty), \mathbb{R}\right), \; c_i(0) = 0 \text{ if } i > m' \right\}$$

where $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $r = \sqrt{x_1^2 + \ldots + x_d^2}$, and $f_n : \mathbb{R}^d \to \mathrm{SO}(d)$ the map defined in Appendix A *with $n = d - 1$, extended to the origin by $f_n(0) = \mathrm{Id}$.*

Proof The set on the right-hand side of the equation will be denoted RHS, that is

$$\mathrm{RHS} = \left\{ x \mapsto \sum_{i=1}^{m} c_i(r) f_n(x) v_i \mid c_i \in C\left([0, \infty), \mathbb{R}\right), \; c_i(0) = 0 \text{ if } i > m' \right\}.$$

We will show first that RHS is contained in $C_{\mathrm{SO}(d)}\left(\mathbb{R}^d \setminus \{0\}, V\right)$. To check that a map $x \mapsto \sum_{i=1}^{m} c_i(r) f_n(x) v_i$ in RHS is equivariant, we pick a generic element $g \in \mathrm{SO}(n)$ and compute

$$\sum_{i=1}^{m} c_i(gr) f_n(gx) v_i = \sum_{i=1}^{m} c_i(r) f_n(gx) v_i$$

$$= \sum_{i=1}^{m} c_i(r) g f_n(x) v_i$$

$$= g \sum_{i=1}^{m} c_i(r) f_n(x) v_i$$

where the second equality follows from Theorem 2.5.

The next thing to do is to check continuity (and we will find that the discontinuity of f_n is irrelevant). First, pick a nonzero $x \in \mathbb{R}^d$ and restrict to the sphere S^n containing x (and centred at the origin). There, $f_n(x) v_i$ is continuous due to equivariance and the properties of the group actions. Indeed, the limit $\lim_{y \to x} f_n(y) v_i$ with $y \in S^n$ can be written as $\lim_{g \to 1} f_n(gx) v_i$ with $g \in \mathrm{SO}(d)$ because the action of $\mathrm{SO}(d)$ on S^n is continuous and transitive. Equivariance equates this to $\lim_{g \to 1} (g f_n(x) v_i)$, and continuity of the action on V shows this limit is $\left(\lim_{g \to 1} g\right) f_n(x) v_i = f_n(x) v_i$.

To see that $c_i(r)f_n(x)v_i$ depends continuously on x at any nonzero x, recall that $f_n(\lambda x) = f_n(x)$ for any $\lambda > 0$, and c_i is continuous.

For continuity at $x = 0$, we use the fact that $v_i \in V^{SO(d)}$ when $i \leq m'$ which implies $c_i(r)f_n(x)v_i = c_i(r)v_i$ where continuity follows from continuity of c_i. Let now $i > m'$. Then, $\lim_{x\to 0} c_i(r)f_n(x)v_i = 0$ since c_i is continuous, $c_i(0) = 0$ and all entries of $f_n(x) \in SO(d)$ are absolutely bounded by 1. This limit corresponds to the value $c_i(0)f_n(0)v_i$. Thus, we see that the sum $\sum_{i=1}^m c_i(r)f_n(x)v_i$ is indeed continuous on $\mathbb{R}^d$, and RHS $\subset C_{SO(d)}(\mathbb{R}^d, V)$.

It remains to prove that the reverse inclusion $C_{SO(d)}(\mathbb{R}^d, V) \subset$ RHS holds. Let F be an element of $C_{SO(d)}(\mathbb{R}^d \setminus \{0\}, V)$. Then, $F(r, 0, \ldots, 0) \in V^H$ for any $r \geq 0$, since $hF(r, 0, \ldots, 0) = F(h(r, 0\ldots, 0)) = F(rh(1, 0\ldots, 0)) = F(r, 0\ldots, 0)$. Hence, we can expand $F(r, 0, \ldots, 0)$ in a basis of V^H, i.e. there exists functions c_i such that

$$F(r, 0, \ldots, 0) = \sum_{i=1}^m c_i(r)v_i.$$

The functions c_i must be continuous because $F(r, 0, \ldots, 0)$ depends continuously on r. Moreover, $c_i(0) = 0$ when $i > m'$ since $F(0) \in V^G$ due to equivariance. Define a map by $D(x) = F(x) - \sum_{i=1}^m c_i(\sqrt{x_1^2 + \ldots + x_d^2})f_n(x)v_i$. This map is equivariant by the first half of this proof, and it vanishes at $(r, 0, \ldots, 0)$ for any $r \geq 0$, since $f_n(r, 0\ldots, 0) = \mathrm{Id}$. Let $x \in \mathbb{R}^d \setminus \{0\}$ and $r = \sqrt{x_1^2 + \ldots + x_d^2}$. Then, there exists $g \in SO(d)$ such that $g(r, 0, \ldots, 0) = x$. Therefore, $D(x) = D(g(r, 0\ldots, 0)) = gD(r, 0, \ldots, 0) = g0 = 0$. That is, $F(x) = \sum_{i=1}^m c_i(r)f_n(x)v_i$ and $F \in$ RHS. $\square$

Hyperbolic space of dimension n can be embedded in $\mathbb{R}^d$, $d = n + 1$, as one component of the level set at -1 of the quadratic form

$$Q(t, x_1, \ldots, x_n) = -t^2 + x_1^2 + \ldots + x_n^2.$$

This is an orbit of the connected group $SO^+(1, n) \subset GL(d, \mathbb{R})$ defined by the preservation of Q.

We treated different realisations of two- and three-dimensional hyperbolic space in Examples 2.1 and 3.8, respectively. Now, we start with the 1-dimensional example. In that case, we have the quadratic form $Q(t, x) = -t^2 + x^2$ which is preserved by the group

$$SO^+(1, 1) = \left\{ \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} \mid \theta \in \mathbb{R} \right\}.$$

Each nonzero point $(t, 0)$ and $(0, x)$ is contained in a unique orbit of this group. This describes all but five of the orbits. The remaining orbits constitute the level set $\{Q = 0\} = \{t^2 = x^2\}$: four diagonal rays and the origin. The latter five orbits cannot be separated by open sets; hence, $\mathbb{R}^2/SO^+(1, 1)$ is not Hausdorff and therefore not a manifold, and we cannot do calculus on this space. However, if we take away the diagonals $t^2 = x^2$, the remaining quotient space $(\mathbb{R}^2 \setminus \{t^2 = x^2\})/SO^+(1, 1)$ is

homeomorphic to the union of four open intervals, which is a manifold. Given that the use of the equivariant maps in deep learning for this example is in an integral over $\mathbb{R}^2$, it is harmless to omit the set $\{t^2 = x^2\}$ of measure zero in $\mathbb{R}^2$ from the domain of these equivariant maps.

To describe continuous $\mathrm{SO}^+(1, 1)$-equivaiant maps $\mathbb{R}^2 \setminus \{Q \neq 0\} \to V$, for any continuous representation V of $\mathrm{SO}^+(1, 1)$, we can proceed as we did in Proposition B.1. The equivariant maps are given by the sums

$$\sum_{i=1}^{m} c_i(t, x) f(t, x) v_i$$

where each c_i is a continuous function, and each $c_i(t, x)$ depends only on the $\mathrm{SO}^+(1, 1)$-orbit containing (t, x), and where $f(t, x)$ is the matrix $\frac{1}{\sqrt{-Q}} \begin{pmatrix} t & x \\ x & t \end{pmatrix}$ when $Q < 0$ and $\frac{1}{\sqrt{Q}} \begin{pmatrix} x & t \\ t & x \end{pmatrix}$ when $Q > 0$.

In higher dimensions, the quotient space $\mathbb{R}^d / \mathrm{SO}^+(1, n)$ behaves similarly. The level set $\{Q = 0\} = \{t^2 = \langle x, x \rangle\}$ consists of 3 orbits of $\mathrm{SO}^+(1, n)$, with $t = 0$, $t > 0$ and $t < 0$. These orbits cannot be separated by open sets, and $\mathbb{R}^d / \mathrm{SO}^+(1, n)$ is not Hausdorff. At positive values of Q, the level set is one orbit of $\mathrm{SO}^+(1, n)$, and at negative values of Q, the level set consists of two orbits (hyperbolic spaces). The quotient of the full measure subset ($\mathbb{R}^d \setminus \{Q = 0\}$) by the action of $\mathrm{SO}^+(1, n)$ is homeomorphic to the union of three open intervals.

Declarations

References

1. Kutyniok, G.: The Mathematics of Artificial Intelligence. In: (2022). arXiv:2203.08890

2. Bronstein, M.M, et al.: Geometric deep learning: grids, groups, graphs, geodesics, and gauges. In: (2021). arXiv:2104.13478

3. Gerken, J.E., et al.: Geometric deep learning and equivariant neural networks. In: Artificial Intelligence Review, pp 1–58 (2023)

4. Weiler, M., et al.: Coordinate independent convolutional networks–isometry and gauge equivariant convolutions on Riemannian manifolds. In: (2021). arXiv:2106.06020

5. Cohen, T., et al.: Gauge equivariant convolutional networks and the icosahedral CNN. In: International Conference on Machine Learning. PMLR. pp. 1321–1330 (2019)

6. Cheng, M.C., et al.: Covariance in physics and convolutional neural networks. In: (2019). arXiv:1906.02481

7. Cohen, T.S., Geiger, M., Weiler, M.: A general theory of equivariant CNNs on homogeneous spaces. In: Advances in Neural Information Processing Systems. (Ed.) by Wallach, H., et al. vol. 32. Curran Associates, Inc., (2019)

8. Weiler, M., Cesa, G.: General E(2)-equivariant steerable CNNs. In: Advances in Neural Information Processing Systems **32** (2019)

9. Weiler, M., et al.: 3D steerable CNNs: learning rotationally equivariant features in volumetric data. In: Advances in Neural Information Processing Systems, pp. 10381–10392 (2018)

10. Lang, L., Weiler, M.: A Wigner-Eckart theorem for group equivariant convolution kernels. In: (2020). arXiv:2010.10952

11. Lombardo, S., Mikhailov, A.V.: Reductions of integrable equations: dihedral group. In: J. Phys. A **37**(31), 7727–7742 (2004). https://doi.org/10.1088/0305-4470/37/31/006. ISSN: 0305-4470

12. Lombardo, S., Mikhailov, A.V.: Reduction groups and automorphic Lie algebras. In: Comm. Math. Phys. **258**(1), 179–202 (2005). https://doi.org/10.1007/s00220-005-1334-5. Issn: 0010-3616

13. Kac, V.G.: Simple irreducible graded Lie algebras of finite growth. In: Izv. Akad. Nauk SSSR Ser. Mat. **32**, 1323–1367 (0373-2436) (1968)

14. Mathieu, O.: Classification of simple graded Lie algebras of finite growth. In: Invent. Math. **108**(3), 455–519 (1992). https://doi.org/10.1007/BF02100615. ISSN: 0020-9910

15. Onsager, L.: Crystal statistics. I. A two-dimensional model with an order-disorder transition. In: Phys. Rev. **65**(2), 117–149 (1944). ISSN: 0031-899X

16. Roan, S.-S.: Onsager's Algebra, Loop Algebra and Chiral Potts Model, Max-Planck-Institut für Mathematik (1991)

17. Lombardo, S., Sanders, J.A.: On the classification of automorphic Lie algebras. In: Comm. Math. Phys. **299**(3), 793–824 (2010). https://doi.org/10.1007/s00220-010-1092-x. ISSN:0010-3616

18. Bury, R.T., Mikhailov, A.V.: Automorphic Lie algebras and corresponding integrable systems. In: Differential geometry application, **74**, 101710 25 (2021). https://doi.org/10.1016/j.difgeo.2020.101710. ISSN: 0926-2245

19. Knibbeler, V., Lombardo, S., Veselov A.P.: Automorphic Lie algebras and modular forms. In: International Mathematics Research Notices. rnab376 (2022). https://doi.org/10.1093/imrn/rnab376. eprint: https://academic.oup.com/imrn/advance-article-pdf/doi/10.1093/imrn/rnab376/42467033/rnab376.pdf. ISSN: 1073-7928

20. Knibbeler, V., Lombardo, S., Oelen, C.: A classification of automorphic Lie algebras on complex tori. Accepted for publication in Proceedings of the Edinburgh Mathematical Society (2024)

21. Olver, P.J.: Classical invariant theory. vol. 44. London Mathematical Society Student Texts. Cambridge University Press, Cambridge, pp. xxii+280 (1999). https://doi.org/10.1017/CBO9780511623660. ISBN: 0-521-55821-2

22. Bruinier., J.H., et al.: The 1-2-3 of modular forms. Universitext. Lectures from the Summer School on Modular Forms and their Applications held in Nordfjordeid, June 2004, Edited by Kristian Ranestad. Springer-Verlag, Berlin, pp. x+266 (2008). https://doi.org/10.1007/978-3-540-74119-0. ISBN: 978-3-540-74117-6

23. Knibbeler, V., Lombardo, S., Sanders, J.A.: Higher-dimensional automorphic Lie algebras. In: Found. Comput. Math. **17**(4), 987–1035 (2017). https://doi.org/10.1007/s10208-016-9312-1. ISSN: 1615-3375

24. Finzi, M., Welling, M., Wilson, A.G.: A practical method for constructing equivariant multilayer perceptrons for arbitrary matrix groups. In: International Conference on Machine Learning. PMLR. pp. 3318–3328 (2021)

25. Aronsson, J.: Homogeneous vector bundles and G-equivariant convolutional neural networks. In: Sampling Theory, Signal Processing, and Data Analysis, **20**(2), 10 (2022)

26. Sepanski, M.R.: Compact Lie groups. vol. 235. Graduate Texts in Mathematics. Springer, New York, pp. xiv+198 (2007). https://doi.org/10.1007/978-0-387-49158-5. ISBN: 978-0-387-30263-8; 0-387-30263-8

27. Mackey, G.W.: Induced representations of locally compact groups. I. In: Ann. of Math. **2**(55), 101–139 (1952). https://doi.org/10.2307/1969423. ISSN: 0003-486X

28. Moore, C.C.: On the Frobenius reciprocity theorem for locally compact groups. In: Pacific J. Math. **12**, 359–365 (1962). ISSN: 0030-8730

29. Čap, A., Slovák, J.: Parabolic geometries. I. vol. 154. Mathematical Surveys and Monographs. Background and general theory. American Mathematical Society, Providence, RI, pp. x+628 (2009). https://doi.org/10.1090/surv/154. ISBN: 978-0-8218-2681-2

30. Toth, G.: Finite Möbius groups, minimal immersions of spheres, and moduli. Universitext. Springer-Verlag, New York, pp. xvi+317 (2002). https://doi.org/10.1007/978-1-4613-0061-8. ISBN: 0-387-95323-X

31. Sharpe, R.W.: Differential geometry. vol. 166. Graduate Texts in Mathematics. Cartan's generalization of Klein's Erlangen program, With a foreword by S. S. Chern. Springer-Verlag, New York, pp. xx+421 (1997). ISBN: 0-387-94732-9

32. Bourbaki, N.: Lie groups and Lie algebras. Chapters 7–9. Elements of Mathematics (Berlin). Translated from the 1975 and 1982 French originals by Andrew Pressley. Springer-Verlag, Berlin, pp. xii+434 (2005). ISBN: 3-540-43405-4

33. Thurston W.P.: Three-dimensional geometry and topology. vol. 1. vol. 35. Princeton Mathematical Series. (Ed.) by Levy, S., Princeton University Press, Princeton, NJ, pp. x+311 (1997). ISBN: 0-691-08304-5

34. Bogatskiy, A., et al.: Lorentz group equivariant neural network for particle physics. In: International Conference on Machine Learning. PMLR, pp. 992–1002 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.