

GLGENN: A Novel Parameter-Light Equivariant Neural Networks Architecture Based on Clifford Geometric Algebras

Ekaterina Filimoshina^{1,2} Dmitry Shirokov^{1,3}

Abstract

We propose, implement, and compare with competitors a new architecture of equivariant neural networks based on geometric (Clifford) algebras: Generalized Lipschitz Group Equivariant Neural Networks (GLGENN). These networks are equivariant to all pseudo-orthogonal transformations, including rotations and reflections, of a vector space with any non-degenerate or degenerate symmetric bilinear form. We propose a weight-sharing parametrization technique that takes into account the fundamental structures and operations of geometric algebras. Due to this technique, GLGENN architecture is parameter-light and has less tendency to overfitting than baseline equivariant models. GLGENN outperforms or matches competitors on several benchmarking equivariant tasks, including estimation of an equivariant function and a convex hull experiment, while using significantly fewer optimizable parameters.

1. Introduction

Equivariant neural networks are a class of neural networks that explicitly incorporate symmetries into their architecture, making them well-suited for tasks that inherently require equivariance to transformations with respect to some group's action (e.g., rotations, permutations, translations, etc.). Equivariant neural networks were introduced in [Cohen & Welling, 2016](#) and have since been extensively developed and applied in a wide range of tasks in computer and natural sciences. These applications include modeling dynamical systems ([Finzi et al., 2021](#)), particle physics ([Ruhe et al., 2023](#); [Finzi et al., 2021](#)), analyzing molecular properties and

protein structures ([Townshend et al., 2021](#); [Liu et al., 2024](#); [Satorras et al., 2021](#); [Pepe et al., 2024](#); [Fuchs et al., 2020](#)), estimating arterial wall-shear stress ([Brehmer et al., 2023](#)), processing tasks involving point clouds ([Thomas et al., 2018](#); [Fuchs et al., 2020](#)), motion capture ([Liu et al., 2024](#)), robotic planning ([Brehmer et al., 2023](#)), etc. Early works addressing equivariance with respect to pseudo-orthogonal transformations (pseudo-orthogonal groups) include [Cohen & Welling, 2016](#); [2017](#); [Weiler & Cesa, 2019](#); [Thomas et al., 2018](#); [Weiler et al., 2018](#); [Anderson et al., 2019](#); [Finzi et al., 2021](#).

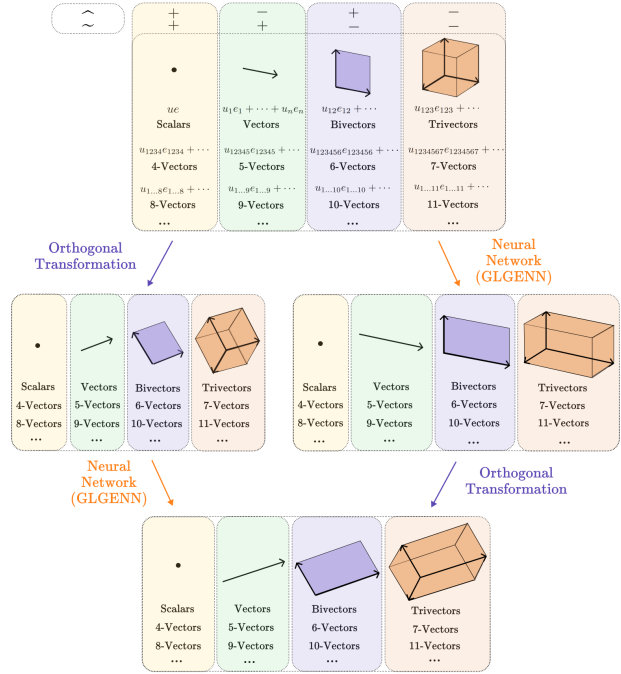


Figure 1. GLGENN is an architecture of neural networks equivariant with respect to any pseudo-orthogonal transformation. Inputs and outputs are represented as multivectors (elements of geometric algebras), which encode various geometric quantities such as scalars, vectors, oriented areas (bivectors) and volumes (trivectors), and higher-dimensional objects (4-vectors, etc.). GLGENN are parameter-light, since they operate in a unified manner across 4 fundamental subspaces of geometric algebras defined by the grade involution ($\sim$) and reversion ($\bar{\cdot}$); they processes geometric objects in groups with a step size of 4.

¹HSE University, Moscow, Russia ² Skolkovo Institute of Science and Technology, Moscow, Russia ³ Institute for Information Transmission Problems of the Russian Academy of Sciences, Moscow, Russia. Correspondence to: Ekaterina Filimoshina <filimoshinaek@gmail.com>, Dmitry Shirokov <dm.shirokov@gmail.com>.

This work introduces a novel neural network architecture that is equivariant with respect to any pseudo-orthogonal transformation, including rotations and reflections. Our approach is based on the well-known mathematical framework of (Clifford) geometric algebras (GAs), which have found applications in various scientific fields, including physics (Doran & Lasenby, 2003; Hestenes & Sobczyk, 1984), computer science (Bayro-Corrochano, 2019; Brandstetter et al., 2023; Dorst et al., 2002), engineering (Dorst et al., 2002), and other scientific fields. Works on applications of GAs in neural networks include Pearson & Bisset, 1994; Bayro-Corrochano & Buchholz, 1997; Buchholz & Sommer, 2008; Kuroe, 2011; Brandstetter et al., 2023.

In particular, GAs provide an efficient and elegant representation of pseudo-orthogonal transformations via spin groups, Lipschitz groups (sometimes also called Clifford groups in the literature), and twisted adjoint representations. Specifically, for any pseudo-orthogonal matrix, there exists a corresponding element of the Lipschitz group in GAs that induces the same transformation of vectors through the twisted adjoint action. The transition from matrix formalism to GA formalism is highly beneficial and inspiring for applications, as it introduces a rich set of operations that are naturally defined within the GA framework.

The idea of leveraging GAs to construct pseudo-orthogonal group equivariant neural networks was first proposed by the Amsterdam Machine Learning Lab (University of Amsterdam). Their pioneering work Ruhe et al., 2023 introduced Clifford Group Equivariant Neural Networks (CGENN) and served as the foundation for many subsequent research works. Brehmer et al., 2023 propose Geometric Algebra Transformer (GATr), which incorporates GA into ordinary transformer architecture and outperforms traditional non-geometric baselines in N -body modeling and robotic planning. Zhdanov et al., 2024 introduce Clifford-Steerable Convolutional Neural Networks based on CGENN approach, demonstrating superior performance in tasks related to fluid dynamics and relativistic electrodynamics forecasting. Liu et al., 2024 introduce Clifford Group Equivariant Simplicial Message Passing Networks, a method for steerable $E(n)$ -equivariant message passing on simplicial complexes. Pepe et al., 2024 use CGENN to predict protein coordinates to estimate the 3D structure of a protein.

The above mentioned papers demonstrate the importance of Lipschitz groups in GA as a fundamental tool for building expressive state-of-the-art equivariant neural networks. However, one major challenge in the design of GA-based equivariant neural networks is overparameterization. This issue often leads to a tendency to overfit, particularly in cases where the training dataset is small – a common scenario in natural science applications where such networks are typically employed. Additionally, the excessive number of parameters results in inefficient training times. The goal

of this work is to present a novel, parameter-light equivariant neural network architecture, which balances between expressiveness of CGENN and parameter efficiency. We call this architecture *Generalized Lipschitz Group Equivariant Neural Networks (GLGENN)*, see Fig. 1. Our approach introduces a new parameter-light parametrization technique, which enables our model to either outperform or match existing equivariant models while significantly reducing the number of trainable parameters. The key idea is to design a novel weight-sharing approach (Lecun et al., 1998) for GA-based neural networks that respects the fundamental algebraic structures of GAs, thereby improving efficiency without sacrificing expressive power.¹

The key contributions are as follows:

- **Introduction of generalized Lipschitz groups.** We introduce and study a new class of Lie groups in arbitrary geometric algebra, which are related to pseudo-orthogonal groups and useful for construction of equivariant neural networks.
- **Design and implementation of GLGENN.** We construct a novel, parameter-light architecture of pseudo-orthogonal groups equivariant neural networks based on geometric algebras. Code is available at <https://github.com/katyafilimoshina/glgen>
- **Superior performance.** GLGENN achieve state-of-the-art performance on benchmark equivariant tasks with significantly fewer trainable parameters.

The paper is organized as follows. Section 2 provides all necessary definitions related to GAs and equivariant neural networks. In Section 3, we present our main theoretical results and introduce and study generalized Lipschitz groups. Section 4 applies these results for GLGENN construction. In Section 5, we evaluate GLGENN through experiments. All the mathematical details and proofs can be found in the appendix.

¹Application of the generalized Lipschitz groups instead of ordinary Lipschitz groups allows to achieve parameter efficiency. The generalized Lipschitz groups are important because they preserve the four fundamental subspaces of GAs under the significant operation of the twisted adjoint representation. We prove that equivariance of a mapping w.r.t. these groups, as well as the ordinary Lipschitz groups, implies its orthogonal groups equivariance. The key distinction is that the generalized Lipschitz groups contain ordinary Lipschitz groups as subgroups. As a result, the set of operations equivariant w.r.t. the generalized Lipschitz groups is a subset of the set of operations equivariant w.r.t. ordinary Lipschitz groups. This reduction in the number of ‘degrees of freedom’ encourages us to parametrize operations in layers in a more ‘economic’ way (there is a smaller number of parameters that we can place). Specifically, in all GLGENN layers, we employ such equivariant operations as projections of inputs-multivectors onto the four fundamental subspaces of GAs mentioned above, whereas CGENN relies on projections onto the subspaces of fixed grades. GLGENN and CGENN layers parameterize linear combinations, products, and normalizations of the corresponding projections.

2. Theoretical Background

2.1. Geometric (Clifford) Algebras

Let us consider (Clifford) geometric algebra (GA) (Hestenes & Sobczyk, 1984; Lounesto, 1997; Porteous, 1995) $\mathcal{Cl}(V) = \mathcal{Cl}_{p,q,r}$, $p + q + r = n \geq 1$, over a vector space V with a symmetric bilinear form $\mathbf{b}$ and the corresponding quadratic form q , where V can be real $\mathbb{R}^{p,q,r}$ or complex $\mathbb{C}^{p+q,0,r}$. We use $\mathbb{F}$ to denote the field of real numbers $\mathbb{R}$ in the first case and the field of complex numbers $\mathbb{C}$ in the second case. In this work, we consider both the case of the non-degenerate GAs $\mathcal{Cl}_{p,q} := \mathcal{Cl}_{p,q,0}$, $r = 0$, and the case of the degenerate GAs $\mathcal{Cl}_{p,q,r}$, $r \neq 0$. We use Λ_r to denote the subalgebra $\mathcal{Cl}_{0,0,r}$, which is the Grassmann (exterior) algebra (Doran & Lasenby, 2003; Lounesto, 1997). A well-known example of geometric algebra is the spacetime algebra $\mathcal{Cl}_{1,3} = \mathcal{Cl}(\mathbb{R}^{1,3})$, associated with the Minkowski space $\mathbb{R}^{1,3}$, which plays a central role in relativistic physics. For the readers' convenience, in Appendix B, we illustrate all key theoretical concepts introduced in this section with examples in the setting of $\mathcal{Cl}_{1,3}$.

The identity element of $\mathcal{Cl}_{p,q,r}$ is denoted by $e \equiv 1$, the generators are denoted by e_a , $a = 1, \dots, n$. The generators satisfy the following conditions: $e_a e_b + e_b e_a = 2\eta_{ab} e$ for any $a, b = 1, \dots, n$, where $\eta = (\eta_{ab})$ is the diagonal matrix with p times $+1$, q times -1 , and r times 0 on the diagonal in the real case $\mathcal{Cl}(\mathbb{R}^{p,q,r})$ and $p + q$ times $+1$ and r times 0 on the diagonal in the complex case $\mathcal{Cl}(\mathbb{C}^{p+q,0,r})$.

Let us consider the subspaces $\mathcal{Cl}_{p,q,r}^k$ of fixed grades $k = 0, \dots, n$. Their elements are linear combinations of the basis elements $e_{a_1 \dots a_k} := e_{a_1} \dots e_{a_k}$, $a_1 < \dots < a_k$. The grade-0 subspace can be denoted by $\mathcal{Cl}^0$ without the lower indices p, q, r , since it does not depend on the GA's signature. We have $\mathcal{Cl}_{p,q,r}^k = \{0\}$ for $k < 0$ and $k > n$. We can represent any element (multivector) $U \in \mathcal{Cl}_{p,q,r}$ as a sum of $n + 1$ elements $U = \langle U \rangle_0 + \dots + \langle U \rangle_n$, $\langle U \rangle_k \in \mathcal{Cl}_{p,q,r}^k$, $k = 0, \dots, n$. We call any operation of the form

$$U \mapsto \sum_{k=0}^n \lambda_k \langle U \rangle_k, \quad \lambda_k = \pm 1, \quad (1)$$

a *conjugation operation* in $\mathcal{Cl}_{p,q,r}$. Consider conjugation operations called *grade involution* and *reversion*. The grade involute of an element $U \in \mathcal{Cl}_{p,q,r}$ is denoted by $\hat{U}$ and the reversion is denoted by $\tilde{U}$, and they are defined for an arbitrary $U \in \mathcal{Cl}_{p,q,r}$ as

$$\hat{U} := \sum_{k=0}^n (-1)^k \langle U \rangle_k, \quad \tilde{U} := \sum_{k=0}^n (-1)^{\frac{k(k-1)}{2}} \langle U \rangle_k. \quad (2)$$

The composition of grade involution and reversion is called *Clifford conjugation*. The Clifford conjugate of $U \in \mathcal{Cl}_{p,q,r}$ is denoted by $\hat{\tilde{U}}$. The grade involution defines the *even* $\mathcal{Cl}_{p,q,r}^{(0)}$ and *odd* $\mathcal{Cl}_{p,q,r}^{(1)}$ subspaces:

$$\mathcal{Cl}_{p,q,r}^{(l)} := \{U \in \mathcal{Cl}_{p,q,r} : \hat{U} = (-1)^l U\}, \quad l = 0, 1. \quad (3)$$

Table 1. Signs of the projections of $U = \langle U \rangle_0 + \langle U \rangle_1 + \langle U \rangle_2 + \langle U \rangle_3 \in \mathcal{Cl}_{p,q,r}$ for the grade involution ($\hat{}$), reversion ($\tilde{}$), and Clifford conjugation ($\hat{\tilde{}}$) acting on it.

$\mathcal{Cl}_{p,q,r}^k$	$k = 0$	$k = 1$	$k = 2$	$k = 3$
$\hat{}$	+	−	+	−
$\tilde{}$	+	+	−	−
$\hat{\tilde{}}$	+	−	−	+

We can represent any element $U \in \mathcal{Cl}_{p,q,r}$ as a sum $U = \langle U \rangle_{(0)} + \langle U \rangle_{(1)}$, $\langle U \rangle_{(l)} \in \mathcal{Cl}_{p,q,r}^{(l)}$, $l = 0, 1$. We use angle brackets $\langle \cdot \rangle_{(l)}$ to denote the operation of *projection* of multivectors and sets onto $\mathcal{Cl}_{p,q,r}^{(l)}$. The grade involution and reversion define four subspaces $\mathcal{Cl}_{p,q,r}^{\bar{0}}$, $\mathcal{Cl}_{p,q,r}^{\bar{1}}$, $\mathcal{Cl}_{p,q,r}^{\bar{2}}$, and $\mathcal{Cl}_{p,q,r}^{\bar{3}}$ (they are called the *subspaces of quaternion types* 0, 1, 2, and 3 respectively in Shirokov, 2012a,b):

$$\mathcal{Cl}_{p,q,r}^{\bar{k}} := \{U \in \mathcal{Cl}_{p,q,r} : \hat{U} = (-1)^k U, \\ \tilde{U} = (-1)^{\frac{k(k-1)}{2}} U\}, \quad k = 0, 1, 2, 3. \quad (4)$$

In other words, $\mathcal{Cl}_{p,q,r}^{\bar{k}} := \mathcal{Cl}_{p,q,r}^k \oplus \mathcal{Cl}_{p,q,r}^{k+4} \oplus \mathcal{Cl}_{p,q,r}^{k+8} \oplus \dots$ for $k = 0, 1, 2, 3$. A discussion on the significance of the grade involution, reversion, and these subspaces is provided in Appendix C. Note that the GA $\mathcal{Cl}_{p,q,r}$ can be represented as a direct sum of the subspaces $\mathcal{Cl}_{p,q,r}^k$, $k = 0, 1, 2, 3$, and viewed as $\mathbb{Z}_2 \times \mathbb{Z}_2$ -graded algebra with respect to the commutator and anticommutator (Shirokov, 2018). We can represent any element $U \in \mathcal{Cl}_{p,q,r}$ as a sum of 4 elements: $U = \langle U \rangle_{\bar{0}} + \langle U \rangle_{\bar{1}} + \langle U \rangle_{\bar{2}} + \langle U \rangle_{\bar{3}}$, where $\langle U \rangle_{\bar{m}} \in \mathcal{Cl}_{p,q,r}^{\bar{m}}$, $m = 0, 1, 2, 3$. The action of the grade involution, reversion, and Clifford conjugation on a multivector U of the form above is summarized in Table 1. Note that $\mathcal{Cl}_{p,q,r}^{\bar{k}} = \mathcal{Cl}_{p,q,r}^k$, $k = 0, 1, 2, 3$, in the cases $n \leq 3$.

2.2. Lipschitz Groups and Twisted Adjoint Representations

We use the upper index $\times$ to denote the subset $H^\times$ of all invertible elements of any set H . Let us consider the *adjoint representation* ad acting on the group of all invertible elements $\text{ad} : \mathcal{Cl}_{p,q,r}^\times \rightarrow \text{Aut}(\mathcal{Cl}_{p,q,r})$ as $T \mapsto \text{ad}_T$, where $\text{ad}_T : \mathcal{Cl}_{p,q,r} \rightarrow \mathcal{Cl}_{p,q,r}$:

$$\text{ad}_T(U) := TUT^{-1}, \quad U \in \mathcal{Cl}_{p,q,r}, \quad T \in \mathcal{Cl}_{p,q,r}^\times. \quad (5)$$

Also let us consider the *twisted adjoint representation*. This notion is introduced by Atiyah, Bott, and Shapiro (Atiyah et al., 1964) in a particular case, and there are two ways how to generalize it (see motivation in Appendix E). The first approach (Choi et al., 2002; Harvey, 1990; Lundholm & Svensson, 2009) is to define it as the operation $\check{\text{ad}}$ acting on the group of all invertible elements $\check{\text{ad}} : \mathcal{Cl}_{p,q,r}^\times \rightarrow \text{Aut}(\mathcal{Cl}_{p,q,r})$ as $T \mapsto \check{\text{ad}}_T$ with $\check{\text{ad}}_T : \mathcal{Cl}_{p,q,r} \rightarrow \mathcal{Cl}_{p,q,r}$:

$$\check{\text{ad}}_T(U) := \hat{T}UT^{-1}, \quad U \in \mathcal{Cl}_{p,q,r}, \quad T \in \mathcal{Cl}_{p,q,r}^\times. \quad (6)$$

The second approach (Helmstetter & Micali, 2008; Knus, 1991; Walpuski, 2022) is to define the twisted adjoint representation as the operation $\tilde{\text{ad}} : \mathcal{C}\ell_{p,q,r}^\times \rightarrow \text{Aut}(\mathcal{C}\ell_{p,q,r})$ acting as $T \mapsto \tilde{\text{ad}}_T$ with $\tilde{\text{ad}}_T : \mathcal{C}\ell_{p,q,r} \rightarrow \mathcal{C}\ell_{p,q,r}$:

$$\tilde{\text{ad}}_T(U) := T\langle U \rangle_{(0)}T^{-1} + \widehat{T}\langle U \rangle_{(1)}T^{-1} \quad (7)$$

for any $U \in \mathcal{C}\ell_{p,q,r}$ and $T \in \mathcal{C}\ell_{p,q,r}^\times$. The representation $\tilde{\text{ad}}$ has the following properties, which we prove in Lemma E.2 and apply in construction of equivariant layers in Section 3.3. The properties of ad and $\tilde{\text{ad}}$, are considered in Lemma E.2 as well.

Lemma 2.1. *Let $W \in \mathcal{C}\ell_{p,q,r}^\times$, $U, V \in \mathcal{C}\ell_{p,q,r}$, and $T \in (\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times})\Lambda_r^\times$, where the notation $(\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times})\Lambda_r^\times := \{ab \mid a \in \mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times}, b \in \Lambda_r^\times\}$ stands for the product of two groups. Then ad_W satisfies:*

$$\tilde{\text{ad}}_W(U + V) = \tilde{\text{ad}}_W(U) + \tilde{\text{ad}}_W(V), \quad \tilde{\text{ad}}_W(c) = c \quad (8)$$

for any $c \in \mathcal{C}\ell^0$. Moreover, $\tilde{\text{ad}}_T$ satisfies multiplicativity:

$$\tilde{\text{ad}}_T(UV) = \tilde{\text{ad}}_T(U)\tilde{\text{ad}}_T(V). \quad (9)$$

Consider the Lipschitz groups $\tilde{\Gamma}_{p,q,r}^1$ (Benn & Tucker, 1987; Lounesto, 1997; Ablamowicz, 1986; Brooke, 1978; 1980):

$$\tilde{\Gamma}_{p,q,r}^1 := \{T \in \mathcal{C}\ell_{p,q,r}^\times : \tilde{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^1)\}, \quad (10)$$

where $\tilde{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^1) = \widehat{T}\mathcal{C}\ell_{p,q,r}^1T^{-1} \subseteq \mathcal{C}\ell_{p,q,r}^1$. These groups are often considered in the literature in the case of the non-degenerate geometric algebra $\mathcal{C}\ell_{p,q}$. The definition (10) straightforwardly generalizes the common definition to the case of arbitrary $\mathcal{C}\ell_{p,q,r}$. The well-known spin groups (see details in Appendix E) in the non-degenerate (Benn & Tucker, 1987; Lounesto, 1997; Porteous, 1995) and degenerate (Ablamowicz, 1986; Crumeyrolle, 1990; Brooke, 1978; 1980; Dereli et al., 2010) cases, which have various applications in physics, are normalized subgroups of $\tilde{\Gamma}_{p,q,r}^1$ defined using the following norm functions $\psi, \chi : \mathcal{C}\ell_{p,q,r} \rightarrow \mathcal{C}\ell_{p,q,r}$ of the GA elements:

$$\psi(U) := \tilde{U}U, \quad \chi(U) := \widehat{U}U. \quad (11)$$

2.3. Equivariant Mappings and Neural Networks

Suppose G is a group and $\circ_X$ and $\circ_Y$ are its actions on the sets X and Y respectively. A function (network) $L : X \rightarrow Y$ is called G -equivariant (see, e.g., Bredon, 1972; Pitts, 2013; Yarotsky, 2022) iff it commutes with these actions:

$$L(g \circ_X x) = g \circ_Y L(x), \quad \forall g \in G, \quad \forall x \in X. \quad (12)$$

The definition means that we get the same output if we transform the input to the neural network or transform the output. We prove several general statements about equivariance with respect to an arbitrary group in Appendix G.

3. Theoretical Results

3.1. Generalized Lipschitz Groups

In this section, we introduce and study the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^k$, $k = 0, 1, 2, 3$, in the case of arbitrary

degenerate or non-degenerate geometric algebra $\mathcal{C}\ell_{p,q,r}$. For GLGENN, we are mainly interested in $\tilde{\Gamma}_{p,q,r}^1$, however other groups $\tilde{\Gamma}_{p,q,r}^k$, $k = 0, 2, 3$, are necessary to prove the main statements about $\tilde{\Gamma}_{p,q,r}^1$ and serve as auxiliary tools. The (ordinary) Lipschitz groups $\tilde{\Gamma}_{p,q,r}^1$ (10) preserve the subspace $\mathcal{C}\ell_{p,q,r}^1$ of the first grade under the twisted adjoint representation $\tilde{\text{ad}}$ (6). The generalized Lipschitz groups preserve the subspaces $\mathcal{C}\ell_{p,q,r}^k$, $k = 0, 1, 2, 3$, determined by the grade involution and reversion (4), under the same representation $\tilde{\text{ad}}$:

$$\tilde{\Gamma}_{p,q,r}^k := \{T \in \mathcal{C}\ell_{p,q,r}^\times : \tilde{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^k) \subseteq \mathcal{C}\ell_{p,q,r}^k\}. \quad (13)$$

The groups $\tilde{\Gamma}_{p,q,r}^k$, $k = 0, 1, 2, 3$, can be considered as generalizations of Lipschitz groups because of the following theorem.

Theorem 3.1. *The (ordinary) Lipschitz groups are subgroups of the generalized Lipschitz groups and coincide with some of them in the low-dimensional case:*

$$\tilde{\Gamma}_{p,q,r}^1 \subseteq \tilde{\Gamma}_{p,q,r}^1 \subseteq \tilde{\Gamma}_{p,q,r}^k, \quad k = 0, 1, 2, 3, \quad \forall n; \quad (14)$$

$$\tilde{\Gamma}_{p,q,r}^1 = \tilde{\Gamma}_{p,q,r}^1, \quad n \leq 4. \quad (15)$$

Proof. The proof is based on the equivalent definitions of $\tilde{\Gamma}_{p,q,r}^k$, $k = 0, 1, 2, 3$, and is provided in Appendix F (see Corollary F.4, Remark F.6, and Theorem F.7). $\square$

In this work, we construct neural networks that are equivariant with respect to the action $\tilde{\text{ad}}$ of the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^1$ (13). To prove the main statements for the design of the layers, we need Theorem 3.2 about these groups.

Consider the following groups preserving the subspaces $\mathcal{C}\ell_{p,q,r}^k$, $k = 0, 1, 2, 3$, under the adjoint representation ad (5) and twisted adjoint representation $\tilde{\text{ad}}$ (6) respectively:

$$\Gamma_{p,q,r}^k := \{T \in \mathcal{C}\ell_{p,q,r}^\times : \text{ad}_T(\mathcal{C}\ell_{p,q,r}^k) \subseteq \mathcal{C}\ell_{p,q,r}^k\}, \quad (16)$$

$$\tilde{\Gamma}_{p,q,r}^k := \{T \in \mathcal{C}\ell_{p,q,r}^\times : \tilde{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^k) \subseteq \mathcal{C}\ell_{p,q,r}^k\}, \quad (17)$$

where $\text{ad}_T(\mathcal{C}\ell_{p,q,r}^k) = T\mathcal{C}\ell_{p,q,r}^kT^{-1}$ and $\tilde{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^k) = \widehat{T}\mathcal{C}\ell_{p,q,r}^kT^{-1}$. Further, we show that these groups can be defined in an equivalent way, using only the norm functions ψ and χ (11) applied in the theory of spin groups.

Let us introduce the families of Lie groups $\mathcal{Q}_{p,q,r}^1$, $\mathcal{Q}_{p,q,r}^2$, $\mathcal{Q}_{p,q,r}^3$, and $\mathcal{Q}_{p,q,r}^0$:

$$\mathcal{Q}_{p,q,r}^k := \{T \in \mathcal{C}\ell_{p,q,r}^\times : \widehat{T}T, \widehat{\widehat{T}}T \in \mathbb{Z}_{p,q,r}^{k \times k}\}, \quad (18)$$

$$\mathcal{Q}_{p,q,r}^0 := \{T \in \mathcal{C}\ell_{p,q,r}^\times : \widehat{T}T, \widehat{\widehat{T}}T \in \mathbb{Z}_{p,q,r}^{4 \times 4}\}, \quad (19)$$

and $\check{\mathcal{Q}}_{p,q,r}^1, \check{\mathcal{Q}}_{p,q,r}^2, \check{\mathcal{Q}}_{p,q,r}^3, \check{\mathcal{Q}}_{p,q,r}^0$:

$$\check{\mathcal{Q}}_{p,q,r}^k := \{T \in \mathcal{C}\ell_{p,q,r}^\times : \widehat{T}T, \widehat{\widehat{T}}T \in \check{\mathbb{Z}}_{p,q,r}^{k \times k}\}, \quad (20)$$

$$\check{\mathcal{Q}}_{p,q,r}^0 := \{T \in \mathcal{C}\ell_{p,q,r}^\times : \widehat{T}T, \widehat{\widehat{T}}T \in \langle \mathbb{Z}_{p,q,r}^4 \rangle_{(0)}^\times\}, \quad (21)$$

with $k = 1, 2, 3$, where the sets

$$Z_{p,q,r}^m := \{X \in \mathcal{C}\ell_{p,q,r}^m : XV = VX, \forall V \in \mathcal{C}\ell_{p,q,r}^m\}$$

are the *centralizers* of the subspaces $\mathcal{C}\ell_{p,q,r}^m$ in $\mathcal{C}\ell_{p,q,r}$, and

$$\check{Z}_{p,q,r}^m := \{X \in \mathcal{C}\ell_{p,q,r} : \hat{X}V = VX, \forall V \in \mathcal{C}\ell_{p,q,r}^m\}$$

are the *twisted centralizers* of $\mathcal{C}\ell_{p,q,r}^m$ in $\mathcal{C}\ell_{p,q,r}$. These sets are studied in detail in Filimoshina & Shirokov, 2024b. We provide explicit forms of $Z_{p,q,r}^m$ and $\check{Z}_{p,q,r}^m$, $m \leq 4$, in Remark D.1.

Theorem 3.2. *In degenerate and non-degenerate geometric algebras $\mathcal{C}\ell_{p,q,r}$, we have for $k = 1, 2, 3$,*

$$\Gamma_{p,q,r}^{\bar{k}} = Q_{p,q,r}^{\bar{k}}, \quad \Gamma_{p,q,r}^{\bar{0}} = Q_{p,q,r}^{\bar{0}}, \quad (22)$$

$$\tilde{\Gamma}_{p,q,r}^{\bar{1}} = \check{Q}_{p,q,r}^{\bar{1}} \subseteq \tilde{\Gamma}_{p,q,r}^{\bar{3}} = \check{Q}_{p,q,r}^{\bar{3}}, \quad (23)$$

$$\tilde{\Gamma}_{p,q,r}^{\bar{2}} = \check{Q}_{p,q,r}^{\bar{2}}, \quad \tilde{\Gamma}_{p,q,r}^{\bar{0}} = \check{Q}_{p,q,r}^{\bar{0}}. \quad (24)$$

Moreover, for $m = 0, 1, 2, 3$,

$$\Gamma_{p,q,r}^{\bar{1}} \subseteq \Gamma_{p,q,r}^{\bar{m}} \subseteq \Gamma_{p,q,r}^{\bar{0}}, \quad (25)$$

$$\tilde{\Gamma}_{p,q,r}^{\bar{m}} \subseteq \tilde{\Gamma}_{p,q,r}^{\bar{0}}, \quad \tilde{\Gamma}_{p,q,r}^{\bar{1}} \subseteq \tilde{\Gamma}_{p,q,r}^{\bar{2}}. \quad (26)$$

Proof. The proof relies on the relation between the preservation of the subspaces $\mathcal{C}\ell_{p,q,r}^{\bar{k}}$, $k = 0, 1, 2, 3$, under $\tilde{\text{ad}}$ and the values of the norm functions ψ and χ (11). It is provided in Appendix F (see Theorems F.1, F.2 and Remark F.3). $\square$

The generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^{\bar{k}}$ (13), $k = 0, 1, 2, 3$, are related to the groups $\Gamma_{p,q,r}^{\bar{k}}$ (16) and $\check{\Gamma}_{p,q,r}^{\bar{k}}$ (17):

$$\tilde{\Gamma}_{p,q,r}^{\bar{k}} = \begin{cases} \check{\Gamma}_{p,q,r}^{\bar{k}}, & k = 1, 3, \\ \Gamma_{p,q,r}^{\bar{k}}, & k = 0, 2, \end{cases} \quad (27)$$

since $\tilde{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^{\bar{k}}) = \text{ad}_T(\mathcal{C}\ell_{p,q,r}^{\bar{k}})$ in the cases $k = 0, 2$ and $\tilde{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^{\bar{k}}) = \check{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^{\bar{k}})$ in the cases $k = 1, 3$.

As a corollary of Theorem 3.2 and (27), the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^{\bar{k}}$ (13) satisfy

$$\tilde{\Gamma}_{p,q,r}^{\bar{1}} = \check{Q}_{p,q,r}^{\bar{1}} \subseteq \tilde{\Gamma}_{p,q,r}^{\bar{2}} = Q_{p,q,r}^{\bar{2}} \subseteq \tilde{\Gamma}_{p,q,r}^{\bar{0}} = Q_{p,q,r}^{\bar{0}}, \quad (28)$$

$$\tilde{\Gamma}_{p,q,r}^{\bar{1}} \subseteq \tilde{\Gamma}_{p,q,r}^{\bar{3}} = \check{Q}_{p,q,r}^{\bar{3}}. \quad (29)$$

Note that the group $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ can be regarded as the most significant among the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^{\bar{k}}$, $k = 0, 1, 2, 3$, because it contains the (ordinary) Lipschitz group as a subgroup and is itself a subgroup of all other generalized Lipschitz groups (see Theorem 3.1 and (28)-(29)). We prove that its elements are of a special form:

Theorem 3.3. *We have $\tilde{\Gamma}_{p,q,r}^{\bar{1}} \subseteq (\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times}) \times \Lambda_r^\times$.*

Proof. The proof is based on the equivalent definition of the group $(\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times}) \times \Lambda_r^\times$ and is provided in Theorem F.5. $\square$

The **key takeaways** from this section for constructing the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ -equivariant layers are as follows: (1) the ordinary Lipschitz groups are subgroups of

the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^{\bar{k}}$ (Theorem 3.1), (2) elements of the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ preserve not only the subspace $\mathcal{C}\ell_{p,q,r}^{\bar{1}}$, but also all other subspaces $\mathcal{C}\ell_{p,q,r}^{\bar{m}}$, $m = 0, 2, 3$, under the twisted adjoint representation $\tilde{\text{ad}}$ (see (28)-(29)), and (3) elements of $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ have a special form (Theorem 3.3).

3.2. Pseudo-orthogonal Groups, Complex Orthogonal Groups, and Generalized Lipschitz Groups Equivariance

Let us consider the degenerate and non-degenerate pseudo-orthogonal group (in the real case $V = \mathbb{R}^{p,q,r}$) or complex orthogonal group (in the complex case $V = \mathbb{C}^{p+q,0,r}$) denoted by $O(V, q)$, which is the Lie group of all linear transformations of an n -dimensional vector space V that leave invariant a quadratic form q of signature (p, q, r) if V is real and $(p + q, 0, r)$ if V is complex (see Appendix H):

$$O(V, q) := \{\Phi : V \rightarrow V : \Phi \text{ is linear, invertible, } q(\Phi(v)) = q(v), \forall v \in V\}. \quad (30)$$

When considering both the real and complex cases, we refer to the group $O(V, q)$ as the orthogonal group. Consider the following subgroup of the orthogonal group, which leaves invariant the *radical subspace* $\Lambda_r^1 := \mathcal{C}\ell_{0,0,r}^1$:

$$O_{\Lambda_r^1}(V, q) := \{\Phi \in O(V, q) : \Phi|_{\Lambda_r^1} = \text{id}_{\Lambda_r^1}\}. \quad (31)$$

We have the following relation between the Lipschitz group $\tilde{\Gamma}_{p,q,r}^1$ and the corresponding restricted orthogonal group $O_{\Lambda_r^1}(V, q)$:

$$\tilde{\text{ad}}^1 : \tilde{\Gamma}_{p,q,r}^1 / \Lambda_r^\times \cong O_{\Lambda_r^1}(V, q), \quad (32)$$

where $\tilde{\text{ad}}^1 : \mathcal{C}\ell_{p,q,r}^\times \rightarrow \text{Aut}(\mathcal{C}\ell_{p,q,r})$ is $\tilde{\text{ad}}$ restricted to V (a real $\mathbb{R}^{p,q,r}$ or complex $\mathbb{C}^{p+q,0,r}$ vector space, which can be identified with the space of vectors $\mathcal{C}\ell_{p,q,r}^1$, see page 3), i.e. it acts as $T \mapsto \tilde{\text{ad}}_T^1 : \mathcal{C}\ell_{p,q,r}^1 \rightarrow \mathcal{C}\ell_{p,q,r}^1$ defined as $\tilde{\text{ad}}_T^1(v) := \hat{T}vT^{-1}$ for any vector $v \in \mathcal{C}\ell_{p,q,r}^1$. From (32), we have that for any $T \in \tilde{\Gamma}_{p,q,r}^1$ there exists $\Phi \in O_{\Lambda_r^1}(V, q)$ such that $\tilde{\text{ad}}_T^1 = \Phi$ and vice versa. We prove (32) in Theorem H.10. In the particular case $\mathcal{C}\ell_{p,q}$, the statement (32) is well-known (see, for example, Benn & Tucker, 1987), and the similar statements are proved in Crumeyrolle, 1990; Ruhe et al., 2023.

The relation (32) implies that $O_{\Lambda_r^1}(V, q)$ acts on the whole $\mathcal{C}\ell_{p,q,r}$ in a well-defined way (Ruhe et al., 2023). Namely, for an arbitrary element $x = \sum_i u_i v_{i,1} \cdots v_{i,k_i} \in \mathcal{C}\ell_{p,q,r}$ with $v_{i,j} \in V$, $u_i \in \mathbb{F}$, for any $T \in \tilde{\Gamma}_{p,q,r}^1$ with the corresponding $\Phi = \tilde{\text{ad}}_T^1$, we have

$$\tilde{\text{ad}}_T(x) = \sum_i u_i \tilde{\text{ad}}_T^1(v_{i,1}) \cdots \tilde{\text{ad}}_T^1(v_{i,k_i}) \quad (33)$$

$$= \sum_i u_i \Phi(v_{i,1}) \cdots \Phi(v_{i,k_i}) =: \Phi(x). \quad (34)$$

The following theorem discusses the relation between $O_{\Lambda_r^1}(V, q)$ - and $\tilde{\Gamma}_{p,q,r}^1$ -equivariance.

Theorem 3.4. *If a mapping $f : \mathcal{C}\ell_{p,q,r} \rightarrow \mathcal{C}\ell_{p,q,r}$ is equivariant with respect to any group H that contains the Lipschitz group $\tilde{\Gamma}_{p,q,r}^1$ as a subgroup, then f is equivariant with respect to the corresponding orthogonal group. That is, if $\text{ad}_T(f(x)) = f(\text{ad}_T(x))$ for any $T \in H$ and $x \in \mathcal{C}\ell_{p,q,r}$, then $f(\Phi(x)) = \Phi(f(x))$ for any $\Phi \in O_{\Lambda_r^1}(V, q)$ and $x \in \mathcal{C}\ell_{p,q,r}$, where Φ acts on x in a sense (34).*

Proof. The proof relies on the direct relation between $\tilde{\Gamma}_{p,q,r}^1$ and $O_{\Lambda_r^1}(V, q)$ -equivariance (see Theorem H.11). $\square$

Since $\tilde{\Gamma}_{p,q,r}^1 \subseteq \tilde{\Gamma}_{p,q,r}^1$ by Theorem 3.1, all generalized Lipschitz group $\tilde{\Gamma}_{p,q,r}^1$ -equivariant mappings are restricted orthogonal group $O_{\Lambda_r^1}(V, q)$ -equivariant as well.

3.3. Generalized and Ordinary Lipschitz Groups Equivariant Mappings

This section proves that several mappings are generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^1$ -equivariant.

Theorem 3.5. *Let $T \in (\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times})\Lambda_r^\times$ and $F \in \mathbb{F}[T_1, \dots, T_l]$ be a polynomial in l variables with coefficients in $\mathbb{F}$. Consider l multivectors $x_1, \dots, x_l \in \mathcal{C}\ell_{p,q,r}$. We have the following equivariance property with respect to the group $(\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times})\Lambda_r^\times$:*

$$\tilde{\text{ad}}_T(F(x_1, \dots, x_l)) = F(\tilde{\text{ad}}_T(x_1), \dots, \tilde{\text{ad}}_T(x_l)). \quad (35)$$

Proof. This fact follows directly from additivity and multiplicativity of $\tilde{\text{ad}}_T$ for $T \in (\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times})\Lambda_r^\times$, which are proved in Lemma 2.1. $\square$

Note that the statement that all polynomials are $\tilde{\Gamma}_{p,q,r}^1$ -equivariant, which follows as a corollary of Theorem 3.5, is well-known (Ruhe et al., 2023).

Theorem 3.6. *For any $x \in \mathcal{C}\ell_{p,q,r}$ and $m = 0, 1, 2, 3$, we have the following equivariant property with respect to the generalized Lipschitz groups:*

$$\tilde{\text{ad}}_T(\langle x \rangle_{\overline{m}}) = \langle \tilde{\text{ad}}_T(x) \rangle_{\overline{m}}, \quad \forall T \in \tilde{\Gamma}_{p,q,r}^1. \quad (36)$$

Proof. Suppose $T \in \tilde{\Gamma}_{p,q,r}^1$ and $x = x_{(0)} + x_{(1)} \in \mathcal{C}\ell_{p,q,r}$, where $x_{(0)} := \langle x \rangle_{(0)}$ and $x_{(1)} := \langle x \rangle_{(1)}$. For any $m = 0, 1, 2, 3$, we get

$$\langle \tilde{\text{ad}}_T(x) \rangle_{\overline{m}} = \langle T x_{(0)} T^{-1} + \hat{T} x_{(1)} T^{-1} \rangle_{\overline{m}} \quad (37)$$

$$= \langle \sum_{i=0,2} T \langle x \rangle_i T^{-1} + \sum_{i=1,3} \hat{T} \langle x \rangle_i T^{-1} \rangle_{\overline{m}} \quad (38)$$

$$= \sum_{i=0,2} \langle T \langle x \rangle_i T^{-1} \rangle_{\overline{m}} + \sum_{i=1,3} \langle \hat{T} \langle x \rangle_i T^{-1} \rangle_{\overline{m}} \quad (39)$$

$$= \begin{cases} \langle T \langle x \rangle_{\overline{m}} T^{-1} \rangle_{\overline{m}} = T \langle x \rangle_{\overline{m}} T^{-1}, & m \text{ is even,} \\ \langle \hat{T} \langle x \rangle_{\overline{m}} T^{-1} \rangle_{\overline{m}} = \hat{T} \langle x \rangle_{\overline{m}} T^{-1}, & m \text{ is odd,} \end{cases} \quad (40)$$

where in (39) we apply linearity of projection $\langle \rangle_{\overline{m}}$, and in (40), we use Theorem 3.2 and the notes below it. $\square$

In Theorem 3.7, we prove that several conjugation operations (1) in $\mathcal{C}\ell_{p,q,r}$ are generalized Lipschitz group $\tilde{\Gamma}_{p,q,r}^1$ -equivariant. Note that it does not hold for any conjugation operation. However, we prove that any conjugation operation is (ordinary) Lipschitz group $\tilde{\Gamma}_{p,q,r}^1$ -equivariant in Theorem 3.8.

Theorem 3.7. *The grade involution, reversion, and Clifford conjugation are $\tilde{\Gamma}_{p,q,r}^1$ -equivariant: $\tilde{\text{ad}}_T(\hat{x}) =$*

$$\widehat{\tilde{\text{ad}}_T(x)}, \quad \tilde{\text{ad}}_T(\tilde{x}) = \widetilde{\tilde{\text{ad}}_T(x)}, \quad \tilde{\text{ad}}_T(\widehat{\tilde{x}}) = \widehat{\tilde{\text{ad}}_T(x)} \text{ for any } T \in \tilde{\Gamma}_{p,q,r}^1 \text{ and } x \in \mathcal{C}\ell_{p,q,r}.$$

Proof. Since the grade involution, reversion, and Clifford conjugate of an element $x \in \mathcal{C}\ell_{p,q,r}$ are linear combinations of the projections onto the subspaces $\mathcal{C}\ell_{p,q,r}^{\overline{m}}$, $m = 0, 1, 2, 3$, (see Table 1) which are $\tilde{\Gamma}_{p,q,r}^1$ -equivariant, we get the statement by Lemma G.3 and Theorem 3.6. $\square$

Theorem 3.8. *Any conjugation operation in $\mathcal{C}\ell_{p,q,r}$ is Lipschitz group $\tilde{\Gamma}_{p,q,r}^1$ -equivariant:*

$$\tilde{\text{ad}}_T \left(\sum_{k=0}^n \lambda_k \langle x \rangle_k \right) = \sum_{k=0}^n \lambda_k \langle \tilde{\text{ad}}_T(x) \rangle_k, \quad \lambda_k = \pm 1,$$

for any $T \in \tilde{\Gamma}_{p,q,r}^1$ and $x \in \mathcal{C}\ell_{p,q,r}$.

Proof. By definition (1), any conjugation operation is a linear combination of projections onto the subspaces $\mathcal{C}\ell_{p,q,r}^k$, $k = 0, \dots, n$, which are $\tilde{\Gamma}_{p,q,r}^1$ -equivariant (see Corollary 3.3 Ruhe et al., 2023): $\tilde{\text{ad}}_T(\langle x \rangle_k) = \langle \tilde{\text{ad}}_T(x) \rangle_k$, for any $T \in \tilde{\Gamma}_{p,q,r}^1$ and $x \in \mathcal{C}\ell_{p,q,r}$. Then, we get the statement by Lemma G.3. $\square$

Theorem 3.9. *The norm functions ψ and χ (11) are $\tilde{\Gamma}_{p,q,r}^1$ -equivariant:*

$$\tilde{\text{ad}}_T(\psi(x)) = \psi(\tilde{\text{ad}}_T(x)), \quad \tilde{\text{ad}}_T(\chi(x)) = \chi(\tilde{\text{ad}}_T(x)) \quad (41)$$

for any $T \in \tilde{\Gamma}_{p,q,r}^1$ and $x \in \mathcal{C}\ell_{p,q,r}$.

Proof. Since $\tilde{\Gamma}_{p,q,r}^1 \subseteq (\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times})\Lambda_r^\times$ (Theorem 3.3), $\tilde{\text{ad}}_T$ is multiplicative for any $T \in \tilde{\Gamma}_{p,q,r}^1$ by Lemma 2.1. The mappings $\psi(x)$ and $\chi(x)$ are $\tilde{\Gamma}_{p,q,r}^1$ -equivariant by Lemma G.4, since they are products of the $\tilde{\Gamma}_{p,q,r}^1$ -equivariant mappings: $x \mapsto \tilde{x}$, $x \mapsto \widehat{\tilde{x}}$, and $x \mapsto x$ (Theorem 3.7). $\square$

4. Methodology

In this section, we describe the architecture of Generalized Lipschitz Groups Neural Networks layers. In this and the following sections, we consider the real geometric algebra $\mathcal{C}\ell(\mathbb{R}^{p,q,r})$. Code is available at <https://github.com/katyafilimoshina/glgenn>. We plan to continue developing this repository.

4.1. Geometric Algebra Data Embeddings

GLGENN, along with other GA-based equivariant neural networks such as CGENN (Ruhe et al., 2023), can be applied to any data representable as multivectors. For example, consider the following task in $\mathbb{R}^{0,3}$. We have data objects $x_1, \dots, x_l$, where each x_i contains a scalar (e.g. $a \in \mathbb{R}$, representing mass, temperature, or a one-hot-encoded feature,

etc.), a vector (e.g. $(b, c, d) \in \mathbb{R}^{0,3}$, where $b, c, d \in \mathbb{R}$, representing position, velocity, etc.), and a pseudoscalar (e.g., a signed volume value $f \in \mathbb{R}$). These components can be embedded in GA as $ae + be_1 + ce_2 + de_3 + fe_{123} \in \mathcal{Cl}_{0,3}$ and processed by our layers. Additional grades beyond 0, 1, and 3 can be utilized as well.² The output can be obtained by projecting the resulting multivectors onto the appropriate subspace based on the task. E.g., projecting onto $\mathcal{Cl}^0$ allows for scalar predictions (which will be both equivariant and invariant), or onto $\mathcal{Cl}_{p,q,r}^1$ if vector predictions are required.

4.2. Conjugation Operations Layers

These layers are based on the concept of conjugation operations in $\mathcal{Cl}_{p,q,r}$ (1). Suppose $x_1, \dots, x_l \in \mathcal{Cl}_{p,q,r}$ are input data multivectors, where l is a number of input channels. For an input data multivector $x_{c_{in}} \in \mathcal{Cl}_{p,q,r}$, $c_{in} = 1, \dots, l$, a conjugation operations layer is constructed using

$$x_{c_{in}} \mapsto \sum_{k=0}^n \phi_{c_{in}k} \langle x_{c_{in}} \rangle_k, \quad \phi_{c_{in}k} = \pm 1, \quad (42)$$

where $\phi_{c_{in}k} \in \{-1, 1\}$ are optimizable coefficients. The conjugation operation layer is $\tilde{\Gamma}_{p,q,r}^1$ -equivariant by Theorem 3.8 and contains $l(n+1)$ optimizable parameters. We suggest the following method to make the optimization of conjugation layers with discrete parameters possible. We first apply a linear transformation of the projections of the input multivector as in (42) but with any parameters. Then we round them to -1 or 1 either by directly applying $\phi_{c_{in}k} \mapsto \text{sgn}(\phi_{c_{in}k})$ or by firstly applying the sigmoid function, then scaling the value to $[-1, 1]$, and applying sign function to the result. Conjugation operations, including $\hat{\sim}$, $\sim$, and $\hat{\sim}$, play a significant role in GAs and their applications (see Appendix C for a detailed discussion), making these layers highly promising for experiments.

4.3. $\mathcal{Cl}_{p,q,r}^k$ -Linear Layers

A $\mathcal{Cl}_{p,q,r}^k$ -linear layer is constructed using

$$\langle y_{c_{out}} \rangle_{\bar{k}} := \sum_{c_{in}=0}^3 \phi_{c_{out}c_{in}\bar{k}} \langle x_{c_{in}} \rangle_{\bar{k}}, \quad (43)$$

where $y_{c_{out}} := \sum_{k=0}^3 \langle y_{c_{out}} \rangle_{\bar{k}}$ is an output channel, $\phi_{c_{in}c_{out}\bar{k}} \in \mathbb{R}$ are optimizable coefficients, and c_{in} and

²Note that in the case of $\mathcal{Cl}_{0,3}$, the equivariant model proposed in this paper has some connection with equivariant models based on irreducible representations (irreps) of $O(3)$ (Han et al., 2022). Generally speaking, these are two different approaches: irreps of $O(3)$ follow the tensor approach, whereas the GAs operate with multivectors. However, there are connections between the two approaches: scalars, vectors, bivectors, and trivectors in GA are equivalent to $l = 0$ irreps with even parity, $l = 1$ irreps with odd parity, $l = 1$ irreps with even parity, and $l = 0$ irreps with odd parity, respectively. The dimensions are the same (1, 3, 3, and 1, respectively), the behavior under rotations/reflections of the coordinate system matches, and the coupling operations (such as scalar multiplication, vector scalar and cross products in various cases) are also the same. Further study of this connection deserves a separate study and is not the topic of the current work, but seems important for building bridges between the two communities.

c_{out} are used to denote the number of the input and the output channels respectively. In other words, for any fixed $k = 0, 1, 2, 3$, the value $\langle y_{c_{out}} \rangle_{\bar{k}}$ is a linear combination of $\langle x_{c_{in}} \rangle_{\bar{k}}$, where $c_{in} = 1, \dots, l$. One $\mathcal{Cl}_{p,q,r}^k$ -linear layer is parametrized with $4lm$ optimizable coefficients, where l and m are the numbers of input and output channels respectively. Such layers are $\tilde{\Gamma}_{p,q,r}^1$ -equivariant by Theorems 3.5 and 3.6.

4.4. $\mathcal{Cl}_{p,q,r}^k$ -Geometric Product Layers

Now let us parameterize product interaction terms. We consider only the second-order interactions, because the higher-order interactions are indirectly modeled via multiple successive layers (Ruhe et al., 2023). A second-order interaction term for the pair of multivectors x_1 and x_2 has the form $\langle \langle x_1 \rangle_{\bar{i}} \langle x_2 \rangle_{\bar{j}} \rangle_{\bar{k}}$, where $i, j, k = 0, 1, 2, 3$. All the terms from $\mathcal{Cl}_{p,q,r}^k$ resulting from the interaction of x_1 and x_2 are parameterized as

$$P(x_1, x_2)_{\bar{k}} := \sum_{i=0}^3 \sum_{j=0}^3 \phi_{ijk} \langle \langle x_1 \rangle_{\bar{i}} \langle x_2 \rangle_{\bar{j}} \rangle_{\bar{k}}, \quad (44)$$

where $\phi_{ijk} \in \mathbb{R}$ are optimizable coefficients. $P(x_1, x_2)_{\bar{k}}$ is $\tilde{\Gamma}_{p,q,r}^1$ -equivariant by Theorem 3.5 and Theorem 3.6. We get 4^3 parameters ϕ_{ijk} for each geometric product of a pair of multivectors. Since we have l^2 possible pairs of multivectors for l input channels, parameterizing such number of coefficients can be computationally expensive. Therefore, if a number of channels l is large, let us compute geometric products for less than l^2 pairs of multivectors by an approach proposed in Ruhe et al., 2023. Firstly, we apply to $x_1, \dots, x_l \in \mathcal{Cl}_{p,q,r}$ a linear map to get $y_1, \dots, y_l \in \mathcal{Cl}_{p,q,r}$ and then compute the products of the pairs x_i, y_i for $i = 1, \dots, l$, which we denote by $z_1, \dots, z_l$ respectively. Using such a trick, the terms that will be multiplied get learned. Now, we need to calculate only l geometric products. We get $\langle z_{c_{out}} \rangle_{\bar{k}} := P(x_{c_{in}}, y_{c_{in}})_{\bar{k}}$, where $z_{c_{out}} := \sum_{k=0}^3 \langle z_{c_{out}} \rangle_{\bar{k}}$ and $c_{in} = c_{out} = 1, \dots, l$. In total, for a $\mathcal{Cl}_{p,q,r}^k$ -geometric product layer, we get $4l^2 + 4^3l$ parameters for l input channels.

4.5. $\mathcal{Cl}_{p,q,r}^k$ -Normalization Layers

For numerical stability of $\mathcal{Cl}_{p,q,r}^k$ -geometric product layers, we apply normalization to the four projections $\langle x_{c_{in}} \rangle_{\bar{0}}$, $\langle x_{c_{in}} \rangle_{\bar{1}}$, $\langle x_{c_{in}} \rangle_{\bar{2}}$, and $\langle x_{c_{in}} \rangle_{\bar{3}}$ for each multivector $x_{c_{in}}$, $c_{in} = 1, \dots, l$ before geometric product:

$$\langle x_{c_{in}} \rangle_{\bar{k}} \mapsto \frac{\langle x_{c_{in}} \rangle_{\bar{k}}}{\sigma(\phi_{c_{in}\bar{k}})(\langle \langle x_{c_{in}} \rangle_{\bar{k}} \langle x_{c_{in}} \rangle_{\bar{k}} \rangle_{\bar{0}} - 1) + 1}, \quad (45)$$

where $\sigma(x) := \frac{1}{1+e^{-x}} \in (0, 1)$ is the logistic sigmoid function and $\phi_{c_{in}\bar{m}} \in \mathbb{R}$ are optimizable parameters. The similar approach is applied in Ruhe et al., 2023. Note that $\langle \langle x_{c_{in}} \rangle_{\bar{k}} \langle x_{c_{in}} \rangle_{\bar{k}} \rangle_{\bar{0}}$ is $\tilde{\Gamma}_{p,q,r}^1$ -equivariant by Theorems 3.6 and 3.9. Each normalization layer contains $4l$ optimizable parameters, where l is the number of input channels.

5. Experiments

We demonstrate that GLGENN either outperforms or matches the performance of other equivariant models, while using significantly fewer optimizable parameters. For CGENN (Ruhe et al., 2023), the training setups are aligned with its public code release. For other models, we use the loss values from the corresponding code repository (Finzi et al., 2021). All experimental details and discussions can be found in Appendix I and our code repository. For simplicity, the current examples focus on the case of the non-degenerate GA $\mathcal{C}\ell_{p,q}$. Experiments involving $\mathcal{C}\ell_{p,q,r}$ will be explored in future research. We construct our models to closely resemble the CGENN architecture, replacing the $\mathcal{C}\ell_{p,q}^k$ -linear, $\mathcal{C}\ell_{p,q}^k$ -geometric product, and $\mathcal{C}\ell_{p,q}^k$ -normalization layers from Ruhe et al., 2023 with the same number of GLGENN’s $\mathcal{C}\ell_{p,q}^k$ counterparts (see Section 4).

O(5, 0)-Regression Task

We consider an O(5, 0)-invariant regression task (Finzi et al., 2021) to estimate the function $\sin(\|x_1\|) - \|x_2\|^3/2 + \frac{x_1^T x_2}{\|x_1\|\|x_2\|}$, where $x_1, x_2 \in \mathbb{R}^{5,0}$ are vectors sampled from a standard Gaussian distribution. The loss function used is Mean squared error (MSE). We evaluate the performance using four different training dataset sizes and compare against the ordinary MLP, MLP with augmentations, the O(5, 0)- and SO(5, 0)-equivariant MLP architectures proposed in (Finzi et al., 2021), and with CGENN (Ruhe et al., 2023). The results, presented in Tables 2 and 6 and Figure 2 (left), demonstrate that GLGENN achieves performance on a par with CGENN, while significantly outperforming the other models. Moreover, GLGENN attains these results with fewer parameters and reduced training time compared to CGENN. In this experiment, CGENN has approximately 1.8K parameters associated with GA layers, while GLGENN has around 0.6K parameters. Other models have approximately the same number of parameters as CGENN. In Tables 2 and 6, MSE for CGENN and GLGENN are averaged over 5 runs. MSE for other models are averaged over 3 runs (from Finzi et al., 2021). Number of iterations is the same for all algorithms. Note that when CGENN has the same size as GLGENN (with ≈ 0.6 K GA-associated parameters), we get the similar results (Table 10).

O(5, 0)- and O(7, 0)-Convex Hull Volume Estimation

In these experiments, the task is to estimate the volume of a convex hull generated by K points in $\mathbb{R}^{5,0}$ and $\mathbb{R}^{7,0}$ respectively. We first consider the setup with $K = 16$ points, as in

Table 2. MSE ($\downarrow$) on the O(5, 0)-Regression Experiment

MODEL	# OF TRAINING SAMPLES			
	$3 \cdot 10^1$	$3 \cdot 10^2$	$3 \cdot 10^3$	$3 \cdot 10^4$
GLGENN	0.1055	0.0020	0.0031	0.0011
MLP	28.1011	0.2482	0.0623	0.0622
MLP+Aug	0.4758	0.0936	0.0889	0.0672
EMLP-O(5)	0.152	0.0344	0.0310	0.0273
EMLP-SO(5)	0.1102	0.0384	0.032	0.0279
CGENN	0.0791	0.0089	0.0012	0.0003

Ruhe et al., 2023; Liu et al., 2024. The average convex hull volume in a training dataset is ≈ 11.4 . The results, which are presented in Figure 2 (middle) and Table 3, demonstrate that GLGENN either outperform or match CGENN, which itself outperforms ordinary MLP, Geometric Vector Perceptrons (Jing et al., 2021), and Vector Neurons (Deng et al., 2021). In Table 3, MSE for $K = 16$ are averaged over 5 runs, number of iterations is the same for both algorithms. Notably, as highlighted Ruhe et al., 2023, CGENN tends to overfit when trained on small datasets. In contrast, GLGENN demonstrate a reduced tendency to overfitting. The reason might be that GLGENN achieve these superior results with more than twice the parameter efficiency, using significantly fewer parameters than CGENN. Note that when CGENN is scaled down to approximately the same size as GLGENN (around 25K trainable parameters), its performance deteriorates compared to its original configuration (see Table 9). In the case $n = 7$, we get similar results (see Appendix I).

Following the suggestion of one of the anonymous reviewers, we further increase the difficulty and real-world relevance of the experiment by estimating the convex hull volumes of $K = 256$ and 512 points in $\mathbb{R}^{5,0}$. In this setup, the average convex hull volumes in the training datasets are ≈ 430.3 and 694 respectively. GLGENN consistently outperforms the best-performing CGENN architecture, with the results summarized in Table 3. Due to its significantly smaller number of trainable parameters, GLGENN also require less training time (see Table 7 in Appendix I).

Combining GLGENN with Typical Activation Functions

Standard activation functions (e.g. SiLU or ReLU) can be applied to elements of $\mathcal{C}\ell^0$ (scalars) without breaking equivariance. For simple tasks, the best performance is achieved by combining GLGENN (applied to all grade subspaces) with a standard neural network, such as MLP, acting on elements of $\mathcal{C}\ell^0$. E.g., in the O(5, 0)-Regression Task: (1) a standalone MLP with three ReLU-activated layers performs poorly; (2) GLGENN alone performs reasonably well but converges slower than (3) the combination of GLGENN (to all grades) + MLP (to scalars). With 300 training samples, this improvement is shown in Figure 3 and Table 11. The main limitation is that applying nonlinearities only to certain subspaces (e.g., scalars) prevents interaction between different grades (e.g., vectors and bivectors), effectively isolating them. In contrast, the nonlinearities introduced via geometric product layers inherently mix all grades, creating strong interactions.

O(5, 0)-N-Body Experiment

In this benchmarking equivariant experiment (Ruhe et al., 2023; Brehmer et al., 2023), we consider a system of $N = 5$ charged particles (bodies) with given masses, initial positions, and velocities. The objective is to predict the final positions of the bodies after the system evolves under

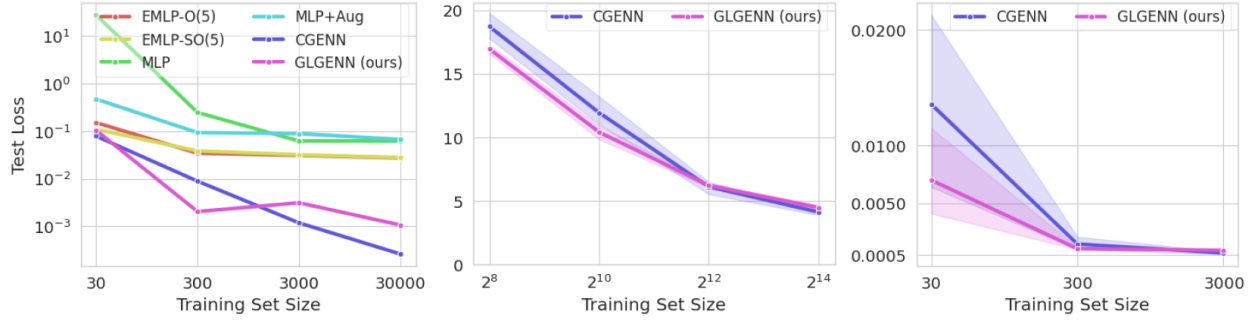


Figure 2. **Left:** $O(5, 0)$ -Regression. **Middle:** $O(5, 0)$ -Convex Hull, 16 points. The shaded regions depict 95% confidence intervals taken over 5 runs. **Right:** $O(5, 0)$ - N -Body. The shaded regions depict 95% confidence intervals taken over 3 runs.

Table 3. MSE ($\downarrow$) on the $O(5, 0)$ -Convex Hull Experiment (K points). The average convex hull volume in the training dataset for $K = 16, 256, 512$ points is $\approx 11.4, \approx 430.3$, and ≈ 694 , respectively. The number of trainable parameters is as follows: for $K = 16$, GLGENN 24.1K vs. CGENN 58.8K; for $K = 256$, GLGENN 791K vs. CGENN 1.72M; for $K = 512$, GLGENN 922K vs. CGENN 1.75M.

K	16							256		512	
	# TRAIN SAMPLES							# TRAIN SAMPLES		# TRAIN SAMPLES	
MODEL	2^8	2^{10}	2^{12}	2^{14}	2^{15}	2^{16}	2^{17}	2^{10}	2^{14}	2^{10}	2^{14}
GLGENN	16.94	10.40	6.2	4.46	3.62	3.04	2.61	2908.16	2918.09	8538.86	4872.39
CGENN	18.71	11.93	6.1	4.11	3.23	2.52	2.08	5176.7	3384.62	14727.6	7212.44
GAP	-1.77	-1.53	0.1	0.35	0.39	0.52	0.53	-2268.54	-466.53	-6188.74	-2340.05

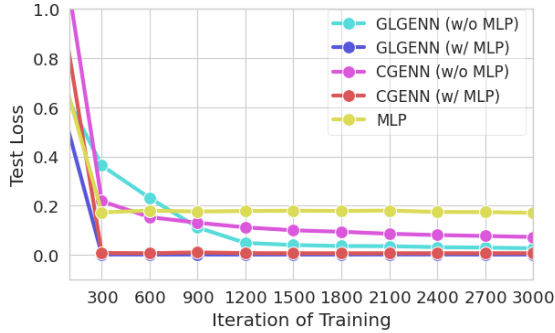


Figure 3. Combination of MLP with GLGENN and CGENN in $O(5, 0)$ -Regression.

Newtonian gravity for 1000 Euler integration steps. Unlike standard low-dimensional setups, we simulate the system in $\mathbb{R}^{5,0}$ and embed the data into $\mathcal{Cl}_{5,0}$. We construct a graph neural network (GNN) based on the message-passing paradigm (Gilmer et al., 2017), where bodies are treated as nodes in a graph, and their pairwise interactions are modeled as edges. For each edge, we compute a message and then aggregate at each node to update the particle states. The message and update networks are equivariant GLGENN. We compare against CGENN, which itself outperforms several state-of-the-art method (see Appendix I). To ensure a fair comparison, we use the best-performing CGENN architecture and then replace its layers with analogous GLGENN counterparts to obtain the GLGENN architecture, which automatically has two times fewer trainable parameters. The results are presented in Table 12 and Figure 2 (right).

6. Conclusions

We introduce a new equivariant neural networks architecture based on geometric algebras (GAs), called Gen-

eralized Lipschitz Groups Equivariant Neural Networks (GLGENN). These networks are equivariant with respect to pseudo-orthogonal transformations of a vector space with any non-degenerate or degenerate bilinear form. GLGENN are parameter-light, since they operate in a unified manner across four fundamental subspaces of GAs defined by the grade involution and reversion. GLGENN strike a balance between the expressiveness of Clifford Group Equivariant Neural Networks (CGENN) and parameter efficiency by respecting the fundamental algebraic structures of GAs³.

To develop GLGENN layers, we introduce and study the notion of generalized Lipschitz groups in GAs. We prove that the equivariance of a mapping with respect to these groups implies equivariance with respect to pseudo-orthogonal groups. As a result, GLGENN are equivariant with respect to a larger group than the pseudo-orthogonal group. However, empirically, we find that these networks are sufficiently expressive for tasks that require pseudo-orthogonal group equivariance. Experimental results demonstrate that GLGENN either outperform or match state-of-the-art models, while having significantly fewer trainable parameters and lower training time. Future research will focus on applying GLGENN to more complex real-world data experiments. Given their lower training time and promising results, we are optimistic about their potential.

³GLGENN goes beyond CGENN in the following: (1) Parameter-sharing approach for GA-based neural networks is presented for the first time. (2) New layers with parameter sharing are proposed, theoretical justification is provided (Section 4). (3) General idea is proposed that if we need orthogonal groups equivariance, then we may search for broader groups equivariance, such as new generalized Lipschitz groups, and get reasonable results (Theorem 3.4 and experiments).

Acknowledgements

The article was prepared within the framework of the project 'Mirror Laboratories' HSE University ('Quaternions, geometric algebras and applications').

The authors are grateful to the four anonymous reviewers for their careful reading of the paper and helpful comments on how to improve the presentation.

Impact Statement

The broader impact of this work is in improvement of the efficiency and effectiveness of equivariant neural networks. The significance of these networks is substantial, especially in the field of natural sciences, where the tasks inherently involve equivariance with respect to pseudo-orthogonal transformations. Parameter-light equivariant architectures reduce the risk of model's overfitting in case of small training datasets, which is a common scenario in the natural science field. Efficient equivariant neural networks have great potential to stimulate breakthroughs and innovations in modeling and simulations of physical systems, robotics, material science molecular biology, geoscience, etc.

References

- Ablamowicz, R. Structure of spin groups associated with degenerate Clifford algebras. *Journal of Mathematical Physics*, 27(1):1–6, 1986.
- Anderson, B., Hy, T. S., and Kondor, R. Cormorant: Covariant molecular neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 14510–14519, 2019.
- Atiyah, M., Bott, R., and Shapiro, A. Clifford modules. *Topology*, 3:3–38, 1964.
- Bayro-Corrochano, E. *Geometric Algebra Applications, vol. I, Computer Vision, Graphics and Neurocomputing*. Springer, 2019.
- Bayro-Corrochano, E. and Buchholz, S. Geometric neural networks. In Sommer, G. and Koenderink, J. J. (eds.), *Algebraic Frames for the Perception-Action Cycle*, pp. 379–394, Berlin, Heidelberg, 1997. Springer Berlin Heidelberg.
- Benn, I. and Tucker, R. *An Introduction to Spinors and Geometry with Applications in Physics*. Bristol, 1987.
- Brandstetter, J., Hesselink, R., van der Pol, E., Bekkers, E. J., and Welling, M. Geometric and physical quantities improve E(3) equivariant message passing. In *International Conference on Learning Representations*, 2022.
- Brandstetter, J., van den Berg, R., Welling, M., and Gupta, J. K. Clifford neural layers for PDE modeling. In *The Eleventh International Conference on Learning Representations*, 2023.
- Bredon, G. *Introduction to Compact Transformation Groups*. Academic Press, 1972.
- Brehmer, J., De Haan, P., Behrends, S., and Cohen, T. Geometric algebra transformer. In *Advances in Neural Information Processing Systems*, volume 37, 2023.
- Brooke, J. A Galileian formulation of spin: Clifford algebras and spin groups. *Journal of Mathematical Physics*, 19: 952–959, 1978.
- Brooke, J. *Clifford Algebras, Spin Groups and Galilei Invariance—New Perspectives*. PhD thesis, University of Alberta, 1980.
- Buchholz, S. and Sommer, G. On clifford neurons and Clifford multi-layer perceptrons. *Neural Networks*, 21(7): 925–935, 2008.
- Choi, H. I., Lee, D. S., and Moon, H. P. Clifford algebra, spin representation, and rational parameterization of curves and surfaces. *Advances in Computational Mathematics*, 17:5–48, 2002.
- Cohen, T. and Welling, M. Group equivariant convolutional networks. In *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of PMLR, pp. 2990–2999, 2016.
- Cohen, T. S. and Welling, M. Steerable CNNs. In *International Conference on Learning Representations*, 2017.
- Crumeysrolle, A. *Orthogonal and Symplectic Clifford Algebras*. Springer, Netherlands, 1st edition, 1990.
- Deng, C., Litany, O., Duan, Y., Poulenard, A., Tagliasacchi, A., and Guibas, L. J. Vector neurons: A general framework for SO(3)-equivariant networks. In *IEEE International Conference on Computer Vision (ICCV)*, pp. 12180–12189. IEEE, 2021.
- Dereli, T., Kocak, S., and Limoncu, M. Degenerate spin groups as semi-direct products. *Advances in Applied Clifford Algebras*, 20:565–573, 10 2010.
- Dieudonné, J. *La géométrie des groupes classiques*. Springer-Verlag, 1971.
- Doran, C. and Lasenby, A. *Geometric Algebra for Physicists*. Cambridge University Press, Cambridge, UK, 2003.
- Dorst, L., Doran, C., and Lasenby, J. *Applications of Geometric Algebra in Computer Science and Engineering*. Birkhäuser, 2002.

-
- Filimoshina, E. and Shirokov, D. On some Lie groups in degenerate Clifford geometric algebras. *Advances in Applied Clifford Algebras*, 33(44), 2023.
- Filimoshina, E. and Shirokov, D. On generalization of Lipschitz groups and spin groups. *Mathematical Methods in the Applied Sciences*, 47(3):1375–1400, 2024a.
- Filimoshina, E. and Shirokov, D. A note on centralizers and twisted centralizers in Clifford algebras. *Advances in Applied Clifford Algebras*, 34(50), 2024b.
- Filimoshina, E. and Shirokov, D. Generalized degenerate Clifford and Lipschitz groups in geometric algebras. *Advances in Applied Clifford Algebras*, 35(29), 2025.
- Finzi, M., Welling, M., and Wilson, A. G. Practical method for constructing equivariant multilayer perceptrons for arbitrary matrix groups. volume 139 of *PMLR*, pp. 3318–3328, 2021.
- Fuchs, F., Worrall, D., Fischer, V., and Welling, M. SE(3)-transformers: 3D roto-translation equivariant attention networks. In *Advances in Neural Information Processing Systems*, volume 33, pp. 1970–1981, 2020.
- Garling, D. *Clifford Algebras: An Introduction*. Cambridge University Press, Cambridge, UK, 2011.
- Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., and Dahl, G. E. Neural message passing for quantum chemistry. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1263–1272. PMLR, 2017.
- Han, J., Rong, Y., Xu, T., and Huang, W. Geometrically equivariant graph neural networks: A survey. *ArXiv*, 2202.07230, 2022.
- Harvey, F. *Spinors and Calibrations*. Academic Press, 1990.
- Helmstetter, J. and Micali, A. *Quadratic Mappings and Clifford Algebras*. Birkhäuser, 2008.
- Hestenes, D. and Sobczyk, G. *Clifford Algebra to Geometric Calculus – A Unified Language for Mathematical Physics*. Reidel Publishing Company, 1984.
- Isaacs, I. M. *Algebra: A Graduate Course*. American Mathematical Society, 2009.
- Jing, B., Eismann, S., Suriana, P., Townshend, R., and Dror, R. Learning from protein structure with geometric vector perceptrons. In *ICLR*, 2021.
- Knus, M.-A. *Quadratic and Hermitian forms over rings*. Springer, Berlin, 1991.
- Köhler, J., Klein, L., and Noé, F. Equivariant flows: exact likelihood generative learning for symmetric densities. In *Proceedings of the 37th International Conference on Machine Learning, ICML’20*, 2020.
- Kuroe, Y. Models of Clifford recurrent neural networks and their dynamics. In *The 2011 International Joint Conference on Neural Networks*, pp. 1035–1041, 2011.
- Langley, P. Crafting papers on machine learning. In Langley, P. (ed.), *Proceedings of the 17th International Conference on Machine Learning (ICML 2000)*, pp. 1207–1216, Stanford, CA, 2000. Morgan Kaufmann.
- Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- Liu, C., Ruhe, D., Eijkelboom, F., and Forré, P. Clifford group equivariant simplicial message passing networks. In *ICLR*, 2024.
- Lounesto, P. *Clifford Algebras and Spinors*. Cambridge University Press, 1997.
- Lundholm, D. and Svensson, L. Clifford algebra, geometric algebra, and applications, 2009. <https://arxiv.org/abs/0907.5356>.
- Pearson, J. and Bisset, D. Neural networks in the Clifford domain. In *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN’94)*, volume 3, pp. 1465–1469 vol.3, 1994.
- Pepe, A., Buchholz, S., and Lasenby, J. Clifford group equivariant neural network layers for protein structure prediction. In *Northern Lights Deep Learning Conference*, 2024.
- Pitts, A. M. *Nominal Sets: Names and Symmetry in Computer Science*, volume 57 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 2013.
- Porteous, I. *Clifford Algebras and the Classical Groups*. Cambridge University Press, 1995.
- Ruhe, D., Brandstetter, J., and Forré, P. Clifford group equivariant neural networks. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.
- Satorras, V. G., Hoogeboom, E., and Welling, M. E(n)-equivariant graph neural networks. In *ICML*, 2021.
- Shirokov, D. Quaternion typification of Clifford algebra elements. *Advances in Applied Clifford Algebras*, 22(1): 243–256, 2012a.

-
- Shirokov, D. Development of the method of quaternion typification of Clifford algebra elements. *Advances in Applied Clifford Algebras*, 22(2):483–497, 2012b.
- Shirokov, D. Clifford algebras and their applications to Lie groups and spinors. In *Proceedings of the 19th International Conference on Geometry, Integrability and Quantization*, pp. 11–53, 2018.
- Shirokov, D. On inner automorphisms preserving fixed subspaces of Clifford algebras. *Advances in Applied Clifford Algebras*, 31:30, 2021.
- Thomas, N., Smidt, T. E., Kearnes, S., Yang, L., Li, L., Kohlhoff, K., and Riley, P. Tensor field networks: Rotation- and translation-equivariant neural networks for 3D point clouds. *CoRR*, 2018.
- Townshend, R. J. L., Vögele, M., Suriana, P. A., Derry, A., Powers, A., Laloudakis, Y., Balachandar, S., Jing, B., Anderson, B. M., Eismann, S., Kondor, R., Altman, R., and Dror, R. O. ATOM3d: Tasks on molecules in three dimensions. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 1)*, 2021.
- Walpuski, T. Differential geometry IV, Lecture notes, 2022. <https://walpu.ski/Teaching/SS22/DifferentialGeometry4/>.
- Weiler, M. and Cesa, G. General E(2)-equivariant steerable CNNs. In *Neural Information Processing Systems*, 2019.
- Weiler, M., Geiger, M., Welling, M., Boomsma, W., and Cohen, T. 3d steerable cnns: Learning rotationally equivariant features in volumetric data. *ArXiv*, abs/1807.02547, 2018.
- Yarotsky, D. Universal approximations of invariant maps by neural networks. *Constructive Approximation*, 55(2): 407–474, 2022.
- Zerouali, A. Twisted conjugation on connected simple Lie groups and twining characters. 2020.
- Zhdanov, M., Ruhe, D., Weiler, M., Lucic, A., Brandstetter, J., and Forré, P. Clifford-steerable convolutional neural networks. volume 235 of *PMLR*, 2024.

Brief Summary of Appendix

In Appendix A, we provide an overview of notation that we use throughout the paper.

Appendix B provides an example of the main theoretical constructions considered in Section 2 using the geometric algebra $\mathcal{Cl}_{1,3}$ of the Minkowski space $\mathbb{R}^{1,3}$.

Appendix C discusses the significance of grade involution and reversion in geometric algebras (GAs) $\mathcal{Cl}_{p,q,r}$. We show that these operations are in some sense unique in geometric algebras. This paragraph serves as motivation for considering the four subspaces of geometric algebras determined by the grade involution and the reversion. Our proposed architecture GLGENN actively operates with these subspaces in its structure.

In Appendix D, we consider the notion of centralizers and twisted centralizers of the subspaces of fixed grades and subspaces determined by the grade involution and reversion. We write out explicit forms of these sets. The centralizers and twisted centralizers play a crucial role in our main theoretical results on generalized Lipschitz groups.

Appendix E considers the (ordinary) Lipschitz groups, spin groups, and adjoint and twisted adjoint representations in arbitrary degenerate and non-degenerate geometric algebras. We prove several properties of these representations and find their kernels.

Appendix F presents and proves the key theoretical results necessary to constructing GLGENN. Specifically, we introduce and study generalized Lipschitz groups in arbitrary degenerate or non-degenerate geometric algebra. These groups preserve the subspaces determined by the grade involution and reversion under the twisted adjoint representation. We prove that these groups can be defined equivalently, using only the norm functions, which are applied in the theory of spin groups. This results allows us to prove that the (ordinary) Lipschitz groups are subgroups of the generalized Lipschitz groups.

Appendix G proves several properties of equivariant mappings, which are applied in Section 4.

Appendix H considers the relation between the well-known degenerate and non-degenerate pseudo-orthogonal (or complex orthogonal) groups and generalized Lipschitz groups. The obtained results enable us to prove that equivariance of a mapping with respect to the generalized Lipschitz groups implies its equivariance with respect to the corresponding pseudo-orthogonal group. This result is crucial for the construction of GLGENN, which are equivariant with respect to pseudo-orthogonal transformations.

Appendix I provides experimental details and discussions.

A. Notation

In Table 4, we provide an overview of notation used throughout the paper. We write out the notation, its meaning, and the place where it is mentioned for the first time.

B. Example of Geometric Algebra $\mathcal{Cl}_{1,3}$ of Minkowski Space $\mathbb{R}^{1,3}$

In this section, we provide an example of the key theoretical concepts and constructions considered in Section 2. Let us consider the Minkowski space $V = \mathbb{R}^{1,3}$ equipped with the Minkowski metric $\eta = \text{diag}(1, -1, -1, -1)$. In this case, we have $p = 1, q = 3, r = 0$, and $n = p + q = 4$. The corresponding geometric (Clifford) algebra is $\mathcal{Cl}_{1,3} = \mathcal{Cl}(\mathbb{R}^{1,3})$, which is the real non-degenerate $2^4 = 16$ -dimensional algebra with the identity element e and generators e_1, e_2, e_3, e_4 that satisfy:

$$e_a e_b + e_b e_a = 2\eta_{ab} e, \quad a, b = 1, \dots, 4, \quad (46)$$

which implies that the distinct generators anticommute, the first generator squared gives $+e$, and the other three generators squared give $-e$:

$$e_a e_b = -e_b e_a, \quad a, b = 1, \dots, 4, \quad a \neq b; \quad (e_1)^2 = +e, \quad (e_2)^2 = (e_3)^2 = (e_4)^2 = -e. \quad (47)$$

The basis elements of the algebra $\mathcal{Cl}_{1,3}$ are constructed as ordered products of distinct generators, such as $e_{123} = e_1 e_2 e_3$ and, in general, $e_{a_1 \dots a_k} = e_{a_1} \dots e_{a_k}$ for $a_1, \dots, a_k \in \{1, 2, 3, 4\}$ and $a_1 < \dots < a_k$. In the case of $\mathcal{Cl}_{1,3}$, we have five non-trivial subspaces of fixed grades:

$$\mathcal{Cl}_{1,3}^0 = \text{span}\{e\} \equiv \mathbb{R}, \quad \mathcal{Cl}_{1,3}^1 = \text{span}\{e_1, e_2, e_3, e_4\} \equiv V, \quad \mathcal{Cl}_{1,3}^2 = \text{span}\{e_{12}, e_{13}, e_{14}, e_{23}, e_{24}, e_{34}\}, \quad (48)$$

Table 4. Summary of notation.

NOTATION	MEANING	FIRST MENTION
$\mathcal{C}\ell_{p,q,r}$	(CLIFFORD) GEOMETRIC ALGEBRA (GA) OVER THE REAL $\mathbb{R}^{p,q,r}$ OR COMPLEX $\mathbb{C}^{p+q,0,r}$ VECTOR SPACE	PAGE 3
$\mathcal{C}\ell_{p,q}$	NON-DEGENERATE GEOMETRIC ALGEBRA $\mathcal{C}\ell_{p,q,0}$	PAGE 3
Λ_r	GRASSMANN SUBALGEBRA $\mathcal{C}\ell_{0,0,r}$ OF $\mathcal{C}\ell_{p,q,r}$	PAGE 3
V	REAL $\mathbb{R}^{p,q,r}$ OR COMPLEX $\mathbb{C}^{p+q,0,r}$ VECTOR SPACE	PAGE 3
$\mathbb{F}$	FIELD OF REAL OR COMPLEX NUMBERS IN THE CASES $\mathbb{R}^{p,q,r}$ AND $\mathbb{C}^{p+q,0,r}$ RESPECTIVELY	PAGE 3
η	MATRIX OF A BILINEAR FORM $V \times V \rightarrow \mathbb{F}$	PAGE 3
$\mathbf{b}$	SYMMETRIC BILINEAR FORM, $\mathbf{b} : V \times V \rightarrow \mathbb{F}$	PAGE 3
$\mathbf{q}$	QUADRATIC FORM, $\mathbf{q} : V \rightarrow \mathbb{F}$	PAGE 3
e	IDENTITY ELEMENT	PAGE 3
$e_1, \dots, e_n$	GENERATORS OF $\mathcal{C}\ell_{p,q,r}$	PAGE 3
$\mathcal{C}\ell^0$	SUBSPACE OF GRADE 0	PAGE 3
$\mathcal{C}\ell_{p,q,r}^k$	SUBSPACE OF GRADE k , WHERE $k = 1, \dots, n$	PAGE 3
$\{\mathcal{C}\ell_{p,q}^k \Lambda_r^l\}$	SUBSPACE OF $\mathcal{C}\ell_{p,q,r}$ SPANNED BY THE ELEMENTS OF THE FORM ab , WHERE $a \in \mathcal{C}\ell_{p,q}^k$ AND $b \in \Lambda_r^l$	PAGE 17
$\langle \rangle_k$	PROJECTION ONTO THE SUBSPACE OF GRADE k	PAGE 3
$\hat{U}$	GRADE INVOLUTE OF $U \in \mathcal{C}\ell_{p,q,r}$	FORMULA (2)
$\tilde{U}$	REVERSION OF $U \in \mathcal{C}\ell_{p,q,r}$	FORMULA (2)
$\hat{\hat{U}}$	CLIFFORD CONJUGATE OF $U \in \mathcal{C}\ell_{p,q,r}$	PAGE 3
$\mathcal{C}\ell_{p,q,r}^{(0)}, \mathcal{C}\ell_{p,q,r}^{(1)}$	EVEN AND ODD SUBSPACES OF $\mathcal{C}\ell_{p,q,r}$	FORMULA (3)
$(\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times}) \Lambda_r^\times$	PRODUCT OF TWO GROUPS: $\{ab \mid a \in \mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times}, b \in \Lambda_r^\times\}$	LEMMA 2.1
$\langle \rangle_{(l)}$	PROJECTION ONTO THE SUBSPACE $\mathcal{C}\ell_{p,q,r}^{(l)}$, WHERE $l = 0, 1$	PAGE 3
$\mathcal{C}\ell_{p,q,r}^m$	SUBSPACES DETERMINED BY GRADE INVOLUTION AND REVERSION	FORMULA (4)
$\langle \rangle_{\overline{m}}$	PROJECTION ONTO THE SUBSPACE $\mathcal{C}\ell_{p,q,r}^{\overline{m}}$, WHERE $m = 0, 1, 2, 3$	PAGE 3
$H^\times$	SUBSET OF ALL INVERTIBLE ELEMENTS OF A SET H	PAGE 3
ad	ADJOINT REPRESENTATION $\text{ad}_T(U) = TUT^{-1}$	FORMULA (5)
$\check{\text{ad}}$	TWISTED ADJOINT REPRESENTATION $\check{\text{ad}}_T(U) = \hat{T}UT^{-1}$	FORMULA (6)
$\tilde{\text{ad}}$	TWISTED ADJOINT REPRESENTATION $\tilde{\text{ad}}_T(U) = T\langle U \rangle_{(0)}T^{-1} + \hat{T}\langle U \rangle_{(1)}T^{-1}$	FORMULA (7)
$\tilde{\text{ad}}^1$	$\tilde{\text{ad}}$ RESTRICTED TO $\mathcal{C}\ell_{p,q,r}^1$	PAGE 5
$\tilde{\Gamma}_{p,q,r}^1$	ORDINARY LIPSCHITZ GROUPS	FORMULA (10)
ψ, χ	NORM FUNCTIONS OF $\mathcal{C}\ell_{p,q,r}$ ELEMENTS	FORMULA (11)
$\tilde{\Gamma}_{p,q,r}^k$	GENERALIZED LIPSCHITZ GROUPS, $k = 0, 1, 2, 3$	FORMULA (13)
$\Gamma_{p,q,r}^k = Q_{p,q,r}^k$	GROUPS PRESERVING THE SUBSPACES $\mathcal{C}\ell_{p,q,r}^k$, $k = 0, 1, 2, 3$, UNDER ad	FORMULAS (16) AND (18)–(19)
$\check{\Gamma}_{p,q,r}^k = \check{Q}_{p,q,r}^k$	GROUPS PRESERVING THE SUBSPACES $\mathcal{C}\ell_{p,q,r}^{\check{k}}$, $k = 0, 1, 2, 3$, UNDER $\check{\text{ad}}$	FORMULAS (17) AND (20)–(21)
$Z_{p,q,r}$	CENTER OF $\mathcal{C}\ell_{p,q,r}$	FORMULA (70)
$Z_{p,q,r}^k, \check{Z}_{p,q,r}^k$	CENTRALIZERS AND TWISTED CENTRALIZERS RESPECTIVELY OF $\mathcal{C}\ell_{p,q,r}^k$, $k = 0, \dots, n$	PAGE 4
$Z_{p,q,r}^{\bar{k}}, \check{Z}_{p,q,r}^{\bar{k}}$	CENTRALIZERS AND TWISTED CENTRALIZERS RESPECTIVELY OF $\mathcal{C}\ell_{p,q,r}^{\bar{k}}$, $k = 0, 1, 2, 3$	FORMULAS (86)–(87)
$\text{GL}(n, \mathbb{F})$	GENERAL LINEAR GROUP ACTING ON AN n -DIMENSIONAL VECTOR SPACE OVER $\mathbb{F}$	FORMULA (30)
$\text{O}(V, \mathbf{q})$	ORTHOGONAL GROUP	FORMULA (30)
$\text{O}_{\Lambda_r^1}(V, \mathbf{q})$	SUBGROUP OF $\text{O}(V, \mathbf{q})$ THAT LEAVES INVARIANT Λ_r^1	FORMULA (31)

$$\mathcal{C}\ell_{1,3}^3 = \text{span}\{e_{123}, e_{124}, e_{134}, e_{234}\}, \quad \mathcal{C}\ell_{1,3}^4 = \text{span}\{e_{1234}\}, \quad (49)$$

which we call the subspaces of scalars, vectors, bivectors, trivectors, and 4-vectors respectively. An arbitrary element

$U \in \mathcal{C}\ell_{1,3}$ can be written as

$$U = ue + \sum_{a=1,\dots,4} u_a e_a + \sum_{\substack{a,b=1,\dots,4, \\ a < b}} u_{ab} e_{ab} + \sum_{\substack{a,b,c=1,\dots,4, \\ a < b < c}} u_{abc} e_{abc} + u_{1234} e_{1234}, \quad u, \dots, u_{1234} \in \mathbb{R}. \quad (50)$$

The even $\mathcal{C}\ell_{1,3}^{(0)}$ and odd $\mathcal{C}\ell_{1,3}^{(1)}$ subspaces are defined as

$$\mathcal{C}\ell_{1,3}^{(0)} = \mathcal{C}\ell_{1,3}^0 \oplus \mathcal{C}\ell_{1,3}^2 \oplus \mathcal{C}\ell_{1,3}^4, \quad \mathcal{C}\ell_{1,3}^{(1)} = \mathcal{C}\ell_{1,3}^1 \oplus \mathcal{C}\ell_{1,3}^3. \quad (51)$$

In $\mathcal{C}\ell_{1,3}$, the four subspaces determined by the grade involution $\hat{}$ and the reversion $\tilde{}$ (the subspaces of quaternion types) are defined as

$$\mathcal{C}\ell_{1,3}^{\bar{0}} = \mathcal{C}\ell_{1,3}^0 \oplus \mathcal{C}\ell_{1,3}^4, \quad \mathcal{C}\ell_{1,3}^{\bar{1}} = \mathcal{C}\ell_{1,3}^1, \quad \mathcal{C}\ell_{1,3}^{\bar{2}} = \mathcal{C}\ell_{1,3}^2, \quad \mathcal{C}\ell_{1,3}^{\bar{3}} = \mathcal{C}\ell_{1,3}^3 \quad (52)$$

and are very similar to the subspaces of fixed grades because of small $n = 4$.

Let us consider the following example on how the projections $\langle \rangle_k$ onto the subspaces of fixed grades $\mathcal{C}\ell_{1,3}^k$, $k = 0, \dots, 4$, the projections $\langle \rangle_{(l)}$ onto the even and odd subspaces $\mathcal{C}\ell_{1,3}^{(l)}$, $l = 0, 1$, and the projections $\langle \rangle_{\bar{m}}$ onto $\mathcal{C}\ell_{1,3}^{\bar{m}}$, $m = 0, 1, 2, 3$ act on the element $W = e + 2e_1 + 3e_2 + 4e_{234} + 5e_{1234}$:

$$\langle W \rangle_0 = e, \quad \langle W \rangle_1 = \langle W \rangle_{\bar{1}} = 2e_1 + 3e_2, \quad \langle W \rangle_2 = \langle W \rangle_{\bar{2}} = 0, \quad \langle W \rangle_3 = \langle W \rangle_{\bar{3}} = 4e_{234}, \quad \langle W \rangle_4 = 5e_{1234}, \quad (53)$$

$$\langle W \rangle_{\bar{0}} = \langle W \rangle_{(0)} = e + 5e_{1234}, \quad \langle W \rangle_{(1)} = 2e_1 + 3e_2 + 4e_{234}. \quad (54)$$

For an arbitrary $U \in \mathcal{C}\ell_{1,3}$ (50), we have

$$\hat{U} = \langle U \rangle_{(0)} - \langle U \rangle_{(1)}, \quad \tilde{U} = \langle U \rangle_{\bar{0}} - \langle U \rangle_{\bar{1}} - \langle U \rangle_{\bar{2}} - \langle U \rangle_{\bar{3}}, \quad \hat{\tilde{U}} = \langle U \rangle_{\bar{0}} - \langle U \rangle_{\bar{1}} - \langle U \rangle_{\bar{2}} + \langle U \rangle_{\bar{3}}. \quad (55)$$

Now let us consider the example on how the adjoint representation ad (5) and twisted adjoint representations $\check{\text{ad}}$ (6) and $\tilde{\text{ad}}$ (7) act. Let us consider the element $T = e_1 + e_{123} \in \mathcal{C}\ell_{1,3}^\times$, which is invertible because $\frac{1}{2}(e_1 + e_{123})(e_1 - e_{123}) = e$. Suppose $U = e + e_4 \in \mathcal{C}\ell_{1,3}$. We have

$$\text{ad}_T(U) = TUT^{-1} = T(e + e_4)T^{-1} = TT^{-1} + Te_4T^{-1} = e + \frac{1}{2}(e_1 + e_{123})e_4(e_1 - e_{123}) \quad (56)$$

$$= e + \frac{1}{2}(e_1 + e_{123})(-e_1 + e_{123})e_4 = e - \frac{1}{2}(e_1 + e_{123})(e_1 - e_{123})e_4 = e - e_4, \quad (57)$$

$$\check{\text{ad}}_T(U) = \hat{T}UT^{-1} = -TUT^{-1} = -e + e_4, \quad (58)$$

$$\tilde{\text{ad}}_T(U) = T\langle U \rangle_{(0)}T^{-1} + \hat{T}\langle U \rangle_{(1)}T^{-1} = TeT^{-1} - Te_4T^{-1} = e + e_4. \quad (59)$$

Note that the results are different for different operations ad , $\check{\text{ad}}$, $\tilde{\text{ad}}$.

The Lipschitz group in the case $\mathcal{C}\ell_{1,3}$ has the form

$$\tilde{\Gamma}_{1,3}^1 = \{T \in \mathcal{C}\ell_{1,3}^\times : \hat{T}\mathcal{C}\ell_{1,3}^1T^{-1} \subseteq \mathcal{C}\ell_{1,3}^1\}. \quad (60)$$

A simple example of an element of $\tilde{\Gamma}_{1,3}^1$ is any basis element $e_{a_1 \dots a_k}$, since $e_{a_1 \dots a_k}^{-1} = \pm e_{a_1 \dots a_k}$, where the sign depends on $a_1, \dots, a_k$ and k , and $\widehat{e_{a_1 \dots a_k} e_a e_{a_1 \dots a_k}^{-1}} = \pm e_a \in \mathcal{C}\ell_{1,3}^1$ for any $a = 1, \dots, 4$. A slightly more complicated example of the Lipschitz group $\tilde{\Gamma}_{1,3}^1$ element is $T = e_1 + e_{123} \in \mathcal{C}\ell_{1,3}^\times$ considered in the paragraph above. We have:

$$\hat{T}e_1T^{-1} = -\frac{1}{2}(e_1 + e_{123})e_1(e_1 - e_{123}) = -e_1, \quad \hat{T}e_2T^{-1} = e_3, \quad \hat{T}e_3T^{-1} = e_2, \quad \hat{T}e_4T^{-1} = e_4. \quad (61)$$

Therefore, $\hat{T}e_aT^{-1} \in \mathcal{C}\ell_{1,3}^1$ for $a = 1, \dots, 4$; thus, $T \in \tilde{\Gamma}_{1,3}^1$.

C. Grade Involution and Reversion as Fundamental Operations in $\mathcal{Cl}_{p,q,r}$

This work introduces the neural networks that are equivariant with respect to the groups $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ (27) preserving the subspace $\mathcal{Cl}_{p,q,r}^{\bar{1}}$ (4) determined by the grade involution and reversion under the twisted adjoint representation $\tilde{\text{ad}}$. These operations can be considered as fundamental operations in geometric algebras $\mathcal{Cl}_{p,q,r}$. In this section, we discuss their significance.

In Subsection 2.1, the grade involute and reversion of an arbitrary $U \in \mathcal{Cl}_{p,q,r}$ are defined in the following way respectively:

$$\hat{U} := \sum_{k=0}^n (-1)^k \langle U \rangle_k, \quad \tilde{U} := \sum_{k=0}^n (-1)^{\frac{k(k-1)}{2}} \langle U \rangle_k. \quad (62)$$

These mappings have the following well-known properties, which we apply for many times in our work:

$$\widehat{(\widehat{UV})} = \widehat{U}\widehat{V}, \quad \widetilde{(\widetilde{UV})} = \widetilde{U}\widetilde{V}, \quad \forall U, V \in \mathcal{Cl}_{p,q,r}. \quad (63)$$

Let us show that these operations are in some sense unique in $\mathcal{Cl}_{p,q,r}$. For this purpose, we need to consider another equivalent definition of geometric algebras $\mathcal{Cl}_{p,q,r}$. More detailed discussion of definitions and statements of this section can be found, for example, in Lounesto, 1997; Helmstetter & Micali, 2008.

The geometric (Clifford) algebra $\mathcal{Cl}(V)$ is an associative algebra with the unity e over a vector space V with a symmetric bilinear form $\mathfrak{b} : V \times V \rightarrow \mathbb{F}$ and the corresponding quadratic form $\mathfrak{q} : V \rightarrow \mathbb{F}$, with a linear map $j : V \rightarrow \mathcal{Cl}(V)$ such that

$$j^2(x) = \mathfrak{q}(x)e, \quad \forall x \in V, \quad (64)$$

and for any other associative algebra A over V with the unity e_A and any linear map $k : V \rightarrow A$, such that $k^2(x) = \mathfrak{q}(x)e_A$ for any $x \in V$, there exists a unique algebra homomorphism $f : \mathcal{Cl}(V) \rightarrow A$, such that $f \circ j = k$, i.e. the corresponding diagram is commutative:

$$\begin{array}{ccc} V & \xrightarrow{j} & \mathcal{Cl}(V) \\ & \searrow k & \downarrow f \\ & & A \end{array}$$

Theorem C.1. *In any geometric algebra $\mathcal{Cl}(V)$, there is a unique antiautomorphism $\tau : \mathcal{Cl}(V) \rightarrow \mathcal{Cl}(V)$, such that*

$$\tau(xy) = \tau(y)\tau(x), \quad \tau \circ \tau = \text{id}, \quad \tau(j(v)) = j(v), \quad \forall x, y \in \mathcal{Cl}(V), \quad \forall v \in V. \quad (65)$$

Proof. Let us consider the opposite algebra $\mathcal{Cl}(V)^o$, in which the product of x and y is given by yx . Then there exists a unique algebra homomorphism $\tau : \mathcal{Cl}(V) \rightarrow \mathcal{Cl}(V)^o$ as in the diagram below

$$\begin{array}{ccc} V & \xrightarrow{j} & \mathcal{Cl}(V) \\ & \searrow j & \downarrow \tau \\ & & \mathcal{Cl}(V)^o \end{array}$$

This completes the proof. □

Since in our case, V has finite dimension n with a basis $e_1, \dots, e_n$, the mapping τ is completely determined by its action on the basis elements. Namely, τ is defined by

$$\tau(e_i) = e_i, \quad \tau(e_{i_1}e_{i_2} \cdots e_{i_k}) = e_{i_k}e_{i_{k-1}} \cdots e_{i_1}, \quad \tau(e) = e, \quad (66)$$

where $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. The mapping τ coincides with the reversion $\sim$ defined as in (62).

Theorem C.2. *In any geometric algebra $\mathcal{Cl}(V)$, there is a unique automorphism $\alpha : \mathcal{Cl}(V) \rightarrow \mathcal{Cl}(V)$, such that*

$$\alpha(xy) = \alpha(x)\alpha(y), \quad \alpha \circ \alpha = \text{id}, \quad \alpha(j(v)) = -j(v), \quad \forall x, y \in \mathcal{Cl}(V), \quad \forall v \in V. \quad (67)$$

Proof. Let us consider the linear mapping $\alpha_0 : V \rightarrow \mathcal{Cl}(V)$ defined as $\alpha_0(v) = -j(v)$ for any $v \in V$. We get a unique homomorphism α as in the diagram below

$$\begin{array}{ccc} V & \xrightarrow{j} & \mathcal{Cl}(V) \\ & \searrow \alpha_0 & \downarrow \alpha \\ & & \mathcal{Cl}(V) \end{array}$$

Furthermore, any $x \in \mathcal{Cl}(V)$ can be represented as

$$x = \sum x_1 \cdots x_m, \quad x_1, \dots, x_m \in j(V), \quad (68)$$

and since $\alpha(x_j) = -x_j$ for any $j = 1, \dots, m$, we get $\alpha \circ \alpha = \text{id}$. This completes the proof. $\square$

Again, since V has finite dimension, the mapping α is completely determined by its action on the basis elements. Namely, α is defined by

$$\alpha(e_i) = -e_i, \quad \alpha(e_{i_1} e_{i_2} \cdots e_{i_k}) = (-1)^k e_{i_1} e_{i_2} \cdots e_{i_k}, \quad \alpha(e) = e, \quad (69)$$

where $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. The mapping α coincides with the grade involution $\hat{}$ defined as in (62).

D. Centralizers and Twisted Centralizers

We denote the *center* of the algebra $\mathcal{Cl}_{p,q,r}$ by $Z_{p,q,r}$. It is a well-known fact (Crumeyrolle, 1990) that the center has the following form

$$Z_{p,q,r} = \begin{cases} (\Lambda_r^{(0)} \oplus \mathcal{Cl}_{p,q,r}^n)^\times, & n \text{ is odd;} \\ \Lambda_r^{(0)\times}, & n \text{ is even.} \end{cases} \quad (70)$$

Consider the *centralizers* (see, for example, Garling, 2011; Isaacs, 2009) of the subspaces of fixed grades $\mathcal{Cl}_{p,q,r}^m$, $m = 0, \dots, n$, in $\mathcal{Cl}_{p,q,r}$:

$$Z_{p,q,r}^m := \{X \in \mathcal{Cl}_{p,q,r} : XV = VX, \quad \forall V \in \mathcal{Cl}_{p,q,r}^m\}. \quad (71)$$

The centralizer $Z_{p,q,r}^m$ contains all the elements of $\mathcal{Cl}_{p,q,r}$ that commute with any grade- m element. The center of the Clifford algebra $\mathcal{Cl}_{p,q,r}$ is the centralizer of the entire Clifford algebra $\mathcal{Cl}_{p,q,r}$.

Similarly, we consider the *twisted centralizers* of $\mathcal{Cl}_{p,q,r}^m$, $m = 0, \dots, n$, in $\mathcal{Cl}_{p,q,r}$:

$$\tilde{Z}_{p,q,r}^m := \{X \in \mathcal{Cl}_{p,q,r} : \hat{X}V = VX, \quad \forall V \in \mathcal{Cl}_{p,q,r}^m\}. \quad (72)$$

Note that $Z_{p,q,r}^m = \tilde{Z}_{p,q,r}^m = \mathcal{Cl}_{p,q,r}$ for $m < 0$ and $m > n$. The particular case $\tilde{Z}_{p,q,r}^1$ is considered in the papers Ruhe et al., 2023; Brooke, 1980; Filimoshina & Shirokov, 2024b; 2023.

Note that the projections $\langle Z_{p,q,r}^m \rangle_{(0)}$ and $\langle \tilde{Z}_{p,q,r}^m \rangle_{(0)}$ of $Z_{p,q,r}^m$ and $\tilde{Z}_{p,q,r}^m$ respectively onto the even subspace $\mathcal{Cl}_{p,q,r}^{(0)}$ coincide by definition:

$$\langle Z_{p,q,r}^m \rangle_{(0)} = \langle \tilde{Z}_{p,q,r}^m \rangle_{(0)}, \quad \forall m = 0, 1, \dots, n. \quad (73)$$

The paper Filimoshina & Shirokov, 2024b studies explicit forms of the sets $Z_{p,q,r}^m$ and $\tilde{Z}_{p,q,r}^m$ in the case of arbitrary $m = 0, \dots, n$. In Remark D.1, for the readers' convenience, we write out several particular cases of these sets, which we further use. A few words about the notation in this remark and below. The spaces $\mathcal{Cl}_{p,q}^k$ and Λ_r^k , $k = 0, \dots, n$, are regarded as subspaces of $\mathcal{Cl}_{p,q,r}$. By $\{\mathcal{Cl}_{p,q}^k \Lambda_r^l\}$, we denote the subspace of $\mathcal{Cl}_{p,q,r}$ spanned by the elements of the form ab , where $a \in \mathcal{Cl}_{p,q}^k$ and $b \in \Lambda_r^l$.

Remark D.1 (Explicit forms of centralizers and twisted centralizers). We have:

$$Z_{p,q,r}^1 = Z_{p,q,r} = \begin{cases} \Lambda_r^{(0)} \oplus C\ell_{p,q,r}^n, & n \text{ is odd,} \\ \Lambda_r^{(0)}, & n \text{ is even;} \end{cases} \quad (74)$$

$$Z_{p,q,r}^2 = \begin{cases} \Lambda_r \oplus C\ell_{p,q,r}^n, & r \neq n, \\ \Lambda_r, & r = n; \end{cases} \quad (75)$$

$$Z_{p,q,r}^3 = \begin{cases} \Lambda_r^{(0)} \oplus \Lambda_r^{n-2} \oplus \{C\ell_{p,q}^1(\Lambda_r^{n-3} \oplus \Lambda_r^{n-2})\} \oplus \{C\ell_{p,q}^2 \Lambda_r^{n-3}\} \oplus C\ell_{p,q,r}^n, & n \text{ is odd,} \\ \Lambda_r^{(0)} \oplus \Lambda_r^{n-1} \oplus \{C\ell_{p,q}^1 \Lambda_r^{\geq n-2}\} \oplus \{C\ell_{p,q}^2 \Lambda_r^{n-2}\}, & n \text{ is even;} \end{cases} \quad (76)$$

$$Z_{p,q,r}^4 = \begin{cases} \Lambda_r \oplus \{C\ell_{p,q}^1(\Lambda_r^{n-3} \oplus \Lambda_r^{n-2})\} \oplus \{C\ell_{p,q}^2(\Lambda_r^{n-4} \oplus \Lambda_r^{n-3})\} \oplus C\ell_{p,q,r}^n, & r \neq n, \\ \Lambda_r, & r = n; \end{cases} \quad (77)$$

and:

$$\check{Z}_{p,q,r}^1 = \Lambda_r; \quad (78)$$

$$\check{Z}_{p,q,r}^2 = \begin{cases} \Lambda_r^{(0)} \oplus \Lambda_r^n \oplus \{C\ell_{p,q}^1 \Lambda_r^{n-1}\}, & n \text{ is odd,} \\ \Lambda_r^{(0)} \oplus \Lambda_r^{n-1} \oplus \{C\ell_{p,q}^1 \Lambda_r^{n-2}\} \oplus C\ell_{p,q,r}^n, & n \text{ is even, } r \neq n; \\ \Lambda_r^{(0)} \oplus \Lambda_r^{n-1}, & n \text{ is even, } r = n; \end{cases} \quad (79)$$

$$\check{Z}_{p,q,r}^3 = \Lambda_r \oplus \{C\ell_{p,q}^1 \Lambda_r^{\geq n-2}\} \oplus \{C\ell_{p,q}^2 \Lambda_r^{\geq n-3}\}. \quad (80)$$

Remark D.2. In the particular case of the non-degenerate geometric algebras $C\ell_{p,q}$, we have

$$Z_{p,q}^1 = Z_{p,q} = \begin{cases} C\ell^0 \oplus C\ell_{p,q}^n, & n \text{ is odd,} \\ C\ell^0, & n \text{ is even;} \end{cases} \quad Z_{p,q}^2 = C\ell^0 \oplus C\ell_{p,q}^n; \quad (81)$$

$$Z_{p,q}^3 = \begin{cases} C\ell_{p,q}, & n = 2, 3, \\ C\ell^0, & n \neq 2, 3; \end{cases} \quad Z_{p,q}^4 = \begin{cases} C\ell_{p,q}, & n = 2, 3, \\ C\ell_{p,q}^{(0)}, & n = 4, \\ C\ell^0 \oplus C\ell_{p,q}^n, & n \neq 2, 3, 4; \end{cases} \quad (82)$$

$$\check{Z}_{p,q}^1 = C\ell^0, \quad \check{Z}_{p,q}^2 = \begin{cases} C\ell_{p,q}, & n = 1, 2, \\ C\ell^0, & n \neq 1 \text{ is odd,} \\ C\ell^0 \oplus C\ell_{p,q}^n, & n \neq 2 \text{ is even;} \end{cases} \quad \check{Z}_{p,q}^3 = \begin{cases} C\ell_{p,q}, & n = 1, 2, \\ C\ell_{p,q}^{(0)}, & n = 3, \\ C\ell^0, & n \geq 4. \end{cases} \quad (83)$$

Using the definition of the even subspace $C\ell_{p,q,r}^{(0)}$ and explicit forms of centralizers of fixed grades, presented in [Filimoshina & Shirokov, 2024b](#), the following lemma can be obtained.

Lemma D.3. *We have*

$$\{X \in C\ell_{p,q,r} : XV = VX, \quad \forall V \in C\ell_{p,q,r}^{(0)}\} = \begin{cases} \Lambda_r \oplus C\ell_{p,q,r}^n, & r \neq n, \\ \Lambda_n, & r = n, \end{cases} \quad (84)$$

$$\{X \in C\ell_{p,q,r} : \hat{X}V = VX, \quad \forall V \in C\ell_{p,q,r}^{(0)}\} = \begin{cases} \Lambda_r^{(0)}, & n \text{ is odd; } r = n \text{ is even;} \\ \Lambda_r^{(0)} \oplus C\ell_{p,q,r}^n, & n \text{ is even, } r \neq n. \end{cases} \quad (85)$$

Now let us consider the centralizers and twisted centralizers respectively of the subspaces $C\ell_{p,q,r}^{\bar{k}}$ (4), $k = 0, 1, 2, 3$, defined by the grade involution and reversion, in $C\ell_{p,q,r}$:

$$Z_{p,q,r}^{\bar{k}} := \{X \in C\ell_{p,q,r} : XV = VX, \quad \forall V \in C\ell_{p,q,r}^{\bar{k}}\}, \quad (86)$$

$$\check{Z}_{p,q,r}^{\bar{k}} := \{X \in C\ell_{p,q,r} : \hat{X}V = VX, \quad \forall V \in C\ell_{p,q,r}^{\bar{k}}\}. \quad (87)$$

Theorem D.4. ([Filimoshina & Shirokov, 2024b](#)) *We have*

$$Z_{p,q,r}^{\bar{m}} = Z_{p,q,r}^m, \quad \check{Z}_{p,q,r}^{\bar{m}} = \check{Z}_{p,q,r}^m, \quad m = 1, 2, 3; \quad Z_{p,q,r}^{\bar{0}} = Z_{p,q,r}^4, \quad \check{Z}_{p,q,r}^{\bar{0}} = \langle Z_{p,q,r}^4 \rangle_{(0)}. \quad (88)$$

The centralizers $Z_{p,q,r}^m$ and twisted centralizers $\tilde{Z}_{p,q,r}^m$, $m = 1, 2, 3, 4$, are written out explicitly in Remark D.1. In the formula (88), we have

$$\langle Z_{p,q,r}^4 \rangle_{(0)} = \begin{cases} \Lambda_r^{(0)} \oplus \{C\ell_{p,q}^1 \Lambda_r^{n-2}\} \oplus \{C\ell_{p,q}^2 \Lambda_r^{n-3}\}, & n \text{ is odd or } r = n; \\ \Lambda_r^{(0)} \oplus \{C\ell_{p,q}^1 \Lambda_r^{n-3}\} \oplus \{C\ell_{p,q}^2 \Lambda_r^{n-4}\} \oplus C\ell_{p,q,r}^n, & n \text{ is even, } r \neq n. \end{cases} \quad (89)$$

Let us consider the following Lemma D.5 about $T \in C\ell_{p,q,r}$ such that $\psi(T)$ and $\chi(T)$ (11) are in the centralizers $Z_{p,q,r}^{\bar{k}}$ (86) and twisted centralizers $\tilde{Z}_{p,q,r}^{\bar{k}}$ (87). The proofs of theorems in Appendix F are based on this lemma.

Lemma D.5. (*Filimoshina & Shirokov, 2025*) For any $T \in C\ell_{p,q,r}^\times$, in the cases $(k, l) = (0, 1), (1, 0), (2, 3), (3, 2)$:

$$TC\ell_{p,q,r}^{\bar{k}} T^{-1} \subseteq C\ell_{p,q,r}^{\bar{k}l} \Leftrightarrow \tilde{T}T \in Z_{p,q,r}^{\bar{k} \times}, \quad (90)$$

$$\hat{T}C\ell_{p,q,r}^{\bar{k}} T^{-1} \subseteq C\ell_{p,q,r}^{\bar{k}l} \Leftrightarrow \hat{\tilde{T}}T \in \tilde{Z}_{p,q,r}^{\bar{k} \times}, \quad (91)$$

and in the cases $(k, l) = (0, 3), (3, 0), (1, 2), (2, 1)$, we have:

$$TC\ell_{p,q,r}^{\bar{k}} T^{-1} \subseteq C\ell_{p,q,r}^{\bar{k}l} \Leftrightarrow \tilde{T}T \in Z_{p,q,r}^{\bar{k} \times}, \quad (92)$$

$$\hat{T}C\ell_{p,q,r}^{\bar{k}} T^{-1} \subseteq C\ell_{p,q,r}^{\bar{k}l} \Leftrightarrow \tilde{T}T \in \tilde{Z}_{p,q,r}^{\bar{k} \times}. \quad (93)$$

E. Ordinary Lipschitz Groups, Spin Groups, Adjoint and Twisted Adjoint Representations

In the non-degenerate geometric algebras $C\ell_{p,q}$, the Lipschitz group is defined as the group of all invertible elements preserving the grade-1 subspace under the twisted adjoint representation ad . Generalizing this definition to the case of the degenerate $C\ell_{p,q,r}$, we similarly define the *Lipschitz group* $\tilde{\Gamma}_{p,q,r}^1$ as

$$\tilde{\Gamma}_{p,q,r}^1 := \{T \in C\ell_{p,q,r}^\times : \hat{T}C\ell_{p,q,r}^1 T^{-1} \subseteq C\ell_{p,q,r}^1\}. \quad (94)$$

The definition (94) of the degenerate Lipschitz group is used, for example, in the works Brooke, 1980; 1978; Crumeyrolle, 1990.

The well-known spin groups (see, for example, Lounesto, 1997; Crumeyrolle, 1990) are defined as normalized subgroups of the Lipschitz groups $\tilde{\Gamma}_{p,q,r}^1$, using the norm functions ψ and χ (11). For example, in the non-degenerate geometric algebras $C\ell_{p,q}$, they are defined as

$$\text{Pin}_{p,q} := \{T \in \tilde{\Gamma}_{p,q}^1 : \tilde{T}T = \pm e\} = \{T \in \tilde{\Gamma}_{p,q}^1 : \hat{\tilde{T}}T = \pm e\}, \quad (95)$$

$$\text{Pin}_{p,q}^\psi := \{T \in \tilde{\Gamma}_{p,q}^1 : \tilde{T}T = +e\}, \quad (96)$$

$$\text{Pin}_{p,q}^\chi := \{T \in \tilde{\Gamma}_{p,q}^1 : \hat{\tilde{T}}T = +e\}, \quad (97)$$

$$\text{Spin}_{p,q} := \{T \in \tilde{\Gamma}_{p,q}^1 \cap C\ell_{p,q}^{(0)\times} : \tilde{T}T = \pm e\} = \{T \in \tilde{\Gamma}_{p,q}^1 \cap C\ell_{p,q}^{(0)\times} : \hat{\tilde{T}}T = \pm e\}, \quad (98)$$

$$\text{Spin}_{p,q}^+ := \{T \in \tilde{\Gamma}_{p,q}^1 \cap C\ell_{p,q}^{(0)\times} : \tilde{T}T = +e\} = \{T \in \tilde{\Gamma}_{p,q}^1 \cap C\ell_{p,q}^{(0)\times} : \hat{\tilde{T}}T = +e\}. \quad (99)$$

Let us consider the *adjoint representation* ad acting on the group of all invertible elements $\text{ad} : C\ell_{p,q,r}^\times \rightarrow \text{Aut}(C\ell_{p,q,r})$ as $T \mapsto \text{ad}_T$, where $\text{ad}_T : C\ell_{p,q,r} \rightarrow C\ell_{p,q,r}$:

$$\text{ad}_T(U) = TUT^{-1}, \quad U \in C\ell_{p,q,r}, \quad T \in C\ell_{p,q,r}^\times. \quad (100)$$

Since $\text{ad}_T(UV) = \text{ad}_T(U)\text{ad}_T(V)$ for all $U, V \in C\ell_{p,q,r}$, the mapping ad_T is an algebra homomorphism, which is moreover an algebra automorphism for all $T \in C\ell_{p,q,r}^\times$. It is called *inner automorphism* (or conjugation).

Also let us consider the *twisted adjoint representation*, which is introduced by Atiyah, Bott, and Shapiro in Atiyah et al., 1964. In this paper, the twisted adjoint representation is defined in the case of the non-degenerate algebra $C\ell_{0,q,0}$. It acts

on the Lipschitz group $\tilde{\Gamma}_{0,q,0}^1$ (94) in the way $\check{\text{ad}} : \Gamma_{0,q,0}^\pm \rightarrow \text{Aut}(C\ell_{0,q,0}^1)$ as $T \rightarrow \check{\text{ad}}_T$, where $\check{\text{ad}}_T : C\ell_{0,q,0}^1 \rightarrow C\ell_{0,q,0}^1$ is defined for elements of the grade-1 subspace (vectors) as

$$\check{\text{ad}}_T(U) = \hat{T}UT^{-1}, \quad U \in C\ell_{0,q,0}^1, \quad T \in \tilde{\Gamma}_{0,q,0}^1. \quad (101)$$

There are two ways how to generalize the definition (101) of the twisted adjoint representation to the case of any degenerate and non-degenerate $C\ell_{p,q,r}$ and arbitrary $T \in C\ell_{p,q,r}^\times$, $U \in C\ell_{p,q,r}$. The first approach (Choi et al., 2002; Harvey, 1990; Lundholm & Svensson, 2009) is to define it as the operation $\check{\text{ad}}$ acting on the group of all invertible elements $\check{\text{ad}} : C\ell_{p,q,r}^\times \rightarrow \text{Aut}(C\ell_{p,q,r})$ as $T \mapsto \check{\text{ad}}_T$ with $\check{\text{ad}}_T : C\ell_{p,q,r} \rightarrow C\ell_{p,q,r}$:

$$\check{\text{ad}}_T(U) = \hat{T}UT^{-1}, \quad U \in C\ell_{p,q,r}, \quad T \in C\ell_{p,q,r}^\times. \quad (102)$$

The operation $\check{\text{ad}}_T$ (102) is similar to some operations considered in representation theory of Lie groups (Zerouali, 2020). Note that unlike ad_T (100) the mapping $\check{\text{ad}}_T$ (102) is not multiplicative in U and, therefore, is not an algebra homomorphism for all $T \in C\ell_{p,q,r}^\times$.

The second approach (Helmstetter & Micali, 2008; Knus, 1991; Walpuski, 2022) is to define the twisted adjoint representation as the operation acting on the group of all invertible grade- m elements $\tilde{\text{ad}} : C\ell_{p,q,r}^{m \times} \rightarrow \text{Aut}(C\ell_{p,q,r})$ as $T \mapsto \tilde{\text{ad}}_T$ with $\tilde{\text{ad}}_T : C\ell_{p,q,r} \rightarrow C\ell_{p,q,r}$ defined as:

$$\tilde{\text{ad}}_T(U) = (-1)^{km}TUT^{-1}, \quad U \in C\ell_{p,q,r}^k, \quad T \in C\ell_{p,q,r}^{m \times}, \quad (103)$$

where we use another notation $\tilde{\text{ad}}_T$ not to confuse it with $\check{\text{ad}}_T$ (102). The mapping $\tilde{\text{ad}}_T$ is also called *twisted inner automorphism*, for example, in Helmstetter & Micali, 2008. The operation $\tilde{\text{ad}}_T$ can be extended by linearity $\tilde{\text{ad}}_T(U + V) = \tilde{\text{ad}}_T(U) + \tilde{\text{ad}}_T(V)$, and we finally obtain

$$\tilde{\text{ad}}_T(U) = T\langle U \rangle_{(0)}T^{-1} + \hat{T}\langle U \rangle_{(1)}T^{-1}, \quad \forall U \in C\ell_{p,q,r}, \quad T \in C\ell_{p,q,r}^\times. \quad (104)$$

Note that for any elements of fixed parity $U_{(0)}$ and $U_{(1)}$, we have

$$\tilde{\text{ad}}_T(U_{(0)}) = \text{ad}_T(U_{(0)}), \quad \forall U_{(0)} \in C\ell_{p,q,r}^{(0)}, \quad T \in C\ell_{p,q,r}^\times, \quad (105)$$

$$\tilde{\text{ad}}_T(U_{(1)}) = \check{\text{ad}}_T(U_{(1)}), \quad \forall U_{(1)} \in C\ell_{p,q,r}^{(1)}, \quad T \in C\ell_{p,q,r}^\times. \quad (106)$$

Note that both generalizations $\check{\text{ad}}$ (102) and $\tilde{\text{ad}}$ (104) of the twisted adjoint representation (101) to the case of arbitrary $T \in C\ell_{p,q,r}^\times$ and $U \in C\ell_{p,q,r}$ are used in the literature for various purposes and have some advantages indicated in the works cited above.

Let us consider the *kernels of the adjoint and the twisted adjoint representations*:

$$\ker(\text{ad}) = \{T \in C\ell_{p,q,r}^\times : TUT^{-1} = U, \quad \forall U \in C\ell_{p,q,r}\},$$

$$\ker(\check{\text{ad}}) = \{T \in C\ell_{p,q,r}^\times : \hat{T}UT^{-1} = U, \quad \forall U \in C\ell_{p,q,r}\},$$

$$\ker(\tilde{\text{ad}}) = \{T \in C\ell_{p,q,r}^\times : TU_{(0)}T^{-1} + \hat{T}U_{(1)}T^{-1} = U, \quad \forall U = U_{(0)} + U_{(1)} \in C\ell_{p,q,r}, \quad U_{(0)} \in C\ell_{p,q,r}^{(0)}, \quad U_{(1)} \in C\ell_{p,q,r}^{(1)}\}.$$

Lemma E.1. (Filimoshina & Shirokov, 2023) *We have*

$$\ker(\text{ad}) = Z_{p,q,r}^\times = \begin{cases} (\Lambda_r^{(0)} \oplus C\ell_{p,q,r}^n)^\times & \text{if } n \text{ is odd,} \\ \Lambda_r^{(0)\times} & \text{if } n \text{ is even;} \end{cases} \quad (107)$$

$$\ker(\check{\text{ad}}) = \Lambda_r^{(0)\times}; \quad (108)$$

$$\ker(\tilde{\text{ad}}) = \Lambda_r^\times. \quad (109)$$

Proof. We obtain (107) from (70).

Let us prove $\Lambda_r^{(0)\times} \subseteq \ker(\check{\text{ad}})$. Suppose $T \in \Lambda_r^{(0)\times}$; then $TUT^{-1} = U$ for any $U \in C\ell_{p,q,r}$. Since T is even, we have $\hat{T} = T$; therefore, $\hat{T}UT^{-1} = U$ for any $U \in C\ell_{p,q,r}$.

Let us prove $\ker(\tilde{\text{ad}}) \subseteq \Lambda_r^{(0)\times}$. Suppose $T \in \mathcal{C}\ell_{p,q,r}^\times$ satisfies $\widehat{T}UT^{-1} = U$ for any $U \in \mathcal{C}\ell_{p,q,r}$. Substituting the element $U = e$, we obtain $\widehat{T} = T$; hence, $T \in \mathcal{C}\ell_{p,q,r}^{(0)\times}$ and $TUT^{-1} = U$ for any $U \in \mathcal{C}\ell_{p,q,r}$. In other words, $T \in \mathcal{C}\ell_{p,q,r}^{(0)\times} \cap \ker(\text{ad})$. Using (107), we obtain $T \in \mathcal{C}\ell_{p,q,r}^{(0)\times} \cap (\Lambda_r^{(0)} \oplus \mathcal{C}\ell_{p,q,r}^n)^\times = \Lambda_r^{(0)\times}$ in the case of odd n , $T \in \Lambda_r^{(0)\times}$ in the case of even n , and the proof is completed.

Let us prove $\ker(\tilde{\text{ad}}) \subseteq \Lambda_r^\times$. Suppose $T \in \ker(\tilde{\text{ad}})$; then $\widehat{T}U_1T^{-1} = U_1$ for any $U_1 \in \mathcal{C}\ell_{p,q,r}^1$. Thus, $T \in \Lambda_r^\times$ by the statement (78) of Remark D.1.

Now we must only prove that $\Lambda_r^\times \subseteq \ker(\tilde{\text{ad}})$. Suppose $T \in \Lambda_r^\times$; then $T\mathcal{C}\ell_{p,q,r}^{(0)} = \mathcal{C}\ell_{p,q,r}^{(0)}T$ and $\widehat{T}\mathcal{C}\ell_{p,q,r}^1 = \mathcal{C}\ell_{p,q,r}^1T$ by Lemma D.3 and the statement (78) of Remark D.1 respectively. Since any odd basis element can be represented as a product of an odd number of generators, we obtain $\widehat{T}\mathcal{C}\ell_{p,q,r}^{(1)} = \mathcal{C}\ell_{p,q,r}^{(1)}T$. Thus, $TU_{(0)}T^{-1} + \widehat{T}U_{(1)}T^{-1} = U$ for all $U = U_{(0)} + U_{(1)} \in \mathcal{C}\ell_{p,q,r}$, where $U_{(0)} \in \mathcal{C}\ell_{p,q,r}^{(0)}$ and $U_{(1)} \in \mathcal{C}\ell_{p,q,r}^{(1)}$, and the proof is completed. $\square$

In the next lemma, we study the properties of ad , $\tilde{\text{ad}}$, and $\widehat{\text{ad}}$. Note that the multiplicativity of $\widehat{\text{ad}}_T$ (115) is proved in the particular case $T \in \mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times}$ in Ruhe et al., 2023, and we generalize this statement to $T \in (\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times})\Lambda_r^\times$.

Lemma E.2 (Lemma 2.1). *Let $H \in \mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times}$, $T \in (\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times})\Lambda_r^\times$, $W \in \mathcal{C}\ell_{p,q,r}^\times$, and $U, V \in \mathcal{C}\ell_{p,q,r}$. Then ad_W , $\tilde{\text{ad}}_W$, and $\widehat{\text{ad}}_W$ satisfy additivity:*

$$\text{ad}_W(U + V) = \text{ad}_W(U) + \text{ad}_W(V), \quad (110)$$

$$\tilde{\text{ad}}_W(U + V) = \tilde{\text{ad}}_W(U) + \tilde{\text{ad}}_W(V), \quad (111)$$

$$\widehat{\text{ad}}_W(U + V) = \widehat{\text{ad}}_W(U) + \widehat{\text{ad}}_W(V), \quad (112)$$

ad_W , $\tilde{\text{ad}}_H$, and $\widehat{\text{ad}}_W$ satisfy for any $c \in \mathcal{C}\ell^0$,

$$\text{ad}_W(c) = \tilde{\text{ad}}_W(c) = c, \quad \tilde{\text{ad}}_H(c) = (-1)^l c, \quad (113)$$

where $l = 0$ if $H \in \mathcal{C}\ell_{p,q,r}^{(0)\times}$ and $l = 1$ if $H \in \mathcal{C}\ell_{p,q,r}^{(1)\times}$. Moreover, ad_W and $\widehat{\text{ad}}_T$ satisfy multiplicativity:

$$\text{ad}_W(UV) = \text{ad}_W(U)\text{ad}_W(V), \quad (114)$$

$$\widehat{\text{ad}}_T(UV) = \widehat{\text{ad}}_T(U)\widehat{\text{ad}}_T(V). \quad (115)$$

Proof. The formulas (110)–(113) can be easily checked using the definitions of ad (100), $\tilde{\text{ad}}$ (102), and $\widehat{\text{ad}}$ (104). Let us prove (114) and (115). We have

$$\text{ad}_W(U)\text{ad}_W(V) = WUW^{-1}WVW^{-1} = \text{ad}_W(UV).$$

Since $U = \langle U \rangle_{(0)} + \langle U \rangle_{(1)}$, we get

$$\widehat{\text{ad}}_T(U)\widehat{\text{ad}}_T(V) = (T\langle U \rangle_{(0)}T^{-1} + \widehat{T}\langle U \rangle_{(1)}T^{-1})(T\langle V \rangle_{(0)}T^{-1} + \widehat{T}\langle V \rangle_{(1)}T^{-1}) \quad (116)$$

$$= T\langle U \rangle_{(0)}\langle V \rangle_{(0)}T^{-1} + \widehat{T}\langle U \rangle_{(1)}\langle V \rangle_{(0)}T^{-1} + T\langle U \rangle_{(0)}T^{-1}\widehat{T}\langle V \rangle_{(1)}T^{-1} + \widehat{T}\langle U \rangle_{(1)}T^{-1}\widehat{T}\langle V \rangle_{(1)}T^{-1}. \quad (117)$$

Now we apply the result of Theorem 4.7 Filimoshina & Shirokov, 2023:

$$(\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times})\Lambda_r^\times = \{T \in \mathcal{C}\ell_{p,q,r}^\times : \widehat{T^{-1}}T \in \Lambda_r^\times\}$$

and get $\widehat{T^{-1}}T \in \Lambda_r^\times$, which implies $T^{-1}\widehat{T} = \widehat{T^{-1}}T \in \Lambda_r^\times$. Since $\Lambda_r^\times = \ker(\tilde{\text{ad}})$ by Lemma E.1, we obtain that $(T^{-1}\widehat{T})X = X(T^{-1}\widehat{T})$ for any even $X \in \mathcal{C}\ell_{p,q,r}^{(0)}$ and $(T^{-1}\widehat{T})X = X(\widehat{T^{-1}}T)$ for any odd $X \in \mathcal{C}\ell_{p,q,r}^{(1)}$. Therefore,

$$T\langle U \rangle_{(0)}T^{-1}\widehat{T}\langle V \rangle_{(1)}T^{-1} = TT^{-1}\widehat{T}\langle U \rangle_{(0)}\langle V \rangle_{(1)}T^{-1} = \widehat{T}\langle U \rangle_{(0)}\langle V \rangle_{(1)}T^{-1}, \quad (118)$$

$$\widehat{T}\langle U \rangle_{(1)}T^{-1}\widehat{T}\langle V \rangle_{(1)}T^{-1} = \widehat{T}\widehat{T^{-1}}T\langle U \rangle_{(1)}\langle V \rangle_{(1)}T^{-1} = T\langle U \rangle_{(1)}\langle V \rangle_{(1)}T^{-1}. \quad (119)$$

Finally,

$$\widehat{\text{ad}}_T(U)\widehat{\text{ad}}_T(V) = T\langle U \rangle_{(0)}\langle V \rangle_{(0)}T^{-1} + \widehat{T}\langle U \rangle_{(1)}\langle V \rangle_{(0)}T^{-1} + \widehat{T}\langle U \rangle_{(0)}\langle V \rangle_{(1)}T^{-1} + T\langle U \rangle_{(1)}\langle V \rangle_{(1)}T^{-1} = \widehat{\text{ad}}_T(UV),$$

where we use that the product of two elements of the same parity is even and the product of two elements of different parity is odd. $\square$

F. Generalized Lipschitz Groups in $\mathcal{C}\ell_{p,q,r}$

In this section, we introduce and study degenerate and non-degenerate generalized Lipschitz groups in geometric algebras $\mathcal{C}\ell_{p,q,r}$ of arbitrary dimension and signature.

The *generalized Lipschitz groups* are defined as the groups preserving the subspaces $\mathcal{C}\ell_{p,q,r}^{\bar{k}}$ (4), $k = 0, 1, 2, 3$, determined by the grade involution and reversion, under the twisted adjoint representation $\tilde{\text{ad}}$ (104):

$$\tilde{\Gamma}_{p,q,r}^{\bar{k}} := \{T \in \mathcal{C}\ell_{p,q,r}^{\times} : \tilde{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^{\bar{k}}) \subseteq \mathcal{C}\ell_{p,q,r}^{\bar{k}}\}. \quad (120)$$

These groups are setwise stabilizers (see, for example, Isaacs, 2009) of the subspaces $\mathcal{C}\ell_{p,q,r}^{\bar{k}}$ (4), $k = 0, 1, 2, 3$, in the group $\mathcal{C}\ell_{p,q,r}^{\times}$ under the group action $\tilde{\text{ad}}$. To study these groups, we introduce and study the following groups, which are setwise stabilizers of $\mathcal{C}\ell_{p,q,r}^{\bar{k}}$, $k = 0, 1, 2, 3$, in $\mathcal{C}\ell_{p,q,r}^{\times}$ under ad (100) and $\check{\text{ad}}$ (102) respectively:

$$\Gamma_{p,q,r}^{\bar{k}} := \{T \in \mathcal{C}\ell_{p,q,r}^{\times} : \text{ad}_T(\mathcal{C}\ell_{p,q,r}^{\bar{k}}) := T\mathcal{C}\ell_{p,q,r}^{\bar{k}}T^{-1} \subseteq \mathcal{C}\ell_{p,q,r}^{\bar{k}}\}, \quad (121)$$

$$\check{\Gamma}_{p,q,r}^{\bar{k}} := \{T \in \mathcal{C}\ell_{p,q,r}^{\times} : \check{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^{\bar{k}}) := \hat{T}\mathcal{C}\ell_{p,q,r}^{\bar{k}}T^{-1} \subseteq \mathcal{C}\ell_{p,q,r}^{\bar{k}}\}, \quad (122)$$

and are related to $\tilde{\Gamma}_{p,q,r}^{\bar{k}}$ (120) as follows:

$$\tilde{\Gamma}_{p,q,r}^{\bar{k}} = \begin{cases} \check{\Gamma}_{p,q,r}^{\bar{k}}, & k = 1, 3, \\ \Gamma_{p,q,r}^{\bar{k}}, & k = 0, 2, \end{cases} \quad (123)$$

since $\tilde{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^{\bar{k}}) = \text{ad}_T(\mathcal{C}\ell_{p,q,r}^{\bar{k}})$ in the cases $k = 0, 2$ and $\tilde{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^{\bar{k}}) = \check{\text{ad}}_T(\mathcal{C}\ell_{p,q,r}^{\bar{k}})$ in the cases $k = 1, 3$ by (105)–(106).

In Subsections F.1 and F.2, we find the equivalent definitions of the groups $\Gamma_{p,q,r}^{\bar{k}}$ and $\check{\Gamma}_{p,q,r}^{\bar{k}}$, $k = 0, 1, 2, 3$.

F.1. The Groups $\mathcal{Q}_{p,q,r}^{\bar{0}}, \mathcal{Q}_{p,q,r}^{\bar{1}}, \mathcal{Q}_{p,q,r}^{\bar{2}}, \mathcal{Q}_{p,q,r}^{\bar{3}}, \Gamma_{p,q,r}^{\bar{0}}, \Gamma_{p,q,r}^{\bar{1}}, \Gamma_{p,q,r}^{\bar{2}},$ and $\Gamma_{p,q,r}^{\bar{3}}$

Let us consider the groups $\mathcal{Q}_{p,q,r}^{\bar{1}}, \mathcal{Q}_{p,q,r}^{\bar{2}}, \mathcal{Q}_{p,q,r}^{\bar{3}}$, and $\mathcal{Q}_{p,q,r}^{\bar{0}}$:

$$\mathcal{Q}_{p,q,r}^{\bar{1}} := \{T \in \mathcal{C}\ell_{p,q,r}^{\times} : \tilde{T}T \in Z_{p,q,r}^{1 \times} = \ker(\text{ad}), \quad \hat{T}T \in Z_{p,q,r}^{1 \times} = \ker(\text{ad})\}, \quad (124)$$

$$\mathcal{Q}_{p,q,r}^{\bar{2}} := \{T \in \mathcal{C}\ell_{p,q,r}^{\times} : \tilde{T}T \in Z_{p,q,r}^{2 \times}, \quad \hat{T}T \in Z_{p,q,r}^{2 \times}\}, \quad (125)$$

$$\mathcal{Q}_{p,q,r}^{\bar{3}} := \{T \in \mathcal{C}\ell_{p,q,r}^{\times} : \tilde{T}T \in Z_{p,q,r}^{3 \times}, \quad \hat{T}T \in Z_{p,q,r}^{3 \times}\}, \quad (126)$$

$$\mathcal{Q}_{p,q,r}^{\bar{0}} := \{T \in \mathcal{C}\ell_{p,q,r}^{\times} : \tilde{T}T \in Z_{p,q,r}^{4 \times}, \quad \hat{T}T \in Z_{p,q,r}^{4 \times}\}, \quad (127)$$

where $\ker(\text{ad})$ (107) is the kernel of the adjoint representation ad (100), $Z_{p,q,r}^{1 \times}$, $Z_{p,q,r}^{2 \times}$, $Z_{p,q,r}^{3 \times}$, and $Z_{p,q,r}^{4 \times}$ (see Remark D.1) are the centralizers of the subspaces $\mathcal{C}\ell_{p,q,r}^{\bar{1}}$, $\mathcal{C}\ell_{p,q,r}^{\bar{2}}$, $\mathcal{C}\ell_{p,q,r}^{\bar{3}}$, and $\mathcal{C}\ell_{p,q,r}^{\bar{4}}$ respectively considered in Section D.

The groups $\mathcal{Q}_{p,q,r}^{\bar{1}}, \mathcal{Q}_{p,q,r}^{\bar{2}}, \mathcal{Q}_{p,q,r}^{\bar{3}}$, and $\mathcal{Q}_{p,q,r}^{\bar{0}}$ are generalizations of the groups $\mathcal{Q}$ and $\mathcal{Q}'$ (Shirokov, 2021) in the non-degenerate geometric algebras $\mathcal{C}\ell_{p,q}$ to the case of the degenerate geometric algebras $\mathcal{C}\ell_{p,q,r}$ and coincide with them if $r = 0$:

$$\mathcal{Q}_{p,q}^{\bar{1}} = \mathcal{Q}_{p,q}^{\bar{3}} = \mathcal{Q}, \quad \mathcal{Q}_{p,q}^{\bar{0}} = \mathcal{Q}_{p,q}^{\bar{2}} = \begin{cases} \mathcal{Q}, & n = 1, 2, 3 \bmod 4, \\ \mathcal{Q}', & n = 0 \bmod 4, \end{cases} \quad n \neq 2, 3,$$

and $\mathcal{Q}_{p,q}^{\bar{1}} = \mathcal{Q}_{p,q}^{\bar{2}} = \mathcal{Q}$ if $n = 2, 3$.

Theorem F.1 (Formulas (22) and (25) of Theorem 3.2). *In the degenerate and non-degenerate geometric algebras $\mathcal{C}\ell_{p,q,r}$, we have*

$$\mathcal{Q}_{p,q,r}^{\bar{1}} = \Gamma_{p,q,r}^{\bar{1}}, \quad \mathcal{Q}_{p,q,r}^{\bar{2}} = \Gamma_{p,q,r}^{\bar{2}}, \quad \mathcal{Q}_{p,q,r}^{\bar{3}} = \Gamma_{p,q,r}^{\bar{3}}, \quad \mathcal{Q}_{p,q,r}^{\bar{0}} = \Gamma_{p,q,r}^{\bar{0}},$$

where

$$\Gamma_{p,q,r}^{\bar{1}} \subseteq \Gamma_{p,q,r}^{\bar{m}} \subseteq \Gamma_{p,q,r}^{\bar{0}}, \quad m = 2, 3. \quad (128)$$

Proof. For fixed $k = 0, 1, 2, 3 \bmod 4$ and $m = (k - 1) \bmod 4$, $l = (k + 1) \bmod 4$, we get

$$\Gamma_{p,q,r}^{\bar{k}} = \{T \in C\ell_{p,q,r}^{\times} : TC\ell_{p,q,r}^{\bar{k}}T^{-1} \subseteq (C\ell_{p,q,r}^{\bar{k}m} \cap C\ell_{p,q,r}^{\bar{k}l})\} \quad (129)$$

$$= \{T \in C\ell_{p,q,r}^{\times} : TC\ell_{p,q,r}^{\bar{k}}T^{-1} \subseteq C\ell_{p,q,r}^{\bar{k}m}, \quad TC\ell_{p,q,r}^{\bar{k}}T^{-1} \subseteq C\ell_{p,q,r}^{\bar{k}l}\} \quad (130)$$

$$= \{T \in C\ell_{p,q,r}^{\times} : \tilde{T}T \in Z_{p,q,r}^{\bar{k} \times}, \quad \hat{T}T \in Z_{p,q,r}^{\bar{k} \times}\} = Q_{p,q,r}^{\bar{k}}, \quad (131)$$

where we use Lemma D.5 in the first equality (131) and $Z_{p,q,r}^{\bar{k}} = Z_{p,q,r}^k$, $k = 1, 2, 3$, and $Z_{p,q,r}^{\bar{0}} = Z_{p,q,r}^4$ by Lemma D.4 in the second equality (131). The inclusions in (128) follow from the definitions (124)–(127) and the following facts about explicit forms of centralizers and twisted centralizers implied by Remark D.1:

$$\ker(\text{ad}) = Z_{p,q,r}^{1 \times} \subseteq Z_{p,q,r}^{m \times} \subseteq Z_{p,q,r}^{4 \times}, \quad m = 2, 3. \quad (132)$$

This completes the proof. $\square$

F.2. The Groups $\check{Q}_{p,q,r}^{\bar{0}}, \check{Q}_{p,q,r}^{\bar{1}}, \check{Q}_{p,q,r}^{\bar{2}}, \check{Q}_{p,q,r}^{\bar{3}}, \check{\Gamma}_{p,q,r}^{\bar{0}}, \check{\Gamma}_{p,q,r}^{\bar{1}}, \check{\Gamma}_{p,q,r}^{\bar{2}},$ and $\check{\Gamma}_{p,q,r}^{\bar{3}}$

Consider the groups $\check{Q}_{p,q,r}^{\bar{1}}, \check{Q}_{p,q,r}^{\bar{2}}, \check{Q}_{p,q,r}^{\bar{3}},$ and $\check{Q}_{p,q,r}^{\bar{0}}$:

$$\check{Q}_{p,q,r}^{\bar{1}} := \{T \in C\ell_{p,q,r}^{\times} : \tilde{T}T \in \check{Z}_{p,q,r}^{1 \times} = \ker(\tilde{\text{ad}}), \quad \hat{T}T \in \check{Z}_{p,q,r}^{1 \times} = \ker(\hat{\text{ad}})\}, \quad (133)$$

$$\check{Q}_{p,q,r}^{\bar{2}} := \{T \in C\ell_{p,q,r}^{\times} : \tilde{T}T \in \check{Z}_{p,q,r}^{2 \times}, \quad \hat{T}T \in \check{Z}_{p,q,r}^{2 \times}\}, \quad (134)$$

$$\check{Q}_{p,q,r}^{\bar{3}} := \{T \in C\ell_{p,q,r}^{\times} : \tilde{T}T \in \check{Z}_{p,q,r}^{3 \times}, \quad \hat{T}T \in \check{Z}_{p,q,r}^{3 \times}\}, \quad (135)$$

$$\check{Q}_{p,q,r}^{\bar{0}} := \{T \in C\ell_{p,q,r}^{\times} : \tilde{T}T \in \langle Z_{p,q,r}^4 \rangle_{(0)}^{\times}, \quad \hat{T}T \in \langle Z_{p,q,r}^4 \rangle_{(0)}^{\times}\}, \quad (136)$$

where $\ker(\tilde{\text{ad}})$ (109) is the kernel of the twisted adjoint representation $\tilde{\text{ad}}$ (104), the sets $\check{Z}_{p,q,r}^2$ (79) and $\check{Z}_{p,q,r}^3$ (80) are the twisted centralizers of the subspaces $C\ell_{p,q,r}^2$ and $C\ell_{p,q,r}^3$ respectively, and $Z_{p,q,r}^4$ (Remark D.1) is the centralizer of the subspace $C\ell_{p,q,r}^4$ (see Appendix D). We use $\check{\cdot}$ in the notation of the groups (133)–(136) due to Theorem F.2 below. We have (see (109)):

$$\langle Z_{p,q,r}^4 \rangle_{(0)} = \begin{cases} \Lambda_r^{(0)} \oplus \{C\ell_{p,q}^1 \Lambda_r^{n-2}\} \oplus \{C\ell_{p,q}^2 \Lambda_r^{n-3}\}, & n \text{ is odd;} \\ \Lambda_r^{(0)} \oplus \{C\ell_{p,q}^1 \Lambda_r^{n-3}\} \oplus \{C\ell_{p,q}^2 \Lambda_r^{n-4}\} \oplus C\ell_{p,q}^n, & n \text{ is even, } r \neq n; \\ \Lambda_r^{(0)}, & n \text{ is even, } r = n. \end{cases} \quad (137)$$

In the particular case of the non-degenerate geometric algebra $C\ell_{p,q}$, we obtain the groups $Q^{\pm}$ and Q' considered in Filimoshina & Shirokov, 2024a:

$$\check{Q}_{p,q}^{\bar{1}} = \check{Q}_{p,q}^{\bar{3}} = Q^{\pm}, \quad \check{Q}_{p,q}^{\bar{0}} = \check{Q}_{p,q}^{\bar{2}} = \begin{cases} Q^{\pm}, & n = 1, 2, 3 \bmod 4, \\ Q', & n = 0 \bmod 4, \end{cases} \quad n \neq 1, 2,$$

and $\check{Q}_{p,q}^{\bar{1}} = \check{Q}_{p,q}^{\bar{0}} = Q^{\pm}$ if $n = 1, 2$.

Theorem F.2 (Formulas (23)–(24) of Theorem 3.2). *In degenerate and non-degenerate geometric algebras $C\ell_{p,q,r}$, we have*

$$\check{Q}_{p,q,r}^{\bar{1}} = \check{\Gamma}_{p,q,r}^{\bar{1}} \subseteq \check{Q}_{p,q,r}^{\bar{3}} = \check{\Gamma}_{p,q,r}^{\bar{3}}, \quad \check{Q}_{p,q,r}^{\bar{2}} = \check{\Gamma}_{p,q,r}^{\bar{2}}, \quad \check{Q}_{p,q,r}^{\bar{0}} = \check{\Gamma}_{p,q,r}^{\bar{0}}.$$

Proof. We get $\check{Q}_{p,q,r}^{\bar{1}} \subseteq \check{Q}_{p,q,r}^{\bar{3}}$, using $\check{Z}_{p,q,r}^1 \subseteq \check{Z}_{p,q,r}^3$ (see Remark D.1). For fixed $k = 0, 1, 2, 3 \bmod 4$ and $m = (k - 1) \bmod 4$, $l = (k + 1) \bmod 4$, we obtain

$$\check{\Gamma}_{p,q,r}^{\bar{k}} = \{T \in C\ell_{p,q,r}^{\times} : \hat{T}C\ell_{p,q,r}^{\bar{k}}T^{-1} \subseteq (C\ell_{p,q,r}^{\bar{k}m} \cap C\ell_{p,q,r}^{\bar{k}l})\} \quad (138)$$

$$= \{T \in C\ell_{p,q,r}^{\times} : \hat{T}C\ell_{p,q,r}^{\bar{k}}T^{-1} \subseteq C\ell_{p,q,r}^{\bar{k}m}, \quad \hat{T}C\ell_{p,q,r}^{\bar{k}}T^{-1} \subseteq C\ell_{p,q,r}^{\bar{k}l}\} \quad (139)$$

$$= \{T \in C\ell_{p,q,r}^{\times} : \tilde{T}T \in \check{Z}_{p,q,r}^{\bar{k} \times}, \quad \hat{T}T \in \check{Z}_{p,q,r}^{\bar{k} \times}\} = \check{Q}_{p,q,r}^{\bar{k}}, \quad (140)$$

where we use Lemma D.5 in the first equality (140) and $\check{Z}_{p,q,r}^{\bar{k}} = \check{Z}_{p,q,r}^k$, $k = 1, 2, 3$, and $\check{Z}_{p,q,r}^{\bar{0}} = \check{Z}_{p,q,r}^4 \cap C\ell_{p,q,r}^{(0)} = Z_{p,q,r}^4 \cap C\ell_{p,q,r}^{(0)}$ by Lemma D.4 and (73) in the second equality (140). $\square$

Remark F.3 (Formula (26) of Theorem 3.2). We have the following relations between the groups $\Gamma_{p,q,r}^{\bar{k}}$ (121) and $\tilde{\Gamma}_{p,q,r}^{\bar{k}}$ (122), $k = 0, 1, 2, 3$:

$$\tilde{\Gamma}_{p,q,r}^{\bar{m}} \subseteq \Gamma_{p,q,r}^{\bar{0}}, \quad m = 0, 1, 2, 3; \quad \tilde{\Gamma}_{p,q,r}^{\bar{1}} \subseteq \Gamma_{p,q,r}^{\bar{2}}. \quad (141)$$

The statements follow from Theorems F.1 and F.2, the definitions of the groups $\check{Q}_{p,q,r}^{\bar{m}}$ (133)–(136), $m = 0, 1, 2, 3$, $Q_{p,q,r}^{\bar{0}}$ (127), $Q_{p,q,r}^{\bar{2}}$ (125), $\check{Z}_{p,q,r}^1 \subseteq Z_{p,q,r}^2$, and $\check{Z}_{p,q,r}^k \subseteq Z_{p,q,r}^4$, $k = 1, 2, 3$, by Remark D.1.

F.3. The Groups $\tilde{\Gamma}_{p,q,r}^{\bar{0}}$, $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$, $\tilde{\Gamma}_{p,q,r}^{\bar{2}}$, and $\tilde{\Gamma}_{p,q,r}^{\bar{3}}$

In Corollary F.4, we summarize the results on the equivalent definitions of the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^{\bar{k}}$ (120), $k = 0, 1, 2, 3$. They follow from Theorems F.1 and F.2 on the equivalent definitions of the groups $\Gamma_{p,q,r}^{\bar{k}}$ (121) and $\tilde{\Gamma}_{p,q,r}^{\bar{k}}$ (122), $k = 0, 1, 2, 3$, respectively, Remark F.3, and the relation (123) between these groups and the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^{\bar{k}}$.

Corollary F.4 (Formula (14) of Theorem 3.1). *In the case of arbitrary $\mathcal{C}\ell_{p,q,r}$, we have*

$$\tilde{\Gamma}_{p,q,r}^{\bar{1}} = \check{Q}_{p,q,r}^{\bar{1}} \subseteq \tilde{\Gamma}_{p,q,r}^{\bar{2}} = Q_{p,q,r}^{\bar{2}} \subseteq \tilde{\Gamma}_{p,q,r}^{\bar{0}} = Q_{p,q,r}^{\bar{0}}, \quad (142)$$

$$\tilde{\Gamma}_{p,q,r}^{\bar{1}} \subseteq \tilde{\Gamma}_{p,q,r}^{\bar{3}} = \check{Q}_{p,q,r}^{\bar{3}} \quad (143)$$

Theorem F.5 (Theorem 3.3). *The elements of the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ have the following special form:*

$$\tilde{\Gamma}_{p,q,r}^{\bar{1}} \subseteq (\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times}) \Lambda_r^\times. \quad (144)$$

Proof. Suppose $T \in \tilde{\Gamma}_{p,q,r}^{\bar{1}}$; then $\tilde{T}T = W \in \ker(\tilde{\text{ad}})$ and $\widehat{\tilde{T}}T = U \in \ker(\tilde{\text{ad}})$. Therefore, $T = \widetilde{T^{-1}}W$ and $\widehat{T^{-1}} = \widehat{U^{-1}}\tilde{T}$. Thus, $\widehat{T^{-1}}T = (\widehat{U^{-1}}\tilde{T})(\widetilde{T^{-1}}W) = \widehat{U^{-1}}W \in \ker(\tilde{\text{ad}}) = \Lambda_r^\times$. Now we apply Theorem 4.7 (Filimoshina & Shirokov, 2023):

$$(\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times}) \Lambda_r^\times = \{T \in \mathcal{C}\ell_{p,q,r}^\times : \widehat{T^{-1}}T \in \Lambda_r^\times\}.$$

and obtain $T \in (\mathcal{C}\ell_{p,q,r}^{(0)\times} \cup \mathcal{C}\ell_{p,q,r}^{(1)\times}) \Lambda_r^\times$. $\square$

F.4. Relation Between Ordinary and Generalized Lipschitz Groups

The group $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ may be considered the most important among the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^{\bar{k}}$ (120), $k = 0, 1, 2, 3$, because its elements preserve under $\tilde{\text{ad}}$ not only the subspace $\mathcal{C}\ell_{p,q,r}^{\bar{1}}$, but also all other subspaces $\mathcal{C}\ell_{p,q,r}^{\bar{k}}$, $k = 1, 2, 3$, by Corollary F.4. Moreover, the group $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ coincides with the ordinary Lipschitz group $\tilde{\Gamma}_{p,q,r}^1$ (10) in the case of the low-dimensional geometric algebras $\mathcal{C}\ell_{p,q,r}$ (Remark F.6 below) and contains it as a subgroup in the case of arbitrary $\mathcal{C}\ell_{p,q,r}$ (Theorem F.7).

Remark F.6 (Formula (15) of Theorem 3.1). In the case of the small dimensions $n \leq 4$, the (ordinary) Lipschitz groups (10) coincide with the generalized Lipschitz groups $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ due to $\mathcal{C}\ell_{p,q,r}^{\bar{1}} = \mathcal{C}\ell_{p,q,r}^1$:

$$\tilde{\Gamma}_{p,q,r}^1 = \tilde{\Gamma}_{p,q,r}^{\bar{1}}, \quad n \leq 4. \quad (145)$$

Note that in the particular case of the non-degenerate geometric algebra $\mathcal{C}\ell_{p,q}$, it is proved (Filimoshina & Shirokov, 2024a) that the groups coincide in the case $n = 5$ as well:

$$\tilde{\Gamma}_{p,q}^1 = \tilde{\Gamma}_{p,q}^{\bar{1}}, \quad n \leq 5. \quad (146)$$

Theorem F.7 (Formula (14) of Theorem 3.1). *In arbitrary $\mathcal{C}\ell_{p,q,r}$, the (ordinary) Lipschitz group $\tilde{\Gamma}_{p,q,r}^1$ (10) is a subgroup of the generalized Lipschitz group $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ (120):*

$$\tilde{\Gamma}_{p,q,r}^1 \subseteq \tilde{\Gamma}_{p,q,r}^{\bar{1}}. \quad (147)$$

Proof. Suppose $T \in \tilde{\Gamma}_{p,q,r}^1$, then $T \in \mathcal{C}\ell_{p,q,r}^\times$ and $\widehat{T}U_1T^{-1} \in \mathcal{C}\ell_{p,q,r}^1$ for any $U_1 \in \mathcal{C}\ell_{p,q,r}^1$ by definition. We get

$$\widehat{T}U_1T^{-1} = (\widehat{T}U_1T^{-1})^\sim = \widetilde{T^{-1}}\widetilde{U_1}\widehat{T} = \widetilde{T^{-1}}U_1\widehat{T}, \quad \forall U_1 \in \mathcal{C}\ell_{p,q,r}^1, \quad (148)$$

$$\widehat{T}U_1T^{-1} = -(\widehat{T}U_1T^{-1})^\wedge = -\widetilde{\widehat{T^{-1}}}\widetilde{\widehat{U_1}}\widetilde{\widehat{T}} = \widetilde{\widehat{T^{-1}}}U_1\widetilde{\widehat{T}}, \quad \forall U_1 \in \mathcal{C}\ell_{p,q,r}^1, \quad (149)$$

where we use $\mathcal{C}\ell_{p,q,r}^1 \subseteq \mathcal{C}\ell_{p,q,r}^{\bar{1}}$, Table 1, and the properties of the reversion $\widetilde{UV} = \widetilde{V}\widetilde{U}$, Clifford conjugation $\widehat{UV} = \widehat{V}\widehat{U}$, and grade involution $\widehat{U} = U$, for any $U, V \in \mathcal{C}\ell_{p,q,r}$. We multiply both sides of the equality (148) on the left by $\widetilde{T}$ and on the right by T , both sides of the equality (149) on the left by $\widetilde{\widehat{T}}$ and on the right by T , and get

$$(\widetilde{\widehat{T}}T)U_1 = U_1(\widetilde{\widehat{T}}T), \quad (\widetilde{\widehat{T}}T)U_1 = U_1(\widetilde{\widehat{T}}T), \quad \forall U_1 \in \mathcal{C}\ell_{p,q,r}^1. \quad (150)$$

Therefore, by the definition (72) of the twisted centralizers, the elements $\widetilde{\widehat{T}}T$ and $\widetilde{\widehat{T}}T$ belong to the twisted centralizer of the grade-1 subspace $\mathcal{C}\ell_{p,q,r}^1$ in $\mathcal{C}\ell_{p,q,r}$:

$$\widetilde{\widehat{T}}T \in \check{Z}_{p,q,r}^1, \quad \widetilde{\widehat{T}}T \in \check{Z}_{p,q,r}^1. \quad (151)$$

We have $\check{Z}_{p,q,r}^1 = \Lambda_r$ by Remark D.1. Therefore, $\widetilde{\widehat{T}}T, \widetilde{\widehat{T}}T \in \Lambda_r^\times$, and $T \in \check{Q}_{p,q,r}^{\bar{1}}$ by definition (133). By Corollary F.4, $\check{Q}_{p,q,r}^{\bar{1}} = \tilde{\Gamma}_{p,q,r}^{\bar{1}}$. Thus, $T \in \tilde{\Gamma}_{p,q,r}^{\bar{1}}$, and the proof is completed. $\square$

G. Equivariant Mappings

Let G be a group and X be a set. A (left) *group action* is a map: $\circ : G \times X \rightarrow X$, $(g, x) \mapsto g \circ x$ that satisfies associativity (i.e. $(gh) \circ x = g \circ (h \circ x)$ for any $g, h \in G$ and $x \in X$) and identity condition (i.e. $e \circ x = x$ for any $x \in X$).

Suppose G is a group and $\circ_X$ and $\circ_Y$ are its actions on the sets X and Y respectively. A *function* (network) $L : X \rightarrow Y$ is called *G-equivariant* iff it commutes with these actions:

$$L(g \circ_X x) = g \circ_Y L(x), \quad \forall g \in G, \quad \forall x \in X. \quad (152)$$

Example G.1. Consider the generalized Lipschitz group $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ with the action $\tilde{\text{ad}}$ and a function $L : \mathcal{C}\ell_{p,q,r} \rightarrow \mathcal{C}\ell_{p,q,r}$. This function is $\tilde{\Gamma}_{p,q,r}^{\bar{1}}$ -equivariant iff

$$L(\tilde{\text{ad}}_T(x)) = \tilde{\text{ad}}_T(L(x)), \quad \forall T \in \tilde{\Gamma}_{p,q,r}^{\bar{1}}, \quad \forall x \in \mathcal{C}\ell_{p,q,r}, \quad (153)$$

i.e.

$$L(T\langle x \rangle_{(0)}T^{-1} + \widehat{T}\langle x \rangle_{(1)}T^{-1}) = T\langle L(x) \rangle_{(0)}T^{-1} + \widehat{T}\langle L(x) \rangle_{(1)}T^{-1}, \quad \forall T \in \tilde{\Gamma}_{p,q,r}^{\bar{1}}, \quad \forall x \in \mathcal{C}\ell_{p,q,r}. \quad (154)$$

We prove several general statements about equivariance with respect to an arbitrary group. Suppose G is a group, $\circ_X$ is its action defined on a set X .

Lemma G.2. Consider k G -equivariant mappings $f_i : X_{i-1} \rightarrow X_i$, $i = 1, \dots, k$. Their composition is G -equivariant:

$$f_k(\cdots(f_2(f_1(g \circ_{X_0} x)))) = g \circ_{X_k} f_k(\cdots(f_2(f_1(x)))) \quad (155)$$

for any $g \in G$ and $x \in X_0$.

Proof. We prove the statement by induction, using

$$f_2(f_1(g \circ_{X_0} x)) = f_2(g \circ_{X_1} (f_1(x))) = g \circ_{X_2} f_2(f_1(x)).$$

$\square$

Lemma G.3. Suppose group G action $\circ$ is linear, i.e. for any $g \in G$, $x, y \in X$, and $\alpha \in \mathbb{F}$

$$g \circ_Y (\alpha x) = \alpha(g \circ_X x), \quad g \circ_Y (x + y) = g \circ_X x + g \circ_X y.$$

Consider k G -equivariant mappings $f_1, \dots, f_k : X \rightarrow Y$. Their linear combinations are G -equivariant:

$$\sum_{i=1}^k \alpha_i f_i(g \circ_X x) = g \circ_Y \left(\sum_{i=1}^k \alpha_i f_i(x) \right) \quad (156)$$

for any $g \in G$, $x \in X$, $\alpha_i \in \mathbb{F}$.

Proof. We have

$$\alpha_1 f_1(g \circ_X x) + \alpha_2 f_2(g \circ_X x) = \alpha_1 (g \circ_Y f_1(x)) + \alpha_2 (g \circ_Y f_2(x)) = g \circ_Y (\alpha_1 f_1(x) + \alpha_2 f_2(x)) \quad (157)$$

and obtain the result similarly for k summands. $\square$

Lemma G.4. Suppose group G action $\circ$ is multiplicative, i.e.

$$g \circ_Y (xy) = (g \circ_X x)(g \circ_X y), \quad \forall g \in G, \quad x, y \in X.$$

The product of two G -equivariant mappings $f_1, f_2 : X \rightarrow Y$ is G -equivariant:

$$g \circ_Y (f_1(x)f_2(x)) = f_1(g \circ_X x)f_2(g \circ_X x) \quad (158)$$

for any $g \in G$ and $x \in X$.

Proof. We have

$$g \circ_Y (f_1(x)f_2(x)) = (g \circ_Y f_1(x))(g \circ_Y f_2(x)) = f_1(g \circ_X x)f_2(g \circ_X x). \quad (159)$$

$\square$

H. Degenerate and Non-degenerate Lipschitz, Generalized Lipschitz, and Orthogonal Groups Equivariance

In this section, we consider the relation between the degenerate and non-degenerate pseudo-orthogonal (complex orthogonal) groups and the ordinary and generalized Lipschitz groups. We extend the work presented in [CrumeYrolle, 1990](#); [Brooke, 1980](#); [Dereli et al., 2010](#); [Ruhe et al., 2023](#), etc.

Consider the degenerate and non-degenerate *pseudo-orthogonal group* (in the real case $V = \mathbb{R}^{p,q,r}$) or *complex orthogonal group* (in the complex case $V = \mathbb{C}^{p+q,0,r}$) denoted by $O(V, q)$. It is defined as the Lie group of all linear transformations of an n -dimensional vector space V that leave invariant a quadratic form q of signature (p, q, r) if V is real and $(p + q, 0, r)$ if V is complex:

$$O(V, q) := \{ \Phi : V \rightarrow V : \quad \Phi \text{ is linear, invertible, } q(\Phi(v)) = q(v), \quad \forall v \in V \}. \quad (160)$$

Note that

$$O(V, q) \cong \{ A \in GL(n, \mathbb{F}) : A^T \eta A = \eta \}.$$

When considering both the real and complex cases, we refer to the group $O(V, q)$ as the orthogonal group.

Let us use the following notation. The *radical subspace* is denoted by $\Lambda_r^1 := \mathcal{C}\ell_{0,0,r}^1$. We have $V = \mathcal{C}\ell_{p,q,r}^1 = \mathcal{C}\ell_{p,q}^1 \oplus \Lambda_r^1$. The subalgebra generated by the basis elements of Λ_r^1 is denoted by Λ_r and is a *Grassmann (exterior) algebra*.

Consider the following subgroup of the orthogonal group, which leaves invariant the radical subspace Λ_r^1 :

$$O_{\Lambda_r^1}(V, q) := \{ \Phi \in O(V, q) : \quad \Phi|_{\Lambda_r^1} = \text{id}_{\Lambda_r^1} \}. \quad (161)$$

The following statement about the matrix forms of the orthogonal $O(V, q)$ and restricted orthogonal $O_{\Lambda_r^1}(V, q)$ groups is well-known and considered, for example, in [CrumeYrolle, 1990](#); [Ruhe et al., 2023](#). We use it in the proof of the main statements about the relation between the Lipschitz and orthogonal groups below.

Lemma H.1. We have the following isomorphisms

$$O(V, \mathfrak{q}) \cong \left\{ \begin{pmatrix} A & 0_{(p+q) \times r} \\ M & G \end{pmatrix} \right\}, \quad A \in O(C\ell_{p,q}^1, \mathfrak{q}|_{C\ell_{p,q}^1}), \quad M \in \text{Mat}_{r \times (p+q)}(\mathbb{F}), \quad G \in \text{GL}(r, \mathbb{F}); \quad (162)$$

$$O_{\Lambda_r^1}(V, \mathfrak{q}) \cong \left\{ \begin{pmatrix} A & 0_{(p+q) \times r} \\ M & I_r \end{pmatrix} \right\}, \quad A \in O(C\ell_{p,q}^1, \mathfrak{q}|_{C\ell_{p,q}^1}), \quad M \in \text{Mat}_{r \times (p+q)}(\mathbb{F}), \quad (163)$$

where $O(C\ell_{p,q}^1, \mathfrak{q}|_{C\ell_{p,q}^1})$ is the non-degenerate orthogonal group of transformations of $C\ell_{p,q}^1$ with a non-degenerate quadratic form $\mathfrak{q}|_{C\ell_{p,q}^1}$; $0_{(p+q) \times r}$ is the zero matrix of size $(p+q) \times r$; $\text{Mat}_{r \times (p+q)}(\mathbb{F})$ is a set of arbitrary matrices of size $r \times (p+q)$ with coefficients in $\mathbb{F}$; I_r is the identity matrix of size $r \times r$; $\text{GL}(r, \mathbb{F})$ is the general linear group acting on an r -dimensional vector space over $\mathbb{F}$.

Theorem H.2 (Kernel of $\tilde{\text{ad}}^1$). The kernel of the restricted twisted adjoint representation $\tilde{\text{ad}}^1$ acting on the Lipschitz group $\tilde{\text{ad}}^1 : \tilde{\Gamma}_{p,q,r}^1 \rightarrow \text{Aut}(C\ell_{p,q,r})$ has the following form:

$$\ker(\tilde{\text{ad}}^1 : \tilde{\Gamma}_{p,q,r}^1 \rightarrow \text{Aut}(C\ell_{p,q,r})) = \Lambda_r^\times. \quad (164)$$

Proof. We have

$$\ker(\tilde{\text{ad}}^1 : \tilde{\Gamma}_{p,q,r}^1 \rightarrow \text{Aut}(C\ell_{p,q,r})) = \{T \in \tilde{\Gamma}_{p,q,r}^1 : \hat{T}vT^{-1} = v, \quad \forall v \in C\ell_{p,q,r}^1\} \quad (165)$$

$$= \tilde{\Gamma}_{p,q,r}^1 \cap \{T \in C\ell_{p,q,r}^\times : \hat{T}vT^{-1} = v, \quad \forall v \in C\ell_{p,q,r}^1\} \quad (166)$$

$$= \tilde{\Gamma}_{p,q,r}^1 \cap \Lambda_r^\times, \quad (167)$$

where in (167), we apply the statement (109) of Lemma E.1. Note that $\Lambda_r^\times \subseteq \tilde{\Gamma}_{p,q,r}^1$, since for any $T \in \Lambda_r^\times$ and any $v \in C\ell_{p,q,r}^1$, we have $\hat{T}vT^{-1} = vTT^{-1} = v \in C\ell_{p,q,r}^1$ again by Lemma E.1. Thus, we obtain (164). $\square$

Remark H.3. Note that for any vectors $v, v_1, v_2 \in V$, we have

$$2\mathfrak{b}(v_1, v_2)e = v_1v_2 + v_2v_1, \quad \mathfrak{q}(v)e = v^2. \quad (168)$$

In Remark H.4, we consider how reflections can be represented in geometric algebras. We use this particular case of the relation between elements of the orthogonal groups $O(V, \mathfrak{q})$ and the Lipschitz groups $\tilde{\Gamma}_{p,q,r}^1$ in the proof of Theorem H.9 below.

Remark H.4. The mapping $\tilde{\text{ad}}_v^1$, where $v \in C\ell_{p,q}^{1\times}$ is an invertible vector, acts on an arbitrary vector $x \in C\ell_{p,q,r}^1$ as a reflection of a vector x across the hyperplane orthogonal to the vector v :

$$\tilde{\text{ad}}_v^1(x) = \hat{v}xv^{-1} = -v xv^{-1} = x - (xv + vx)v^{-1} = x - 2\mathfrak{b}(x, v)\frac{v^2}{\mathfrak{q}(v)}v^{-1} = x - 2\frac{\mathfrak{b}(x, v)}{\mathfrak{b}(v, v)}v, \quad (169)$$

where in the right-hand side, we have the difference between the vector x and twice the projection of the vector x onto the vector v .

In Remarks H.5 and H.6, we consider several examples of elements that belong to the Lipschitz group $\tilde{\Gamma}_{p,q,r}^1$ (10).

Remark H.5. All invertible non-degenerate vectors belong to the Lipschitz group:

$$C\ell_{p,q}^{1\times} \subseteq \tilde{\Gamma}_{p,q,r}^1, \quad (170)$$

since, by Remark H.4, we have $\tilde{\text{ad}}_v^1(x) = x - 2\frac{\mathfrak{b}(x, v)}{\mathfrak{b}(v, v)}v \in C\ell_{p,q,r}^1$ for any $x \in C\ell_{p,q,r}^1$ and $v \in C\ell_{p,q}^{1\times}$.

Remark H.6. Any element $U \in C\ell_{p,q,r}$ of the form

$$U = e + me_{ab} \in C\ell^0 \oplus C\ell_{p,q,r}^2, \quad \forall m \in \mathbb{F}, \quad \forall e_a \in C\ell_{p,q}^1, \quad \forall e_b \in \Lambda_r^1, \quad (171)$$

is invertible and belongs to the Lipschitz group:

$$U \in \tilde{\Gamma}_{p,q,r}^1. \quad (172)$$

Let us prove these statement. Firstly, U is invertible, since $(e + me_{ab})(e - me_{ab}) = e$. Secondly U preserves the grade-1 subspace under $\tilde{\text{ad}}$, since

$$\hat{U}e_a U^{-1} = (e + me_{ab})e_a(e - me_{ab}) = (e_a - m\eta_{aa}e_b)(e - me_{ab}) = e_a - 2m\eta_{aa}e_b \in \mathcal{C}\ell_{p,q,r}^1, \quad (173)$$

$$\hat{U}e_b U^{-1} = (e + me_{ab})e_b(e - me_{ab}) = e_b \in \mathcal{C}\ell_{p,q,r}^1, \quad (174)$$

$$\hat{U}e_c U^{-1} = (e + me_{ab})e_c(e - me_{ab}) = (e_c + me_{ab}e_c)(e - me_{ab}) = e_c \in \mathcal{C}\ell_{p,q,r}^1, \quad \forall c = 1, \dots, n, \quad c \neq a, b. \quad (175)$$

Thus, by definition of the Lipschitz group $\tilde{\Gamma}_{p,q,r}^1$ (10), we get (172).

Let us prove auxiliary Lemmas H.7 and H.8, which we use in the proof of Theorem H.9.

Lemma H.7. *In the case $r \neq 0$ and $r \neq n$, consider any matrix of the form*

$$B_{i,j} = \begin{pmatrix} I_{p+q} & 0 \\ M_{i,j} & I_r \end{pmatrix} \in \text{O}_{\Lambda_r^1}(V, \mathfrak{q}), \quad (176)$$

where I_{p+q} and I_r are the identity matrices of the sizes $(p+q) \times (p+q)$ and $r \times r$ respectively, and $M_{i,j} \in \text{Mat}_{r,p+q}(\mathbb{F})$ has at most one non-zero element m_{ij} , which is in the i -th row and j -th column, for fixed $i = 1, \dots, r$, $j = 1, \dots, p+q$. Then for the following multivector

$$U_{i,j} := e + c_{ij}e_j e_{p+q+i} \in (\mathcal{C}\ell^0 \oplus \mathcal{C}\ell_{p,q,r}^2)^\times, \quad c_{ij} := -\frac{m_{ij}}{2\eta_{jj}} \in \mathbb{F}, \quad e_j \in \mathcal{C}\ell_{p,q}^1, \quad e_{p+q+i} \in \Lambda_r^1,$$

we have

$$B_{i,j} = \tilde{\text{ad}}_{U_{i,j}}^1. \quad (177)$$

In the cases $r = n$ or $r = 0$, for the identity matrices I_r or I_{p+q} respectively, we have $I_r = \tilde{\text{ad}}_e^1$ and $I_{p+q} = \tilde{\text{ad}}_e^1$ respectively.

Proof. In the cases $r = n$ or $r = 0$, we have $\tilde{\text{ad}}_e^1(v) = eve^{-1} = v$ for any $v \in \mathcal{C}\ell_{p,q,r}^1$ and get the statement.

Further in the proof, we consider the case $r \neq 0$ and $r \neq n$. Without loss of generality, let us prove the statement for $B_{1,1}$.

Since $B_{1,1} \in \text{O}_{\Lambda_r^1}(V, \mathfrak{q})$ (163), it leaves invariant Λ_r^1 . Thus, $B_{1,1}v = \text{id}v$ for any $v \in \Lambda_r^1$. On the other hand, for $U_{1,1} := e + c_{11}e_1 e_{p+q+1}$, we have $\tilde{\text{ad}}_{U_{1,1}}^1(v) = U_{1,1}vU_{1,1}^{-1} = vU_{1,1}U_{1,1}^{-1} = v$, for any $c_{ij} \in \mathbb{F}$, where we apply the statement (84) of Lemma D.3, since $U_{1,1} \in \mathcal{C}\ell_{p,q,r}^{(0)}$ and $v \in \Lambda_r^1$.

Let us consider how $B_{1,1}$ acts on the vectors from a canonical basis of $\mathcal{C}\ell_{p,q}^1$.

For the vector $v_1 = (1, 0, \dots, 0) \in V$ corresponding to e_1 , we have

$$B_{1,1}v_1 = (1, 0, \dots, 0, m_{11}, 0, \dots, 0), \quad (178)$$

where m_{11} is on the $(p+q+1)$ -th position. On the other hand, for $U_{1,1}$ and e_1 , we get

$$\tilde{\text{ad}}_{U_{1,1}}^1(e_1) = (e_1 - c_{11}\eta_{11}e_{p+q+1})(e - c_{11}e_1 e_{p+q+1}) = e_1 - 2\eta_{11}c_{11}e_{p+q+1} = e_1 + m_{11}e_{p+q+1}, \quad (179)$$

which corresponds to (178).

For any other vector $v_k \in V$, $k = 2, \dots, p+q$, corresponding to e_k , we have

$$B_{1,1}v_k = (0, \dots, 0, 1, 0, \dots, 0), \quad (180)$$

where 1 is on the k -th position. On the other hand, for $U_{1,1}$ and e_k , we get

$$\tilde{\text{ad}}_{U_{1,1}}^1(e_k) = (e + c_{11}e_1 e_{p+q+1})e_k(e - c_{11}e_1 e_{p+q+1}) = e_k, \quad (181)$$

which corresponds to (180). We have proved that $B_{1,1}v = \tilde{\text{ad}}_{U_{1,1}}^1(v)$ for any v from a canonical basis of V . By linearity, we get the statement (177). This completes the proof. $\square$

Lemma H.8. *We have*

$$\widehat{X}V = VX, \quad \forall X \in (C\ell_{p,q,r}^{(0)\times} \cup C\ell_{p,q,r}^{(1)\times})\Lambda_r^\times, \quad \forall V \in \Lambda_r^{1\times}. \quad (182)$$

Proof. Suppose $V \in \Lambda_r^{1\times}$ and $X = WH$, where $W \in (C\ell_{p,q,r}^{(0)\times} \cup C\ell_{p,q,r}^{(1)\times})$ and $H \in \Lambda_r^\times$.

Note that $\widehat{H}V = VH$ by the statement (109) of Lemma E.1.

If $W \in C\ell_{p,q,r}^{(0)\times}$, then we have $\widehat{W}V = WV = VW$ by the statement (84) of Lemma D.3.

Consider the case $W \in C\ell_{p,q,r}^{(1)\times}$. If $r = n$, then $W \in \Lambda_r^{(1)\times}$, and we get $\widehat{W}V = VW$ again by the statement (109) of Lemma E.1. If $r \neq n$, then there exists such generator e_i that $(e_i)^2 \neq 0$, and we can always represent W as $W = e_i(\eta_{ii}e_iW)$. Since $\eta_{ii}e_iW \in C\ell_{p,q,r}^{(0)}$, we again get $(\widehat{\eta_{ii}e_iW})V = V(\eta_{ii}e_iW)$ by the statement (84) of Lemma D.3. Also we have $\widehat{e_i}V = -e_iV = Ve_i$, since $V \in \Lambda_r^1$ does not contain e_i . So, in this case, we get $\widehat{W}V = \widehat{e_i}(\widehat{\eta_{ii}e_iW})V = \widehat{e_i}V\eta_{ii}e_iW = Ve_i\eta_{ii}e_iW = VW$ as well.

Using the notes above, we finally get

$$\widehat{X}V = \widehat{W}\widehat{H}V = \widehat{W}VH = VWH = VX, \quad (183)$$

and the statement is proved. $\square$

Theorem H.9 (Image of $\tilde{\text{ad}}^1$). *The image of the restricted twisted adjoint representation $\tilde{\text{ad}}^1$ acting on the Lipschitz group $\tilde{\text{ad}}^1 : \tilde{\Gamma}_{p,q,r}^1 \rightarrow \text{Aut}(C\ell_{p,q,r})$ has the form:*

$$\text{im}(\tilde{\text{ad}}^1 : \tilde{\Gamma}_{p,q,r}^1 \rightarrow \text{Aut}(C\ell_{p,q,r})) = O_{\Lambda_r^1}(V, \mathfrak{q}). \quad (184)$$

Proof. Firstly, we prove that $\text{im}(\tilde{\text{ad}}^1 : \tilde{\Gamma}_{p,q,r}^1 \rightarrow \text{Aut}(C\ell_{p,q,r})) \subseteq O_{\Lambda_r^1}(V, \mathfrak{q})$, i.e. that $\tilde{\text{ad}}_T^1 \in O_{\Lambda_r^1}(V, \mathfrak{q})$ for any $T \in \tilde{\Gamma}_{p,q,r}^1$. Suppose $T \in \tilde{\Gamma}_{p,q,r}^1$, then $\tilde{\text{ad}}_T^1$ is linear by Lemma 2.1. Moreover, $\tilde{\text{ad}}_T^1$ is invertible with the inverse $\tilde{\text{ad}}_{T^{-1}}^1$, since

$$\tilde{\text{ad}}_{T^{-1}}^1(\tilde{\text{ad}}_T^1(v)) = \widehat{T^{-1}}(\widehat{T}vT^{-1})T = v, \quad \forall v \in C\ell_{p,q,r}^1. \quad (185)$$

Now, we need to show that

$$\mathfrak{q}(\tilde{\text{ad}}_T^1(v)) = \mathfrak{q}(v), \quad \forall v \in C\ell_{p,q,r}^1. \quad (186)$$

We have

$$\mathfrak{q}(\tilde{\text{ad}}_T^1(v)) = \mathfrak{q}(\tilde{\text{ad}}_T(v)) = \tilde{\text{ad}}_T(v)\tilde{\text{ad}}_T(v) = \tilde{\text{ad}}_T(v^2), \quad (187)$$

where we apply Remark H.3, since $\tilde{\text{ad}}_T(v) \in V$, and multiplicativity of $\tilde{\text{ad}}_T$ (Lemma 2.1), since $T \in \tilde{\Gamma}_{p,q,r}^1 \subseteq (C\ell_{p,q,r}^{(0)\times} \cup C\ell_{p,q,r}^{(1)\times})\Lambda_r^\times$ by Theorem 3.3. Since $v^2 = \mathfrak{q}(v)$ again by Remark H.3, we finally get

$$\mathfrak{q}(\tilde{\text{ad}}_T^1(v)) = \tilde{\text{ad}}_T(\mathfrak{q}(v)) = \mathfrak{q}(v), \quad (188)$$

where in the last equality we use the property that $\tilde{\text{ad}}_T$ leaves invariant the scalars by the statement (8) of Lemma 2.1. The statement (186) is proved, i.e. we have shown that $\tilde{\text{ad}}_T^1 \in O_{\Lambda_r^1}(V, \mathfrak{q})$. To prove $\tilde{\text{ad}}_T^1 \in O_{\Lambda_r^1}(V, \mathfrak{q})$, we finally need to show that $\tilde{\text{ad}}_T^1|_{\Lambda_r^1} = \text{id}_{\Lambda_r^1}$. This statement is true because for any $v \in \Lambda_r^1$ and $T \in \tilde{\Gamma}_{p,q,r}^1 \subseteq (C\ell_{p,q,r}^{(0)\times} \cup C\ell_{p,q,r}^{(1)\times})\Lambda_r^\times$ (Theorem 3.3), we have $\tilde{\text{ad}}_T^1(v) = \widehat{T}vT^{-1} = vTT^{-1} = v$ by Lemma H.8.

Now let us prove $O_{\Lambda_r^1}(V, \mathfrak{q}) \subseteq \text{im}(\tilde{\text{ad}}^1 : \tilde{\Gamma}_{p,q,r}^1 \rightarrow \text{Aut}(C\ell_{p,q,r}))$, i.e. let us prove the surjectivity of the mapping $\tilde{\text{ad}}^1$. We have by Lemma H.1:

$$O_{\Lambda_r^1}(V, \mathfrak{q}) \cong \left\{ \begin{pmatrix} A & 0 \\ M & I_r \end{pmatrix} \right\}, \quad A \in O(C\ell_{p,q}^1, \mathfrak{q}|_{C\ell_{p,q}^1}), \quad M \in \text{Mat}_{r \times (p+q)}(\mathbb{F}). \quad (189)$$

Suppose $\Phi \in O_{\Lambda_r^1}(V, \mathfrak{q})$. Then, there exist $A \in O(C\ell_{p,q}^1, \mathfrak{q}|_{C\ell_{p,q}^1})$ and $M \in \text{Mat}_{r \times (p+q)}(\mathbb{F})$ such that

$$\Phi = \begin{pmatrix} A & 0 \\ M & I_r \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & I_r \end{pmatrix} \begin{pmatrix} I_{p+q} & 0 \\ M & I_r \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & I_r \end{pmatrix} \cdots \begin{pmatrix} A_k & 0 \\ 0 & I_r \end{pmatrix} \begin{pmatrix} I_{p+q} & 0 \\ M_{1,1} & I_r \end{pmatrix} \cdots \begin{pmatrix} I_{p+q} & 0 \\ M_{r,p+q} & I_r \end{pmatrix}, \quad (190)$$

where $A = A_1 \cdots A_k$, $k \leq p+q$, and $A_1, \dots, A_k \in O(C\ell_{p,q}^1, \mathfrak{q}|_{C\ell_{p,q}^1})$ are matrices representing reflections, by the well-known Cartan–Dieudonné theorem (Dieudonné, 1971), which states that every orthogonal transformation in a $(p+q)$ -dimensional space with a non-degenerate symmetric bilinear form can be represented as a composition of at most $p+q$ reflections; and $M_{i,j} \in \text{Mat}_{r \times (p+q)}(\mathbb{F})$, $i = 1, \dots, r$, $j = 1, \dots, p+q$ is a matrix with at most one non-zero element, which is on the i -th row and j -th column, such that $M = \sum_{i=1, \dots, r} \sum_{j=1, \dots, p+q} M_{i,j}$.

Note that each reflection matrix A_i , $i = 1, \dots, k$, can be associated with some vector $v_i \in C\ell_{p,q}^1$, so that $A_i = \tilde{\text{ad}}_{v_i}^1$, by Remark H.4. Any matrix

$$B_{i,j} = \begin{pmatrix} I_{p+q} & 0 \\ M_{i,j} & I_r \end{pmatrix} \quad (191)$$

can be associated with some multivector $\gamma_{i,j} = e + m_{i,j}e_j e_{p+q+i} \in C\ell^0 \oplus C\ell_{p,q,r}^2$, where $m_{i,j} \in \mathbb{F}$, $e_j \in C\ell_{p,q}^1$, and $e_{p+q+i} \in \Lambda_r^1$, so that $B_{i,j} = \tilde{\text{ad}}_{\gamma_{i,j}}^1$, by Lemma H.7. Consider the following multivector, which consists of $k+p+q$ factors:

$$T = v_1 \cdots v_k \gamma_{1,1} \cdots \gamma_{r,p+q}. \quad (192)$$

Note that $T \in \tilde{\Gamma}_{p,q,r}^1$, since $v_1, \dots, v_k \in \tilde{\Gamma}_{p,q,r}^1$ by Remark H.5 and $\gamma_{1,1}, \dots, \gamma_{r,p+q} \in \tilde{\Gamma}_{p,q,r}^1$ by Remark H.6. We obtain for any $v \in V = C\ell_{p,q,r}^1$:

$$\Phi(v) = \begin{pmatrix} A_1 & 0 \\ 0 & I_r \end{pmatrix} \cdots \begin{pmatrix} A_k & 0 \\ 0 & I_r \end{pmatrix} \begin{pmatrix} I_{p+q} & 0 \\ M_{1,1} & I_r \end{pmatrix} \cdots \begin{pmatrix} I_{p+q} & 0 \\ M_{r,p+q} & I_r \end{pmatrix} v \quad (193)$$

$$= \tilde{\text{ad}}_{v_1}^1 \left(\tilde{\text{ad}}_{v_k}^1 \left(\tilde{\text{ad}}_{\gamma_{1,1}}^1 \left(\cdots \left(\tilde{\text{ad}}_{\gamma_{r,p+q}}^1 (v) \right) \right) \right) \right) \quad (194)$$

$$= \widehat{v_1} \cdots \widehat{v_k} (\widehat{\gamma_{1,1}} \cdots \widehat{\gamma_{r,p+q}} v \gamma_{r,p+q}^{-1} \cdots \gamma_{1,1}^{-1} v_k^{-1}) \cdots v_1^{-1}, \quad (195)$$

$$= (v_1 \cdots v_k \widehat{\gamma_{1,1} \cdots \gamma_{r,p+q}} v (v_1 \cdots v_k \gamma_{1,1} \cdots \gamma_{r,p+q})^{-1})^{-1} \quad (196)$$

$$= \tilde{\text{ad}}_{v_1 \cdots v_k \gamma_{1,1} \cdots \gamma_{r,p+q}}^1 (v) \quad (197)$$

$$= \tilde{\text{ad}}_T^1 (v). \quad (198)$$

Thus, for any $\Phi \in O_{\Lambda_r^1}(V, \mathfrak{q})$, there exists $T \in \tilde{\Gamma}_{p,q,r}^1$ such that $\Phi = \tilde{\text{ad}}_T^1$, and this completes the proof. $\square$

Theorem H.10. *In the case of any degenerate or non-degenerate $C\ell_{p,q,r}$, the mapping $\tilde{\text{ad}}^1$ defines the following isomorphism:*

$$\tilde{\text{ad}}^1 : \tilde{\Gamma}_{p,q,r}^1 / \Lambda_r^\times \cong O_{\Lambda_r^1}(V, \mathfrak{q}), \quad (199)$$

Proof. We have the homomorphism $\tilde{\text{ad}}^1$ of the groups $\tilde{\Gamma}_{p,q,r}^1$ and $O_{\Lambda_r^1}(V, \mathfrak{q})$, which is surjective by the statement (184) of Theorem H.9. By the fundamental homomorphism theorem, the group $O_{\Lambda_r^1}(V, \mathfrak{q})$ is isomorphic to the quotient group $\tilde{\Gamma}_{p,q,r}^1 / \ker(\tilde{\text{ad}}^1)$. By applying Theorem H.2, we complete the proof. $\square$

Note that in the particular case of the non-degenerate geometric algebra $C\ell_{p,q}$, Theorem H.10 has the form

$$\tilde{\text{ad}}^1 : \tilde{\Gamma}_{p,q}^1 / C\ell^{0 \times} \cong O(V, \mathfrak{q}) \quad (200)$$

and is well-known (see, for example, Benn & Tucker, 1987).

Theorem H.11 (Theorem 3.4). *If a mapping $f : \mathcal{C}\ell_{p,q,r} \rightarrow \mathcal{C}\ell_{p,q,r}$ is equivariant with respect to any group H that contains the Lipschitz group $\tilde{\Gamma}_{p,q,r}^1$ as a subgroup, then f is equivariant with respect to the corresponding restricted orthogonal group. In other words, if*

$$\tilde{\text{ad}}_T(f(x)) = f(\tilde{\text{ad}}_T(x)), \quad \forall T \in H, \quad \forall x \in \mathcal{C}\ell_{p,q,r},$$

then

$$f(\Phi(x)) = \Phi(f(x)), \quad \forall \Phi \in O_{\Lambda_r^1}(V, \mathfrak{q}), \quad \forall x \in \mathcal{C}\ell_{p,q,r},$$

where Φ acts on x in a sense (34) by applying an orthogonal transformation to its vector components.

Proof. Suppose H is a group, $\tilde{\Gamma}_{p,q,r}^1 \subseteq H$, and $\tilde{\text{ad}}_T(f(x)) = f(\tilde{\text{ad}}_T(x))$ for some mapping f , for any $T \in H$, and $x \in \mathcal{C}\ell_{p,q,r}$. Then it holds, in particular, for any $T \in \tilde{\Gamma}_{p,q,r}^1$, and f is $\tilde{\Gamma}_{p,q,r}^1$ -equivariant. By Theorem H.10, a mapping is equivariant w.r.t. $\tilde{\Gamma}_{p,q,r}^1$ iff it is equivariant w.r.t. $O_{\Lambda_r^1}(V, \mathfrak{q})$. Therefore, f is $O(V, \mathfrak{q})_{\Lambda_r^1}$ -equivariant, and the statement is proved. $\square$

I. Experimental Details

The implementation of GLGENN and summary of each experiment are available at <https://github.com/katyafilimoshina/glgenn>.

For CGENN (Ruhe et al., 2023), we follow the training setups from the corresponding public code release. For other models, we use the loss values from the corresponding code repository (Finzi et al., 2021).

We construct our models to closely resemble the CGENN architecture, replacing the $\mathcal{C}\ell_{p,q}^k$ -linear, $\mathcal{C}\ell_{p,q}^k$ -geometric product, and $\mathcal{C}\ell_{p,q}^k$ -normalization layers from Ruhe et al., 2023 with the same number of GLGENN’s $\mathcal{C}\ell_{p,q}^{\bar{k}}$ counterparts (see Section 4). This design choice results in GLGENN having significantly fewer parameters than CGENN, as they operate in a unified manner across 4 fundamental subspaces of geometric algebras defined by the grade involution ($\hat{\cdot}$) and reversion ($\tilde{\cdot}$); they processes geometric objects in groups with a step size of 4.

However, this parameter efficiency difference between GLGENN and CGENN becomes apparent only for tasks with $n > 3$. In lower-dimensional cases ($n \leq 3$), subspaces of fixed grades coincide with the subspaces determined by grade involution and reversion (4), i.e., $\mathcal{C}\ell_{p,q,r}^k = \mathcal{C}\ell_{p,q,r}^{\bar{k}}$ for $k = 0, 1, 2, 3$. Nonetheless, as experiments show, even in 5-dimensional cases, the performance gap between CGENN and GLGENN is substantial. Moreover, GLGENN’s parameter efficiency advantage increases with dimension n , as the subspaces $\mathcal{C}\ell_{p,q,r}^{\bar{k}}$ diverge further from $\mathcal{C}\ell_{p,q,r}^k$.

O(5, 0)-Regression Task

In our first experiment, we consider an O(5, 0)-invariant regression task proposed in (Finzi et al., 2021). The task is to estimate the function $\sin(\|x_1\|) - \|x_2\|^3/2 + \frac{x_1^T x_2}{\|x_1\|\|x_2\|}$, where $x_1, x_2 \in \mathbb{R}^{5,0}$ are vectors sampled from a standard Gaussian distribution. The loss function used is Mean squared error (MSE).

We evaluate the performance using four different training dataset sizes and compare against the ordinary MLP, MLP with augmentations, the O(5, 0)- and SO(5, 0)-equivariant MLP architectures proposed in (Finzi et al., 2021), and with CGENN (Ruhe et al., 2023). CGENN, MLP, MLP with augmentations, O(5, 0)- and SO(5, 0)-equivariant MLP architectures have approximately the same number of parameters, we use the setup from the similar experiment provided in Ruhe et al., 2023 (they state that they compare their model CGENN with other models with the same number of trainable parameters).

The number of parameters in Table 5 (O(5, 0)-Regression Experiment) is as follows. All the models besides GLGENN and CGENN have $\approx 150.3\text{K}$ parameters in total. In CGENN and GLGENN, the architecture contains sequentially applied geometric algebra-based layers (applied to the subspaces of all grades) and ordinary MLP layers (applied only to the subspace of 0-grade (scalars)). The most significant layers and at the same time the most consuming for training are geometric algebra-based ones, and CGENN possess $\approx 1.8\text{K}$ of parameters associated with such layers, while GLGENN has $\approx 0.6\text{K}$ of such parameters. For the MLP part, which is very fast and easy for training, both CGENN and GLGENN both have $\approx 148.5\text{K}$ parameters.

Table 5. MSE ($\downarrow$) on the $O(5, 0)$ -Regression Experiment.

MODEL	# OF TRAINING SAMPLES			
	$3 \cdot 10^1$	$3 \cdot 10^2$	$3 \cdot 10^3$	$3 \cdot 10^4$
GLGENN	0.1055	0.0020	0.0031	0.0011
MLP	28.1011	0.2482	0.0623	0.0622
MLP+AUG	0.4758	0.0936	0.0889	0.0672
EMLP- $O(5)$	0.152	0.0344	0.0310	0.0273
EMLP-SO(5)	0.1102	0.0384	0.032	0.0279
CGENN	0.0791	0.0089	0.0012	0.0003

Table 6. MSE ($\downarrow$) on the $O(5, 0)$ -Regression Experiment in the case of large training set size.

MODEL	# OF TRAINING SAMPLES	
	$6 \cdot 10^4$	$1 \cdot 10^5$
GLGENN	0.0001	0.0001
CGENN	0.0002	0.0002

The results, presented in Tables 5, 6 and Figure 2 (left), demonstrate that GLGENN achieves performance on a par with CGENN, while significantly outperforming the other models. Moreover, GLGENN attains these results with fewer parameters and reduced training time compared to CGENN. In Table 5, MSE for CGENN and GLGENN are averaged over 5 runs. MSE for other models are averaged over 3 runs (from Finzi et al., 2021). Number of iterations is the same for all algorithms.

$O(5, 0)$ -Convex Hull Volume Estimation

In this equivariant experiment, the task is to estimate the volume of a convex hull generated by K points in $\mathbb{R}^{5,0}$. We consider three different values of K : 16 (as used in (Ruhe et al., 2023; Liu et al., 2024)) and two additional settings, $K = 256$ and 512, which are more relevant for real-world applications.

In the case $K = 16$, we consider 4 different sizes of the training set: 256, 1024, 4096, and 16384 samples. We compare GLGENN with the state-of-the-art model CGENN (Ruhe et al., 2023). While both models share a similar architecture, GLGENN has 24.1K trainable parameters, whereas CGENN has 58.8K. The number of training steps is the same for both models. The results are presented in Table 3 and Figure 2 (middle). GLGENN either outperforms or matches CGENN across all training set sizes. The most noticeable difference occurs with smaller training sets. We attribute this to GLGENN’s lower parameter count, which reduces its tendency to overfit—an issue commonly observed in small datasets.

To study the behavior of CGENN and GLGENN on different iterations of training, we provide Figure 4. This figure shows the training and test loss at different iterations. Notably, CGENN tends to achieve lower training loss compared to GLGENN. However, at a certain point, while CGENN’s training loss continues to decrease, its test loss plateaus and remains almost constant. In contrast, GLGENN’s training loss decreases at a slower rate, but its test loss is consistently lower than that of CGENN, indicating better generalization.

For $K = 256$ and $K = 512$, the results are also presented in Table 3. To ensure a fair comparison, we first select CGENN architectures that perform best for each setup. We then construct corresponding GLGENN models by replacing CGENN layers with their GLGENN counterparts, resulting in significantly fewer trainable parameters. Note that GLGENN consistently outperform CGENN in these real-world settings.

There is a contrast in GLGENN and CGENN generalization behavior in cases of large K . We illustrate it in Figure 5, which shows the training and test losses for $K = 256$ in the case of 1024 (left) and 16384 (right) training set sizes. GLGENN demonstrate stable convergence without signs of overfitting: the training and test losses decrease in a similar way throughout the optimization trajectory. CGENN show a clear overfitting pattern: while the training loss quickly drops to near-zero, the test loss plateaus early and remains significantly higher, especially in the small-data regime.

Following the recommendation of one of the anonymous reviewers, in Table 7, we report the average wall-clock time required to process one training batch for GLGENN and CGENN across different training set sizes and number of points K .

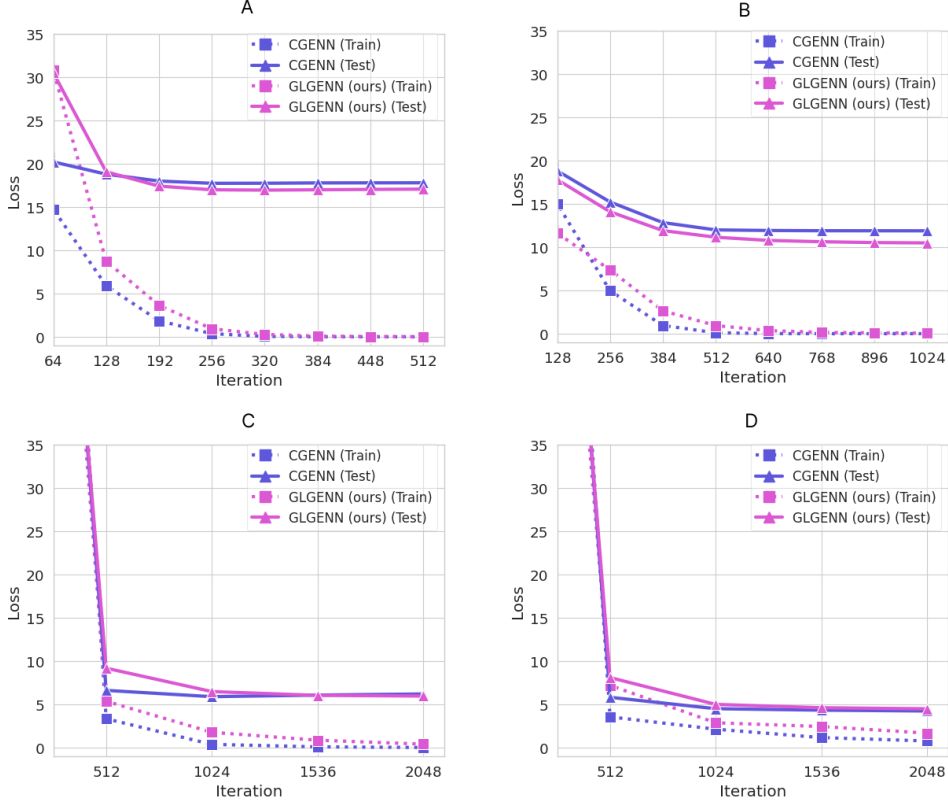


Figure 4. $O(5, 0)$ -Convex Hull, $K = 16$. The plots illustrate the training and test loss curves for CGENN and GLGENN across different training iterations. Subfigures (A)–(D) correspond to different training set sizes: 256 (A), 1024 (B), 4096 (C), and 16384 (D), respectively.

The results demonstrate that GLGENN consistently achieve faster training times compared to CGENN. The performance gap remains notable across training dataset scales. The majority of memory usage is attributed to the Python environment and PyTorch’s internal components rather than to the models themselves; as a result, the memory footprint is comparable between GLGENN and CGENN.

Table 7. Average time (in seconds) for processing one training batch in $O(5, 0)$ -Convex Hull Experiment. In case of the number of points $K = 16$ and 256, for datasets with 2^8 , 2^{10} , and 2^{12} training samples, the batch size is 128; for 2^{14} samples, the batch size is 256. In case of $K = 512$, for datasets with 2^{10} and 2^{12} samples, batch sizes are 256 and 512 respectively.

K	16				256		512	
	# TRAIN SAMPLES				# TRAIN SAMPLES		# TRAIN SAMPLES	
MODEL	2^8	2^{10}	2^{12}	2^{14}	2^{10}	2^{14}	2^{10}	2^{14}
GLGENN	0.22352	0.16708	0.1712	0.3408	3.9931	12.8758	13.3317	27.4383
CGENN	0.33692	0.2331	0.23418	0.43854	8.658	17.7424	17.6603	35.5605
GAP	-0.1134	-0.06602	-0.06298	-0.09774	-4.6649	-4.8666	-4.3286	-8.1222

$O(7, 0)$ -Convex Hull Volume Estimation

In this experiment, we extend the previous task with number of points $K = 16$ to a higher-dimensional setting and evaluate the performance of GLGENN and CGENN in estimating the volume of a convex hull formed by 16 points in $\mathbb{R}^{7,0}$. We consider three different training set sizes: 256, 512, and 1024 samples.

While both models share a similar architecture, GLGENN maintains 24.1K trainable parameters, as in the previous experiment, whereas CGENN’s parameter count increases to 83.7K. The number of training steps remains the same for both models.

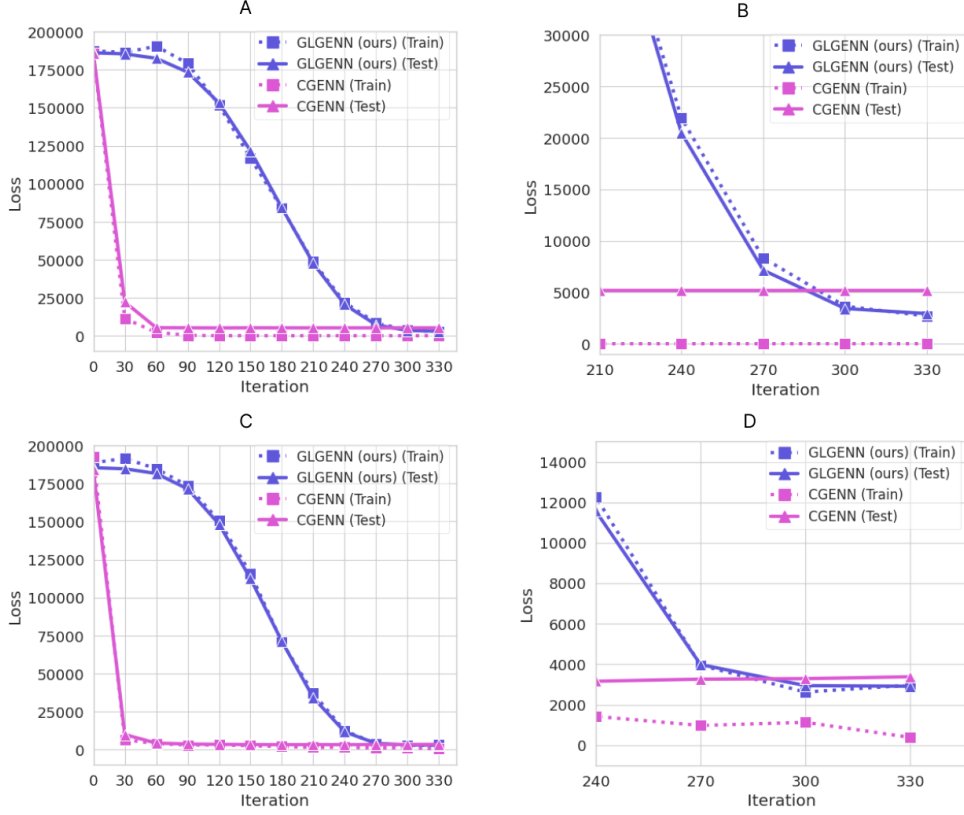


Figure 5. $O(5, 0)$ -Convex Hull, $K = 256$. The plots illustrate the training and test loss curves for CGENN and GLGENN across different training iterations. Subfigures (A)–(B) correspond to training with 1024 samples; subfigures (C)–(D) correspond to training with 16384 samples. The right-hand plots zoom in on the final iterations of training.

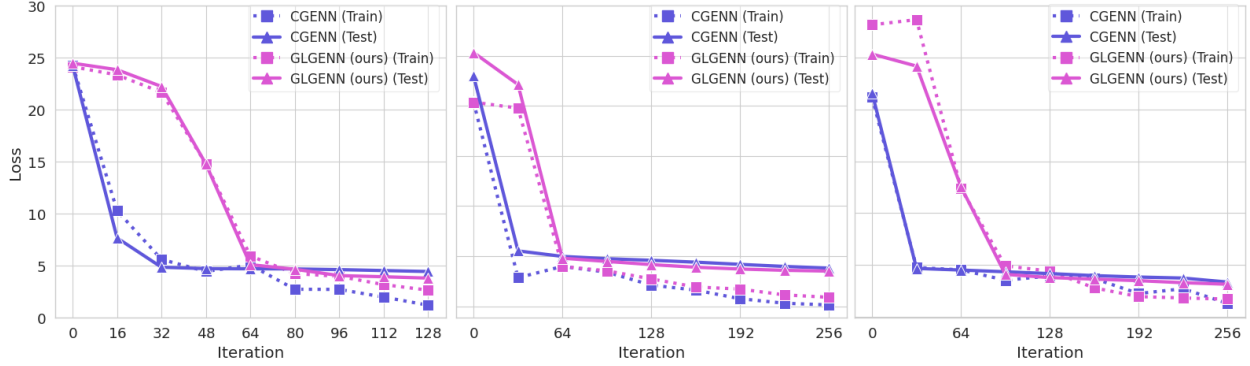


Figure 6. $O(7, 0)$ Convex Hull. The plots illustrate the training and test loss curves for CGENN and GLGENN across different training iterations. Subfigures correspond to different training set sizes: 256 (left), 512 (middle), 1024 (right), respectively.

The results, presented in Table 8, show that GLGENN consistently outperforms or matches CGENN. The smaller is the training set size, the higher is the difference.

Figure 6 compares the training and test loss for CGENN and GLGENN across different iterations. The observed behavior is consistent with the results from the $O(5, 0)$ -Convex Hull Experiment: CGENN tends to achieve lower training loss, while GLGENN exhibits better generalization, particularly for smaller training sets.

Table 8. MSE ($\downarrow$) on the $O(7, 0)$ -Convex Hull Experiment on the test dataset.

MODEL	# OF TRAINING SAMPLES			# OF PARA-METERS
	2^8	2^9	2^{10}	
GLGENN	4.0032	3.7378	3.5343	24.1K
CGENN	4.4914	3.7756	3.5408	83.7K

CGENN Same Size as GLGENN

In all our experiments, GLGENN architecture, by construction, has fewer trainable parameters than CGENN. This is because, for each task, we first construct the best-performing CGENN architecture and then obtain GLGENN by replacing CGENN layers with their corresponding GLGENN counterparts considered in Section 4, which have fewer parameters. Across all experiments, GLGENN either matches or outperforms CGENN, despite using fewer parameters.

In this subsection, we explore another setting: we compare GLGENN with CGENN architectures constrained to have the same number of parameters as the best-performing GLGENN. Our results show that, in almost all cases, CGENN performs worse under this constraint.

Table 9 presents results from the $O(5, 0)$ -Convex Hull Volume Estimation Experiment. In this setup, the CGENN model is downsized to have approximately 25K parameters (compared to 58.8K in its best-performing version), while GLGENN has 24.1K parameters. As shown, GLGENN outperforms the size-matched CGENN across all training set sizes.

Table 9. MSE ($\downarrow$) on the $O(5, 0)$ -Convex Hull Experiment ($K = 16$ Points) on the test dataset. Comparison of CGENN of different sizes with GLGENN.

MODEL	# OF PARA-METERS	# OF TRAINING SAMPLES			
		2^8	2^{10}	2^{12}	2^{14}
GLGENN	24.1K	16.94	10.40	6.2	4.46
CGENN (SAME SIZE AS GLGENN)	25K	19.79	15.94	7.69	4.23
CGENN (BEST PERFORMING)	58.8K	18.71	11.93	6.1	4.11

Table 10 reports results for the $O(5, 0)$ -Regression Task, where both GLGENN and CGENN are constrained to have approximately 149.1K total parameters, including about 0.6K parameters in geometric algebra-based layers. Again, GLGENN achieves superior or comparable performance.

Table 10. MSE ($\downarrow$) on the $O(5, 0)$ -Regression Task on the Test Dataset. Comparison of CGENN of different sizes with GLGENN. The second column shows the number of trainable parameters associated with geometric algebra-based layers.

MODEL	# OF GA-PARA-METERS	# OF TRAINING SAMPLES			
		$3 \cdot 10^1$	$3 \cdot 10^2$	$3 \cdot 10^3$	$3 \cdot 10^4$
GLGENN	0.6K	0.1055	0.0020	0.0031	0.0011
CGENN (SAME SIZE AS GLGENN)	0.6K	0.2856	0.0076	0.0017	0.0005
CGENN (BEST PERFORMING)	1.8K	0.0791	0.0089	0.0012	0.0003

Application of Typical Activation Functions

It is possible to combine geometric algebra-based layers with standard neural network layers, such as MLPs, while preserving equivariance — provided that the non-geometric algebra-based layers are applied only to scalars (i.e., elements of the subspace $C\ell^0$). The diversity of layers in this combination may lead to better results, although in equivariant tasks, non-geometric algebra-based layers on their own generally underperform compared to GLGENN layers or other geometric algebra-based layers. Applying non-geometric layers to $C\ell^0$ has inherent limitations compared to geometric algebra-based nonlinearities: they do not mix grades and instead isolate them, thereby hindering interactions across subspaces.

We provide an example in the case of the $O(5, 0)$ -Regression Task with 300 training samples. The results presented in Table 11 and Figure 3 show that the best performance is achieved by a combination of GLGENN (applied to all grades) and MLP (applied to scalars).

Table 11. MSE ($\downarrow$) on combination of MLP with GLGENN and CGENN in $O(5, 0)$ -Regression.

MODEL	MSE ON TEST
MLP	0.1706
GLGENN w/ MLP	0.0010
CGENN w/ MLP	0.0066
GLGENN w/o MLP	0.0265
CGENN w/o MLP	0.0723

$O(5, 0)$ - N -Body Experiment

We consider a system of $N = 5$ charged particles (bodies) in $\mathbb{R}^{5,0}$ with given masses, initial positions, and velocities. The task is to predict the final positions of the bodies after the system evolves under Newtonian gravity for 1000 Euler integration steps. For this purpose, we embed all the data in the geometric algebra $\mathcal{C}\ell_{5,0}$.

We construct a graph neural network (GNN) based on the message-passing paradigm (Gilmer et al., 2017), where bodies are treated as nodes in a graph, and their pairwise interactions are modeled as edges. The message and update networks are equivariant GLGENN. We use $\mathcal{C}\ell_{5,0}^k$ -linear, $\mathcal{C}\ell_{5,0}^k$ -normalization, and $\mathcal{C}\ell_{5,0}^k$ -geometric product GLGENN layers combined with MVSILU layer from CGENN (Ruhe et al., 2023). We compare against CGENN, which itself outperforms several state-of-the-art methods, including steerable SE(3)-Transformers (Fuchs et al., 2020), Tensor Field Networks (Thomas et al., 2018), SEGNN (Brandstetter et al., 2022), Radial Field (Köhler et al., 2020), EGNN (Satorras et al., 2021), and NMP (Gilmer et al., 2017). To ensure a fair comparison, we use the best-performing CGENN architecture and then replace its layers with analogous GLGENN counterparts to obtain the GLGENN architecture, which automatically has two times fewer trainable parameters. The results are presented in Table 12 and Figure 2 (right).

Table 12. MSE ($\downarrow$) on the $O(5, 0)$ - N -Body Experiment.

MODEL	# OF TRAINING SAMPLES			# OF PARAMETERS
	$3 \cdot 10^1$	$3 \cdot 10^2$	$3 \cdot 10^3$	
GLGENN	0.007	0.0011	0.0009	103K
CGENN	0.0136	0.0015	0.0007	210K