

ADS-GNN - A CONFORMALLY EQUIVARIANT GRAPH NEURAL NETWORK

Maksim Zhdanov*
 AMLab, University of Amsterdam
 m.zhdanov@uva.nl

Nabil Iqbal*
 Department of Mathematical Sciences, Durham University
 AMLab, University of Amsterdam
 nabil.iqbal@durham.ac.uk

Erik Bekkers
 AMLab, University of Amsterdam
 e.j.bekkers@uva.nl

Patrick Forré
 AMLab, AI4Science Lab, University of Amsterdam
 p.d.forre@uva.nl

ABSTRACT

Conformal symmetries, i.e. coordinate transformations that preserve angles, play a key role in many fields, including physics, mathematics, computer vision and (geometric) machine learning. Here we build a neural network that is equivariant under general conformal transformations. To achieve this, we lift data from flat Euclidean space to Anti de Sitter (AdS) space. This allows us to exploit a known correspondence between conformal transformations of flat space and isometric transformations on the AdS space. We then build upon the fact that such isometric transformations have been extensively studied on general geometries in the geometric deep learning literature. We employ message-passing layers conditioned on the proper distance, yielding a computationally efficient framework. We validate our model on tasks from computer vision and statistical physics, demonstrating strong performance, improved generalization capacities, and the ability to extract conformal data such as scaling dimensions from the trained network.

1 INTRODUCTION

The notion of *symmetry* is a key tool both in our understanding of nature and for the construction of machine learning systems that perceive nature. The construction of *equivariant* neural network architectures that encode specific symmetries has powerful advantages both conceptual and computational. In particular, much work has been dedicated to building networks that are equivariant under familiar symmetries such as rotations and translations.

In this work, we will study the symmetry group of *conformal transformations* i.e. the set of transformations on \mathbb{R}^d that preserve angles. This includes translations, rotations, reflections, scalings and so called special conformal transformations. Scale transformations – i.e. rigid rescalings of the whole system – are an extremely important subgroup of the conformal group. Conformal and scale invariance play a central role in many diverse fields. To give some examples: biological visual systems seem to exhibit insensitivity to scale (Logothetis et al., 1995; Han et al., 2020). Physical systems undergoing a second-order phase transition have fluctuations at all scales; they are generally conformally invariant at the critical point (Cardy, 1996), and can usefully be described by *conformal field theories*

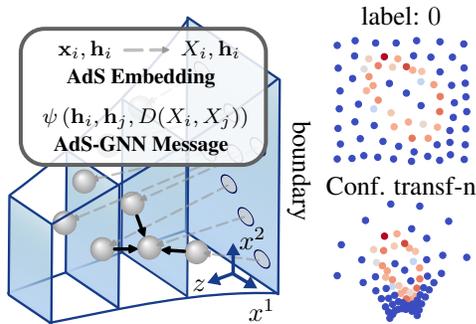


Figure 1: AdS-GNN lifts points from Euclidean space to Anti de Sitter space and computes message passing conditioned on the proper distance.

*Equal contribution.

(Di Francesco et al., 1997). Indeed in physics most systems exhibiting scale invariance also exhibit conformal invariance “for free” (Polchinski, 1988; Nakayama, 2015). Diverse applications also exist in computational geometry and computer vision (see e.g. (Sharon & Mumford, 2006; Lei et al., 2023))

In this work, we construct a neural network that is equivariant under conformal transformations, motivated by applications to computer vision and to physical systems. Our approach acts on point clouds and lifts the data into an auxiliary higher dimensional space called Anti de Sitter (AdS) space. As outlined below, this approach is inspired by ideas in conformal field theory in theoretical physics. We validate our construction on some tasks drawn from computer vision and statistical physics.

2 PREVIOUS WORK

Construction of equivariant neural networks under general symmetry groups began with (Cohen & Welling, 2016) and has since evolved into a well-studied field (see, e.g., (Bronstein et al., 2021; Weiler et al., 2023) for reviews). Advancements include treatments of isometric transformations on general Riemannian manifolds (Weiler et al., 2023) and extensions to semi-Riemannian manifolds (Zhdanov et al., 2023; 2024). The significant benefits of incorporating equivariance, particularly for point cloud data and even at large computational scales, are increasingly demonstrated (Vadgama et al., 2025; Brehmer et al., 2024), highlighting advantages in performance and efficiency. Lie group theory has been employed to develop equivariant networks for broad classes of transformations, such as general affine transformations (Mironenco & Forré, 2024), or to leverage specific algebraic structures like adjoint actions on Lie algebras (Lin et al., 2024). However, despite progress in handling Euclidean symmetries, scale equivariance, or even these more general Lie groups, these methods do not address the full conformal group, which uniquely includes non-affine special conformal transformations.

The pursuit of scale equivariance, a critical component of conformal symmetry, is the closest related area to our work on full conformal equivariance, with various approaches developed from specialized convolutional architectures (Bekkers, 2020; Sosnovik et al., 2019) to techniques like Fourier layers for robust scale handling (Rahman & Yeh, 2023). Our use of AdS space connects conceptually to *scale space* theories (Witkin, 1987; Worrall & Welling, 2019), where the extra dimension corresponds to scale. However, a crucial distinction is that the geometry of AdS space enforces equivariance under the larger group of all conformal transformations by construction, not just scale or isometric transformations. This approach also finds resonance in physics with recent explorations of using neural networks to model aspects of conformal field theory (Halverson et al., 2024).

3 CONFORMAL SYMMETRY AND ANTI DE SITTER SPACE

3.1 CONFORMAL TRANSFORMATIONS

A detailed and self-contained account of conformal transformations is given in Appendix D; here we present a brief review. Formally, a *global conformal transformation* of the Euclidean space \mathbb{R}^d is an injective smooth map $\varphi : \mathbb{R}^d \setminus \{x_\varphi\} \rightarrow \mathbb{R}^d, x \mapsto \varphi(x)$, defined on \mathbb{R}^d except on a possible point x_φ ¹ such that local angles are preserved, i.e. for all $x \in \mathbb{R}^d \setminus \{x_\varphi\}$ and $v_1, v_2 \in \mathbb{R}^d \setminus \{0\}$ we require²:

$$\frac{\langle \varphi'(x)v_1, \varphi'(x)v_2 \rangle}{\|\varphi'(x)v_1\| \cdot \|\varphi'(x)v_2\|} =: \cos(\angle(\varphi'(x)v_1, \varphi'(x)v_2)) \stackrel{!}{=} \cos(\angle(v_1, v_2)) := \frac{\langle v_1, v_2 \rangle}{\|v_1\| \cdot \|v_2\|}, \quad (1)$$

where $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denote the standard Euclidean scalar product and norm, resp., and $\varphi'(x)$ the Jacobian matrix of φ at x . Note, that, in contrast to *isometric transformations*, we do not require that the distances/norms are preserved.

The *group of all global conformal transformations* of \mathbb{R}^d is denoted by $\text{Conf}_g(\mathbb{R}^d)$. We also define the (*restricted*) *conformal group* $\text{Conf}(\mathbb{R}^d)$ as the connected component of the identity of $\text{Conf}_g(\mathbb{R}^d)$.

¹By allowing φ to not be defined on a certain point $x_\varphi \in \mathbb{R}^d$, we effectively allow φ to map x_φ to the “points at infinity” ∞ in the conformal compactification \mathbb{S}^d of \mathbb{R}^d . In fact, every *global* conformal transformation of \mathbb{R}^d uniquely extends to an angle-preserving diffeomorphism of \mathbb{S}^d for $d \geq 2$, see (Schottenloher, 2008) Thm. 2.6-2.11, or D.4.2 and D.4.3. This is why we introduce the definition with x_φ in this way.

²An equivalent, but more general and abstract, definition is provided in Definition D.1.2 and Lemma D.1.1.

It is shown³ that for $d \geq 2$ the global conformal group $\text{Conf}_g(\mathbb{R}^d)$ is isomorphic to the *projective orthogonal group*:

$$\text{PO}(d+1, 1) := \text{O}(d+1, 1)/\{\pm 1\},$$

where the role of the quotient is described in Appendix D in Proposition D.4.1.

It is instructive to consider the action of the group on points in terms of separated parameters. In particular, a general element G of the group can be written in terms of parameters $(\lambda, t, b, M) \in \mathbb{R}_{>0} \times \mathbb{R}^d \times \mathbb{R}^d \times \text{O}(d)$ and acts on a point $x \in \mathbb{R}^d$ through the composition of maps:

$$x' = x + t \quad x' = Mx \quad x' = \lambda x \quad \frac{x'}{\|x'\|^2} = \frac{x}{\|x\|^2} - b, \quad (2)$$



translations



rotations



scalings



special conformal transformations

resulting in a transformed point $x' = Gx$. These $(d+1)(d+2)/2$ parameters then assemble into an element of $\text{O}(d+1, 1)$, see e.g. (Di Francesco et al., 1997) for a review.

Above we have explained how points $x \in \mathbb{R}^d$ transform under the global conformal group. In a typical application we will often be dealing with *conformal fields* $\phi(x)$ defined on this space. The transformation of these fields is governed by the representation theory of the global conformal group⁴, which is non-trivial and we do not review it here, except to state that a privileged role is played by a basis of fields called *conformal primaries*, which we denote by $\mathbb{O}(x)$. Associated with each primary field is a number $\Delta_{\mathbb{O}}$ called the *conformal dimension* which characterizes the field in question⁵. Conformal primaries are distinguished by the fact that they transform as simply as possible under the group, and in particular under scaling of the coordinate they transform by a multiplicative factor:

$$\mathbb{O}'(x') := \mathbb{O}'(\lambda x) = \lambda^{-\Delta_{\mathbb{O}}} \mathbb{O}(x) \quad (3)$$

Example. A simple example is the electric field in ordinary electrodynamics; the usual inverse square law near a point charge $\vec{E}(x) \sim \frac{\hat{x}}{|x|^2}$ is consistent with equation 3 if $\Delta_E = 2$. Another example is image data viewed as a scalar field $p(x)$ denoting the pixel value; under a conformal transformation only the argument of the field changes and not its value, and thus we would take $\Delta_{\text{pixel}} = 0$.

In physical applications the values of $\Delta_{\mathbb{O}}$ for a given conformally invariant system – e.g. a particular kind of phase transition – are often of great interest⁶, as they are pure numbers that are *universal*: they generally do not depend on the microscopic details of the system, only on the symmetries and the spatial dimensionality.

3.2 THE ANTI DE SITTER SPACE AdS^{d+1}

To organize the data, we will instead lift it from \mathbb{R}^d to an auxiliary space with *one higher dimension*, i.e. *Anti de Sitter space* AdS^{d+1} . This space can conveniently be understood in terms of a submanifold of $\mathbb{R}^{d+1,1}$, equipped with its natural metric

$$\eta = \text{diag}(-1, 1, 1, \dots, 1) \quad (4)$$

Consider the submanifold given by the constraint $\|Y\| = -1$, where $Y \in \mathbb{R}^{d+1,1}$. As shown in Figure 2, this submanifold has two connected components, each of which is defined to be a copy of

³see Theorem D.4.2, Appendix D, Proposition D.4.1 and (Schottenloher, 2008) Thm. 2.6-2.11 for details.

⁴Elementary reviews can be found in (Simmons-Duffin, 2017; Di Francesco et al., 1997).

⁵It may be helpful to consider fields transforming under the rotation group $\text{SO}(3)$: there are scalar fields, spinor fields, vector fields, etc. which are characterized by a discrete parameter called the *spin* s which takes discrete values $s \in \frac{\mathbb{Z}}{2}$, i.e. $s_{\text{scalar}} = 0, s_{\text{vector}} = 1$, etc. Δ can very loosely be thought of as the analogue of “spin” for scale transformations. The fact that it is now a continuous variable is related to the fact that the conformal group is non-compact.

⁶See e.g. the 1982 Nobel Prize lecture by (Wilson, 1983) for a historical overview, or (Cardy, 1996; Kardar, 2007; Zinn-Justin, 2021) for textbook treatments.

AdS^{d+1} .⁷ For concreteness, we will present all formulas working with the “top” branch, i.e. the one which has $Y^0 > 0$, and we define:

$$\text{AdS}^{d+1} := \{Y \in \mathbb{R}^{d+1,1} \mid \|Y\| = -1, Y^0 > 0\}. \quad (5)$$

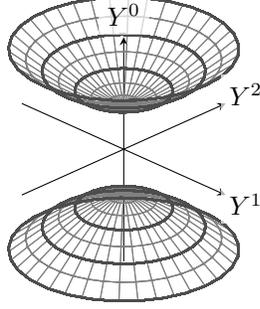


Figure 2: Example of hyperboloid $\|Y\| = -1$ embedded in $\mathbb{R}^{2,1}$, constituting two copies of AdS^2 .

There is a natural action of $\text{PO}(d+1, 1) = \text{O}(d+1, 1)/\{\pm 1\}$ on this submanifold, given by $\text{O}(d+1, 1)$ rotations on Y on each connected component. This forms the *isometry group* $\text{Isom}(\text{AdS}^{d+1})$ of AdS^{d+1} , see (McKay, 2023) Theorem 7.5 or D.2.6. As mentioned before, note that $\text{PO}(d+1, 1)$ is also precisely the global conformal group $\text{Conf}_g(\mathbb{R}^d)$ of \mathbb{R}^d . This is not an accident: the fact that local operations in the interior of AdS^{d+1} result in conformally invariant operations on \mathbb{R}^d is very well known in the context of the AdS/CFT correspondence in quantum gravity (Maldacena, 1999; Gubser et al., 1998; Witten, 1998), which states that under some circumstances a quantum gravity theory on AdS^{d+1} is equivalent to a quantum field theory with conformal symmetry on \mathbb{R}^d . Also, see Remark D.2.7.

Here we will not use any of the dynamics of quantum gravity, but we will exploit some of the well-studied kinematics of that correspondence as a convenient tool to build convolutional kernels on AdS^{d+1} . Indeed, a general framework for constructing convolutional layers that are equivariant under isometry groups of any pseudo-Riemannian manifold was given in (Weiler et al., 2023; Zhdanov et al., 2024), and our work can be viewed as building on a special case of that.

To be more precise, we first place explicit coordinates $X = (X^1, \dots, X^{d+1})$ on AdS^{d+1} . It will be convenient to separate these coordinates as $X = (X^1, \dots, X^{d+1}) = (x^1, \dots, x^d, z) = (x, z) \in \mathbb{R}^d \times \mathbb{R}_{>0}$ where $x \in \mathbb{R}^d$ and $z \in \mathbb{R}_{>0}$ is an extra dimension. We can now solve the hyperboloid constraint equation 5 in terms of these coordinates as

$$Y^0 := \frac{z^2 + \|x\|^2 + 1}{2z}, \quad Y^a := \frac{x^a}{z}, \quad Y^{d+1} := \frac{z^2 + \|x\|^2 - 1}{2z} \quad (6)$$

The Riemannian metric on AdS^{d+1} is the induced metric on this hyperboloid⁸.

$$ds^2 = \sum_{\mu, \nu=1}^{d+1} g_{\mu\nu}(X) dX^\mu dX^\nu = \frac{1}{z^2} \left(\sum_{a=1}^d (dx^a)^2 + dz^2 \right). \quad (7)$$

Finally, we record how the isometry group acts on AdS^{d+1} as $X' = GX$. Using the same parameters as in equation 2 we have:

$$(x', z') = (x + t, z) \quad (x', z') = (Mx, z) \quad (x', z') = (\lambda x, \lambda z), \quad (8)$$

$$\left(\frac{x'}{\|x'\|^2 + z'^2}, \frac{z'}{\|x'\|^2 + z'^2} \right) = \left(\frac{x}{\|x\|^2 + z^2} + b, \frac{z}{\|x\|^2 + z^2} \right). \quad (9)$$

A key point is that the manifold (7) has a d -dimensional boundary at $z = 0$. This boundary is mapped to itself under the isometries. Furthermore, the isometry group acts on the boundary points $(x^a, z = 0)$ precisely as in (2). Thus one should imagine that conformal data on \mathbb{R}^d “lives on the boundary of AdS^{d+1} ”. In what follows a key role will be played by the $\text{PO}(d+1, 1)$ -invariant proper distance between $D(X, X')$ between two points in AdS^{d+1} , which is related to the absolute inner product $|\langle Y, Y' \rangle|$ and is given explicitly by:

$$\cosh D(X, X') := |\langle Y, Y' \rangle| \stackrel{!}{=} \frac{z^2 + z'^2 + \sum_{a=1}^d (x^a - x'^a)^2}{2zz'}. \quad (10)$$

⁷For a more abstract and general definition of AdS for general quadratic spaces we refer to Definition D.2.3 and the corresponding (partial) parameterization presented here to Appendix D.6.

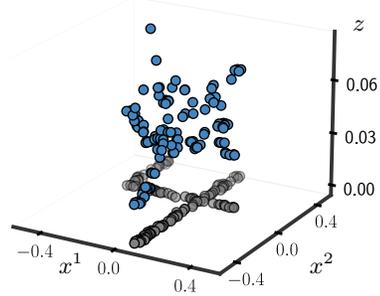
⁸For the details, see Proposition D.6.1.

4 ADS-GNN

We will now describe how to use these ideas to formulate a conformally equivariant neural network by extending the data from the boundary into the bulk of AdS^{d+1} .

Algorithm 1 AdS Embedding

Require: $X = \{\mathbf{x}_i\}_{i=1}^N \subset \mathbb{R}^d$, $k_{\text{lift}} \in \mathbb{N}$, $z_0 \in \mathbb{R}$
 1: **for** each point $i \in \{1, \dots, N\}$ **do**
 2: $z_i \leftarrow z_0$
 3: $\text{neighbors}_i \leftarrow \text{KNN}(x_i, X, k_{\text{lift}})$
 4: $(\hat{x}_i, \hat{z}_i) \leftarrow \text{ComputeAdSCoM}(\text{neighbors}_i)$
 5: **end for**
 6: **return** $\{(x_i, \hat{z}_i)\}_{i=1}^N \subset \text{AdS}^{d+1}$



4.1 EMBEDDING POINTS IN AdS

Consider a point cloud of N points in \mathbb{R}^d , $\{x_i\}$, $i \in \{1, \dots, N\}$. We would now like to lift this data into the bulk of AdS^{d+1} in a manner that preserves the symmetries.

A first attempt from the correspondence of symmetries shown in (8) is to simply embed each of the points directly into the boundary $z = 0$, i.e. with $X_i^\mu = (x_i^a, z = 0)$. However the metric (7) has a singularity at $z = 0$ – e.g. from (10) note that each point at $z = 0$ is at infinite proper AdS distance from any points at $z > 0$ – and thus such an attempt will require us to pick a regulating value of z . Using a fixed constant will explicitly break the symmetries.

We instead determine a value of z_i for each point using an approach outlined in Algorithm 1. We first embed each point into AdS using

$$X_i^\mu = (x_i^a, z = z_0) \quad (11)$$

with a small regulator z_0 . For each point, we then compute the *AdS center of mass* $\hat{X} = (\hat{x}_i, \hat{z}_i)$ of its k_{lift} nearest neighbours. The AdS center of mass is a generalization to hyperbolic space of the familiar notion of the center of mass from flat Euclidean space. It can be computed easily using an approach due to (Galperin, 1993) which we review in Appendix A.2.

The geometry of AdS implies that the center of mass will generally be deeper inside than the original points. Importantly the z value of the centroid now has a finite limit as $z_0 \rightarrow 0$, in which case it depends on the (appropriately averaged) relative separation of the points. We then perform a final embedding of the point using this z value, i.e.

$$X_i^\mu = (x_i^a, \hat{z}_i). \quad (12)$$

Intuitively, the z coordinate corresponds to the *length scale* of the degrees of freedom of a system⁹. Our choice above amounts to saying that the appropriate length scale for a point x_i is related to its distance from its neighbours. This exactly preserves scale invariance, but it gently breaks special conformal transformations. This is expected on physical grounds, as generally any choice of regulator necessarily breaks conformal invariance (Cardy, 1996). In experiments, we check generalization under special conformal transformations empirically and verify that the breaking is mild.

Proposition 4.1.1. *The lifting procedure (i.e. the map from $x^a \in \mathbb{R}^d$ to $X^\mu \in \text{AdS}^{d+1}$ described in equation 12) is equivariant under the subgroup of the conformal group generated by translations, rotations and dilatations.*

This is proved in Appendix A.3. Finally, we have discussed lifting the points x_i . The input data may also have some *features* h_i^{input} associated to each point x_i . They should be interpreted as a sample of an underlying conformal field $\mathbb{O}(x)$ with dimension Δ , where $h_i^{\text{input}} = \mathbb{O}(x_i)$. This boundary feature should be contrasted with the feature associated with a bulk point in AdS^{d+1} ; this is a scalar

⁹This is familiar from the physics of the AdS/CFT correspondence, where it is well-understood that the infrared physics lives deeper in the bulk (Susskind & Witten, 1998).

and does not transform with an associated factor of λ^Δ as in equation 3. Said differently, the full dependence of bulk features under scaling arises from its dependence on an extra coordinate z . This difference in representation is taken into account by lifting the feature as follows

$$h_i^{\text{lifted}} = \hat{z}_i^\Delta h_i^{\text{input}} \quad (13)$$

For many cases (e.g. image data) the input feature will have $\Delta = 0$ and this step may be skipped. This relation has an analogue in terms of bulk-to-boundary propagators from the AdS/CFT correspondence, where such factors of z^Δ relate physics in the bulk (i.e. h_i^{lifted}) to that of the boundary (i.e. h_i^{input})¹⁰

Given this set of points $\{X_i\}$ in AdS, we now operate on it using a graph neural network.

4.2 INVARIANT MESSAGE PASSING

We begin by studying scalar features. To orient ourselves, we recall first an earlier model, that of E(n) Equivariant Graph Neural Networks (EGNNs) (Satorras et al., 2021; Liu et al., 2024). These are graph neural networks that are equivariant to flat-space rotations, translations, reflections and permutations. The input to the model is a graph $G = (\mathcal{V}, \mathcal{E})$ whose vertices are embedded into Euclidean space \mathbb{R}^d . We denote the position of node v_i as $\mathbf{p}_i \in \mathbb{R}^d$ and its latent D -dimensional feature vector of node v_i as \mathbf{h}_i . The l -th layer of EGNN is then defined as

$$\begin{aligned} \mathbf{m}_{ij} &= \psi_e(\mathbf{h}_i^l, \mathbf{h}_j^l, \|\mathbf{p}_i - \mathbf{p}_j\|_2), & \text{EGNN message} & \quad (14) \\ \mathbf{h}_i^{l+1} &= \psi_h(\mathbf{h}_i^l, \mathbf{m}_i), \quad \mathbf{m}_i = \sum_{j \in \mathcal{N}(i)} \mathbf{m}_{ij}, & \text{aggregate + update} & \end{aligned}$$

where here $\mathcal{N}(i)$ represents the set of neighbours of node v_i , ψ_e, ψ_h are message and update MLPs, which we see depend only on the pairwise distance in \mathbb{R}^d between nodes.

We adopt the model above to operate on AdS where a graph G is embedded. As above, each node v_i has a position $X_i \in \text{AdS}^{d+1}$ and a latent feature vector \mathbf{h}_i . If edges \mathcal{E} are not provided in the graph, we induce connectivity with k_{con} nearest neighbours using the AdS proper distance (10).

In the message function (14), we also use the AdS proper distance instead of the Euclidean one, i.e.

$$\mathbf{m}_{ij} = \psi_e(\mathbf{h}_i^l, \mathbf{h}_j^l, D(X_i, X_j)), \quad \text{AdS-GNN message} \quad (15)$$

which yields an efficient conformal group equivariant GNN without substantial computational overhead compared to its Euclidean counterpart. Note that by making use of the AdS proper distance, we introduce a notion of locality both in ordinary space and in scale (as represented by the z -coordinate). Although the embedding of the point cloud mildly breaks special conformal transformations, the graph neural network itself is exactly invariant under all of $\text{Conf}(\mathbb{R}^d)$.

The message function in Eq. 15 only depends on distance and therefore is restricted to invariant features. To enable more expressive equivariance, we want to exploit the fact that AdS^{d+1} is a $\text{Conf}(\mathbb{R}^d)$ -equivariantly embedded sub-manifold of $\mathbb{R}^{d+1,1}$, as discussed in Section A.1. This allows us to employ $O(d+1, 1)$ -equivariant neural networks from (Ruhe et al., 2023), which operate on multivectors, i.e. elements of the Clifford Algebra $\text{Cl}(\mathbb{R}^{d+1,1})$. This yields the following message structure:

$$\mathbf{M}_{ij} = \psi_e(\mathbf{H}_i^l, \mathbf{H}_j^l, X_i, X_j), \quad \text{AdS-CEGNN message} \quad (16)$$

where $\mathbf{M}_{ij}, \mathbf{H}_i \in \text{Cl}(\mathbb{R}^{d+1,1})^c$, $X_i \in \text{AdS}^{d+1} \subseteq \mathbb{R}^{d+1,1}$.

We summarize the equivariance properties of both networks:

Proposition 4.2.1. *After the lifting procedure described in equation 12), AdS-GNN is equivariant to the global conformal group: $\text{Conf}_g(\mathbb{R}^d) = \text{PO}(d+1, 1)$, and, AdS-CEGNN to the restricted conformal group: $\text{Conf}(\mathbb{R}^d) = \text{PO}^0(d+1, 1)$, the identity component of $\text{Conf}_g(\mathbb{R}^d)$, see D.5.1.*

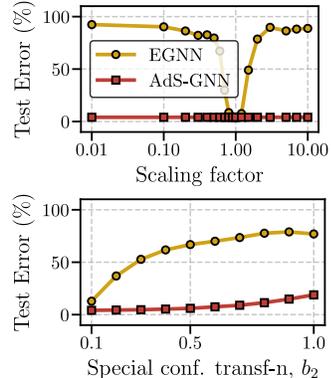
The framework above will result in features \mathbf{h}_i that are manifestly invariant under the conformal group. If performing an invariant task (e.g. classification) we can now aggregate the information in a permutation-invariant manner by summing over nodes as usual at the final layer. On the other hand,

¹⁰See e.g. Section 2.5 of (Witten, 1998) for more details.

Table 1: Classification error on SuperPixel MNIST. We augment the test set with random rotations and scaling. \times indicates random-guess performance. Baseline results are taken from (Bekkers et al., 2024).

Model	Error rate, %		
	non-augmented	rotated	rotated+scaled
MONET	8.89	\times	\times
SplineCNN	4.78	\times	\times
GCGP	4.2	\times	\times
GAT	3.81	\times	\times
PNCNN	1.24 ± 0.12	\times	\times
PΘNITA	1.17 ± 0.11	1.17	\times
EGNN	4.17 ± 0.45	4.17	\times
AdS-GNN (Ours)	4.09 ± 0.27	4.09	4.09

Figure 3: Test error on augmented data, SuperPixel MNIST.



for a regression task we should specify the transformation properties of the output; e.g. if the output from the network is a conformal field $\mathbb{O}(x_i)$ living on the boundary it must transform under scale transformations with a specified Δ as in equation 3. This can be accomplished by taking the output $\mathbb{O}(x_i)$ to be

$$\mathbb{O}(x_i) = \hat{z}_i^{-\Delta} h_i^{l_{\text{final}}} \quad (17)$$

where $h_i^{l_{\text{final}}}$ is the conformally invariant output from the node associated with x_i and \hat{z}_i the z -coordinate of the embedding found in equation 12. This relation is the inverse transformation of equation 13, and guarantees that the output satisfies equation 3.

Computational cost The computational cost of our approach decomposes into two components: the AdS embedding procedure (Algorithm 1) and applying the graph neural network. In the former, we employ k -nearest neighbors which has $\mathcal{O}(N \log N)$ complexity. An MPNN scales linearly with the number of nodes N in the graph. We note that in the case when a graph is available, one potentially can use it during the lifting, thus alleviating the overhead of k -nearest neighbors.

5 EXPERIMENTAL RESULTS

We test the framework above on tasks that are loosely divided into two types: computer vision tasks and applications from physics. For most of the tasks we use the scalar AdS-GNN.

5.1 COMPUTER VISION TASKS

SuperPixel MNIST We benchmark AdS-GNN on the super-pixel MNIST dataset (Monti et al., 2017), which consists of 2D point clouds of MNIST digits segmented into 75 superpixels. Results are given in Table 1. For in-distribution data, AdS-GNN performs on par with its roto-equivariant counterparts. It does however fall slightly behind PΘNITA, which has orientational information; we feel this happens as AdS-GNN is unable to handle orientation and relies on invariant descriptors. We also study the response of a model to various augmentations (see Fig. 3); for this we compare to EGNN and find that AdS-GNN has much stronger generalization capacities. As expected, AdS-GNN is precisely scale-invariant. For special conformal transformations, there is a small breaking of symmetry arising from the uplift. We empirically measure the effect on augmented performance in Figure 3 for a special conformal transformation as in (2) parametrized by $b = (0, b_2)$ and verify that it is very small.

Shape segmentation We further benchmark against a shape recognition dataset, where we sample points from a selection of randomly sized and rotated shapes (either a square, a circle, or a triangle). The task is to assign a shape class to each point in the cloud. We benchmark against EGNN and a message-passing neural network MPNN which conditions messages against $x_i - x_j$ (and thus has only translational equivariance, not rotational). Here, as shown in the left panel of Figure 4 we find that AdS-GNN outperforms EGNN even on in-distribution performance. We believe that this happens as the training data contains structure at different sizes to which AdS-GNN adapts more efficiently

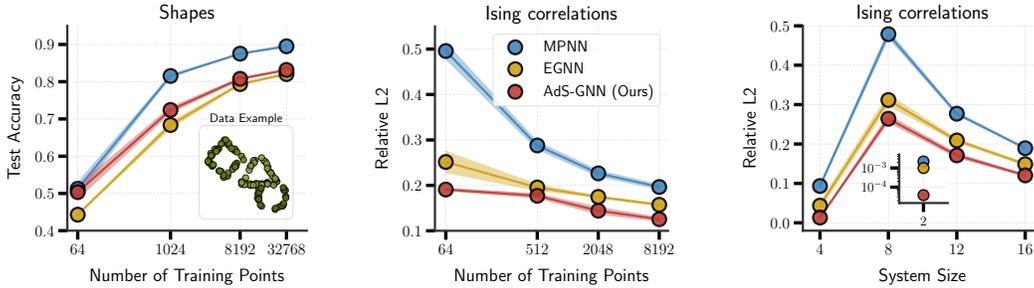


Figure 4: Performance on shape segmentation and the Ising task. *Left*, test accuracy of shape segmentation as a function of the number of training points. *Middle*, relative L2 as a function of the number of training points, with system size fixed at $N = 16$. *Right*, loss as a function of system size with 8192 training points. Inset shows $N = 2$.

than EGNN, particular when there is a small number of training points. Both of them fall behind MPNN, which we again attribute to their lack of orientational information.

A further segmentation task – PascalVOC – is discussed in the Appendix, where we find that AdS-GNN is essentially the same as EGNN.

5.2 PHYSICS TASKS

2d Ising We next consider a task from statistical physics, that of predicting N -point correlation functions in the 2d Ising model. As we review in more detail in Appendix C, this is a simple model of magnetism with a tunable temperature. As one increases it the model undergoes a phase transition, and at this point fluctuations of the magnetization exhibit conformal symmetry and are precisely understood in terms of a continuum model called the 2d Ising conformal field theory. This model is exactly solvable and has two non-trivial types of conformal field, the spin $\sigma(x)$ and the energy $\epsilon(x)$, which have conformal dimensions $\Delta_\sigma = \frac{1}{8}$ and $\Delta_\epsilon = 1$ respectively¹¹.

To benchmark AdS-GNN we consider the task of predicting the N -th moment of the energy and spin operators $\langle \epsilon(x_1)\epsilon(x_2)\cdots\epsilon(x_N) \rangle$ as a function of their input points. These functions are known explicitly in 2d (see equation 44) and are rather complicated: they have an erratic pattern of spikes when two points come close, on top of a more gentle modulation arising from the background of the other points.

We create a training dataset where we sample the coordinates of input points uniformly in $[-2, 2]$ for various values of N and use conformal field theory results for the N -th moment as the training data. We simultaneously predict the spin and energy moments, using the sum of relative L2 losses. As the output is the N -point function of conformal fields with a non-trivial Δ we use N copies of equation 17 to find our prediction $\text{Pred}_a(\{x_i\})$ in each channel a to be

$$\log(\text{Pred}_a(\{x_i\})) = \text{AdSGNN}_a(\{x_i\}) - \Delta_a \sum_{i=1}^N \log(\hat{z}_i) \quad (18)$$

where $a \in \{\sigma, \epsilon\}$ and the Δ_a 's are trainable parameters which can be interpreted as the learned conformal dimension of the σ and ϵ fields respectively.

We benchmark against EGNN and a baseline message-passing neural network (MPNN) whose messages are conditioned against $x_i - x_j$, and which thus only has translational equivariance. The results are shown in Figure 4 for the choice $k_{\text{lift}} = 1$. We see that the performance of AdS-GNN is *superior in all regimes*. We also note various useful features. Note that for 2 points (inset on right panel of Figure 4) AdSGNN performs more than an order of magnitude better than the others. This is because the form of the two-point function is completely fixed by conformal invariance to be proportional to $|x_1 - x_2|^{-2\Delta}$, and thus AdS-GNN need only learn two numbers, whereas the others must learn to reconstruct the functional form of the power law.

¹¹There are other slightly non-local fields that we will not discuss here.

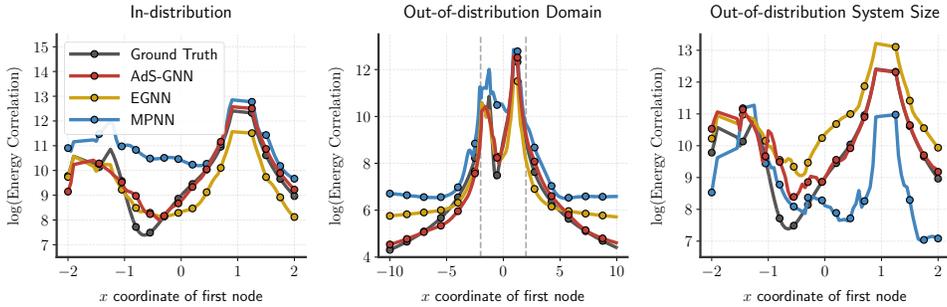


Figure 5: Visualization of the output from various models on the Ising task; here all points are fixed except for the x coordinate of the first, which is varied. *Left*, models tested on in-distribution data. *Center*, testing on values of x which are far out of the training range, which is shown with dashed lines. *Right*, testing a model trained on $N = 8$ on a test set with $N = 16$.

Interpretability: from equation 18, the learned values of the conformal dimensions Δ_ϵ and Δ_σ may be read off at the end of training, as shown in Table 3. They are very close to the ground truth, showing the useful ability of this model to extract conformal dimensions from data.

Generalization: AdS-GNN generalizes better than EGNN in two different directions. As shown in Figure 5, it generalizes well to values of x that are outside the training data, as one might have expected from scale-equivariance. AdS-GNN also generalizes well *across* N ; as shown in an example in Figure 5, AdS-GNN trained with one value of N works well when asked to predict correlation functions with a different N , suggesting that it has robustly understood the underlying physics. It always generalizes better than EGNN, particularly when the graph connectivity is more dense. This is discussed quantitatively in Appendix B.2.

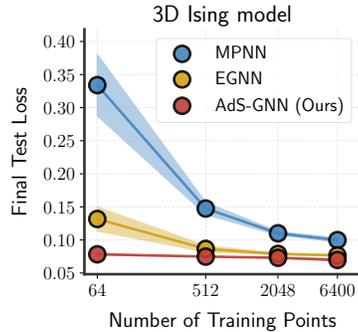


Figure 6: Relative L2 loss for the Ising 3D task. AdS-GNN also correctly recovers the conformal dimension of spin $\Delta_\sigma = 0.518$.

3d Ising We perform a similar analysis for the 3d Ising model. This model is not exactly solvable, and there are no simple formulas for the N -point functions to use for training data. Nevertheless it is possible to use modern conformal bootstrap methods (reviewed in Appendix C.2) to compute a good approximation in the case $N = 4$ for the spin operator. We use this to build a training dataset and then perform the same regression as above. Our results are shown in Figure 6. Just as in the case of the 2d Ising model, AdS-GNN clearly has superior performance.

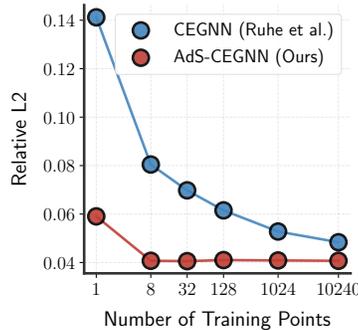


Figure 7: Relative L2 loss for the N-body task. The L_2 error is computed across long trajectories.

N-body simulation In the last experiment, we employ a neural network model for learning the dynamics of a system of 10 charged particles. The particles interact through electric fields in three spatial dimension; the inverse-square law of the electric field means that this a conformally invariant problem in 3d. The acceleration is expected to be conformal vector field with $\Delta_a = 2$ (as it is determined by an electric field, whose conformal dimension we discussed below equation 3).

We forecast acceleration value $a_t \in \mathbb{R}^{10 \times 3}$ for each body given only the position $x_t \in \mathbb{R}^{10 \times 3}$ at the time step t . At training, the model is learned from random slices of short ($T = 4$) training trajectories. During evaluation, we generate a set of very long trajectories (2,000 time steps) and measure the

relative L_2 error between the ground-truth trajectories and those unrolled by the models. The task involves handling vector information, hence we use AdS-CEGNN. We benchmark against CEGNN (Ruhe et al., 2023), which is SOTA at N-body simulation.

AdS-CEGNN outperforms CEGNN, as evident from the results in Fig.7. Furthermore, it indeed correctly recovers the conformal dimension associated with acceleration $\Delta_a = 2$. Therefore, AdS-CEGNN effectively provides useful feedback when the value of the conformal dimension is known a priori: if at the end of the training Δ is correct, the model has processed the data correctly. This is in general not the case for deep learning models whose quality is based exclusively on performance.

6 CONCLUSION

In this paper, we introduced AdS-GNN and AdS-CEGNN - neural networks that are equivariant with respect to conformal transformations. We have demonstrated strong performance on various tasks, including computer vision and N-body simulation. We found particularly interesting an application to the Ising model, where the model exhibited impressive generalization capacities and *interpretability*, in that the conformal dimensions Δ_a – important universal quantities – could be extracted from the trained model. One might ask what other such universal information exists: a general conformally invariant field theory in physics is defined in terms of these dimensions Δ_a and a set of 3-point coefficients c_{abc} which turn out to determine higher-order moments of the conformal fields¹². It is interesting to ask whether this information could be usefully extracted from the trained network. We anticipate further applications to critical phenomena and long-range interactions in physical systems.

Future Work Besides physics-related tasks, we see potential applications of AdS-GNN and AdS-CEGNN to computed vision and robotics tasks due to the inherent consistency of our models to local scale variations. That is, proposed models maintain consistent predictions when objects appear at varying scales within scenes, and do not require extensive multi-scale training data to achieve robustness.

ACKNOWLEDGMENTS

MZ was supported by Microsoft Research AI4Science. This work was supported by a grant from the Simons Foundation (PD-Pivot Fellow-00004147, NI and EB). NI is supported in part by the STFC under grant number ST/T000708/1. This publication is part of the project SIGN with file number VI.Vidi.233.220 of the research programme Vidi which is (partly) financed by the Dutch Research Council (NWO) under the grant <https://doi.org/10.61686/PKQGZ71565>

REFERENCES

- Ali R Amir-Moéz. Conformal Linear Transformations. *Mathematics Magazine*, 40(5):268–270, 1967.
- Erik J Bekkers. B-spline cnns on lie groups. In *International Conference on Learning Representations*, 2020. URL <https://openreview.net/forum?id=HlgBhkBFDH>.
- Erik J. Bekkers, Sharvaree P. Vadgama, Rob Hesselink, Putri A. van der Linden, and David W. Romero. Fast, expressive se(n) equivariant networks through weight-sharing in position-orientation space. In *ICLR*, 2024.
- Johann Brehmer, Sönke Behrends, Pim de Haan, and Taco Cohen. Does equivariance matter at scale? *arXiv preprint arXiv:2410.23179*, 2024.
- Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021.
- John Cardy. *Scaling and renormalization in statistical physics*, volume 5. Cambridge university press, 1996.

¹²See e.g. (Di Francesco et al., 1997; Simmons-Duffin, 2017) for textbook treatments.

- Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International conference on machine learning*, pp. 2990–2999. PMLR, 2016.
- Miguel S. Costa, Tobias Hansen, João Penedones, and Emilio Trevisani. Radial expansion for spinning conformal blocks. *JHEP*, 07:057, 2016. doi: 10.1007/JHEP07(2016)057.
- P. Di Francesco, P. Mathieu, and D. Senechal. *Conformal Field Theory*. Graduate Texts in Contemporary Physics. Springer-Verlag, New York, 1997. ISBN 978-0-387-94785-3, 978-1-4612-7475-9. doi: 10.1007/978-1-4612-2256-9.
- Vijay Prakash Dwivedi, Ladislav Rampásek, Michael Galkin, Ali Parviz, Guy Wolf, Anh Tuan Luu, and Dominique Beaini. Long range graph benchmark. In *NeurIPS*, 2022.
- Sheer El-Showk, Miguel F. Paulos, David Poland, Slava Rychkov, David Simmons-Duffin, and Alessandro Vichi. Solving the 3D Ising Model with the Conformal Bootstrap. *Phys. Rev. D*, 86: 025022, 2012. doi: 10.1103/PhysRevD.86.025022.
- Matthias Fey, Jan Eric Lenssen, Frank Weichert, and Heinrich Müller. Splinecnn: Fast geometric deep learning with continuous b-spline kernels. In *CVPR*, 2018.
- Marc Anton Finzi, Roberto Bondesan, and Max Welling. Probabilistic numeric convolutional neural networks. In *ICLR*, 2021.
- GA Galperin. A concept of the mass center of a system of material points in the constant curvature spaces. *Communications in Mathematical Physics*, 154:63–84, 1993.
- S. S. Gubser, Igor R. Klebanov, and Alexander M. Polyakov. Gauge theory correlators from noncritical string theory. *Phys. Lett. B*, 428:105–114, 1998. doi: 10.1016/S0370-2693(98)00377-3.
- James Halverson, Joydeep Naskar, and Jiahua Tian. Conformal Fields from Neural Networks. 9 2024.
- Yena Han, Gemma Roig, Gad Geiger, and Tomaso Poggio. Scale and translation-invariance for novel objects in human vision. *Scientific reports*, 10(1):1411, 2020.
- Mehran Kardar. *Statistical physics of fields*. Cambridge University Press, 2007.
- Zohar Komargodski and David Simmons-Duffin. The Random-Bond Ising Model in 2.01 and 3 Dimensions. *J. Phys. A*, 50(15):154001, 2017. doi: 10.1088/1751-8121/aa6087.
- Na Lei, Feng Luo, Shing-Tung Yau, and Xianfeng Gu. Computational conformal geometric methods for vision. In *Handbook of Mathematical Models and Algorithms in Computer Vision and Imaging: Mathematical Imaging and Vision*, pp. 1739–1790. Springer, 2023.
- Tzu-Yuan Lin, Minghan Zhu, and Maani Ghaffari. Lie neurons: Adjoint-equivariant neural networks for semisimple lie algebras. In *International Conference on Machine Learning*, pp. 30529–30545. PMLR, 2024.
- Cong Liu, David Ruhe, Floor Eijkelboom, and Patrick Forré. Clifford Group Equivariant Simplicial Message Passing Networks. *International Conference on Learning Representations (ICLR)*, 2024.
- Nikos K Logothetis, Jon Pauls, and Tomaso Poggio. Shape representation in the inferior temporal cortex of monkeys. *Current biology*, 5(5):552–563, 1995.
- Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *ICLR*, 2019.
- Juan Martin Maldacena. The Large N limit of superconformal field theories and supergravity. *Int. J. Theor. Phys.*, 38:1113–1133, 1999. doi: 10.1023/A:1026654312961. [Adv. Theor. Math. Phys.2,231(1998)].
- Benjamin McKay. An introduction to Cartan geometries. *arXiv preprint arXiv:2302.14457*, 2023.
- Mircea Mironenco and Patrick Forré. Lie group decompositions for equivariant neural networks. In *The Twelfth International Conference on Learning Representations*, 2024. URL <https://openreview.net/forum?id=p34fRKp8qA>.

- Federico Monti, Davide Boscaini, Jonathan Masci, Emanuele Rodolà, Jan Svoboda, and Michael M. Bronstein. Geometric deep learning on graphs and manifolds using mixture model cnns. In *CVPR*, 2017.
- Yu Nakayama. Scale invariance vs conformal invariance. *Phys. Rept.*, 569:1–93, 2015. doi: 10.1016/j.physrep.2014.12.003.
- David Poland, Slava Rychkov, and Alessandro Vichi. The Conformal Bootstrap: Theory, Numerical Techniques, and Applications. *Rev. Mod. Phys.*, 91:015002, 2019. doi: 10.1103/RevModPhys.91.015002.
- Joseph Polchinski. Scale and Conformal Invariance in Quantum Field Theory. *Nucl. Phys. B*, 303: 226–236, 1988. doi: 10.1016/0550-3213(88)90179-4.
- Md Ashiqur Rahman and Raymond A Yeh. Truly scale-equivariant deep nets with fourier layers. *Advances in Neural Information Processing Systems*, 36:6092–6104, 2023.
- David Ruhe, Johannes Brandstetter, and Patrick Forré. Clifford group equivariant neural networks. *Advances in Neural Information Processing Systems*, 36:62922–62990, 2023.
- Slava Rychkov, David Simmons-Duffin, and Bernardo Zan. Non-gaussianity of the critical 3d Ising model. *SciPost Phys.*, 2(1):001, 2017. doi: 10.21468/SciPostPhys.2.1.001.
- Victor Garcia Satorras, Emiel Hooeboom, and Max Welling. E(n) equivariant graph neural networks. In Marina Meila and Tong Zhang (eds.), *ICML*, 2021.
- Martin Schottenloher. *A Mathematical Introduction to Conformal Field Theory*. Lecture Notes in Physics. Springer: Berlin, Heidelberg, 2nd edition, 2008.
- Eitan Sharon and David Mumford. 2d-shape analysis using conformal mapping. *International Journal of Computer Vision*, 70(1):55–75, 2006.
- David Simmons-Duffin. The Conformal Bootstrap. In *Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings*, pp. 1–74, 2017. doi: 10.1142/9789813149441_0001.
- Ivan Sosnovik, Michał Szmaja, and Arnold Smeulders. Scale-equivariant steerable networks. *arXiv preprint arXiv:1910.11093*, 2019.
- Leonard Susskind and Edward Witten. The Holographic bound in anti-de Sitter space. 5 1998.
- Sharvaree Vadgama, Mohammad Mohaiminul Islam, Domas Buracus, Christian Shewmake, and Erik Bekkers. On the utility of equivariance and symmetry breaking in deep learning architectures on point clouds. *arXiv preprint arXiv:2501.01999*, 2025.
- Petar Velickovic, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. In *ICLR*, 2018.
- Ian Walker and Ben Glocker. Graph convolutional gaussian processes. In *ICML*, 2019.
- Maurice Weiler, Patrick Forré, Erik Verlinde, and Max Welling. *Equivariant and Coordinate Independent Convolutional Networks*. 2023. URL https://maurice-weiler.gitlab.io/cnn_book/EquivariantAndCoordinateIndependentCNNs.pdf.
- Kenneth G Wilson. The renormalization group and critical phenomena. *Reviews of Modern Physics*, 55(3):583, 1983.
- Andrew P Witkin. Scale-space filtering. In *Readings in computer vision*, pp. 329–332. Elsevier, 1987.
- Edward Witten. Anti de Sitter space and holography. *Adv. Theor. Math. Phys.*, 2:253–291, 1998. doi: 10.4310/ATMP.1998.v2.n2.a2.
- Daniel Worrall and Max Welling. Deep scale-spaces: Equivariance over scale. *Advances in Neural Information Processing Systems*, 32, 2019.

Maksim Zhdanov, Nico Hoffmann, and Gabriele Cesa. Implicit convolutional kernels for steerable CNNs. In *Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS), 2023*. URL <https://openreview.net/forum?id=2YtdxqvjdjX>.

Maksim Zhdanov, David Ruhe, Maurice Weiler, Ana Lucic, Johannes Brandstetter, and Patrick Forré. Clifford-steerable convolutional neural networks. In *Forty-first International Conference on Machine Learning, 2024*. URL <https://openreview.net/forum?id=XTglHJjzQI>.

Jean Zinn-Justin. *Quantum field theory and critical phenomena*, volume 171. Oxford university press, 2021.

A AdS^{d+1} DETAILS

We provide a few further details about aspects of AdS^{d+1}:

A.1 THE GROUP ACTION ON AdS

Recall that AdS^{d+1} can be understood as the hyperboloid defined in equation 5. The action of $[\Lambda] \in O(d+1, 1)/\{\pm I\}$ then works concretely as follows

$$[\Lambda].Y := \text{sign}((\Lambda Y)^0) \cdot (\Lambda Y) \quad (19)$$

where ΛY is just vector matrix multiplication and where the scalar multiplication with $\text{sign}((\Lambda Y)^0)$ corrects the sign of the 0-th entry (the component of (-1) -signature). Together we get:

$$(([\Lambda].Y)^0)^2 - \sum_{i=1}^{d+1} (([\Lambda].Y)^i)^2 = 1, \quad ([\Lambda].Y)^0 > 0, \quad (20)$$

and thus: $[\Lambda].Y \in \text{AdS}^{d+1}$. This gives us a well-defined group action of $O(d+1, 1)/\{\pm I\}$ on AdS^{d+1} , and, in particular, a well-defined group action of $\text{Conf}(\mathbb{R}^d)$ on AdS^{d+1} . Note, however, that this action does not act linearly (since $-I$ acts trivially). This shows that we have now two different group actions: i) the linear action, given just by matrix multiplication, and ii) the sign-corrected action, given by Equation 19. It is a subtle, but important point to notice that for every $\Lambda \in O^0(d+1, 1)$, the connected component of the identity I , we have $\text{sign}((\Lambda Y)^0) = 1$ whenever $Y^0 > 0$. So for those elements $\Lambda \in O^0(d+1, 1)$ these two actions agree. Note that $-I \notin O^0(d+1, 1)$, see D.1.5. Since we can canonically identify the (restricted) conformal group with the above:

$$\text{Conf}(\mathbb{R}^d) = \text{PO}^0(d+1, 1) \cong O^0(d+1, 1), \quad (21)$$

we get a well-defined *linear* action of $\text{Conf}(\mathbb{R}^d)$ on AdS^{d+1} , meaning that the inclusion map:

$$j : \text{AdS}^{d+1} \subseteq \mathbb{R}^{d+1,1} \quad (22)$$

is equivariant w.r.t. $\text{Conf}(\mathbb{R}^d)$, considered as the subgroup $\text{Conf}(\mathbb{R}^d) \cong O^0(d+1, 1) \subseteq O(d+1, 1)$. This implies that any $O(d+1, 1)$ -equivariant map (like CGENNs) $\psi : \mathbb{R}^{d+1,1} \rightarrow \mathcal{H}$ induces a $\text{Conf}(\mathbb{R}^d)$ -equivariant map:

$$\psi \circ j : \text{AdS}^{d+1} \rightarrow \mathbb{R}^{d+1,1} \rightarrow \mathcal{H}. \quad (23)$$

This insight will be used to construct conformally equivariant graph neural networks (with multi-vector features).

A.2 THE CENTER OF MASS OF A SET OF POINTS ON ADS

We will require an expression for the ‘‘center of mass’’ $C(\{X_i\})$ of a set of points on AdS. This problem was solved in (Galperin, 1993); the basic idea is to view the hyperboloid as a submanifold of $\mathbb{R}^{d+1,1}$ as above, use additivity properties there to find a vector, and then find the intersection of the ray in the direction of that vector with the hyperboloid.

In practice, this is quite simple to implement. Denote the center of mass by \bar{Y}^A , and the set of N points for which we want the centroid by (x_i^a, z_i) . We would like to find the analogous coordinates for the centroid (\bar{x}^a, \bar{z}) .

We have that

$$\bar{Y}^0 \equiv \frac{\bar{z}}{2} \left(1 + \frac{1}{\bar{z}^2} \left(1 + \sum_a \bar{x}^a \bar{x}^a \right) \right) = \frac{1}{\mathcal{N}\mathcal{N}} \sum_i \frac{z_i}{2} \left(1 + \frac{1}{z_i^2} \left(1 + \sum_a x_i^a x_i^a \right) \right) \quad (24)$$

$$\bar{Y}^a \equiv \frac{\bar{x}^a}{\bar{z}} = \frac{1}{\mathcal{N}\mathcal{N}} \sum_i \frac{x_i^a}{z_i} \quad (25)$$

$$\bar{Y}^{d+1} \equiv \frac{\bar{z}}{2} \left(1 - \frac{1}{\bar{z}^2} \left(1 - \sum_a \bar{x}^a \bar{x}^a \right) \right) = \frac{1}{\mathcal{N}\mathcal{N}} \sum_i \frac{z_i}{2} \left(1 - \frac{1}{z_i^2} \left(1 - \sum_a x_i^a x_i^a \right) \right) \quad (26)$$

The first equality is the definition of the embedding, the second is the definition of the centroid from Galperin. Here \mathcal{N} is a normalization constant which is picked to guarantee that

$$(\bar{Y}^0)^2 - \sum_a (\bar{Y}^a)^2 - (\bar{Y}^{d+1})^2 = 1 \quad (27)$$

So to find the centroid, the easiest thing to do is to compute the sums on the right hand side of the second equality, which thus determines the vector \bar{Y}^A up to an overall scale \mathcal{N} ; then we enforce the norm constraint above which lets us find \mathcal{N} and thus fixes the vector \bar{Y}^A completely. We then express the answer in useful coordinates by solving for (\bar{z}, \bar{x}^a) through

$$\bar{z} = \frac{1}{\bar{Y}^0 - \bar{Y}^{d+1}}, \quad \bar{x}^a = \frac{\bar{Y}^a}{\bar{Y}^0 - \bar{Y}^{d+1}}. \quad (28)$$

To get some intuition for the procedure, we study it in the case of two points $X_1 = (x_1^a, \epsilon)$ and $X_2 = (x_2^a, \epsilon)$ starting at the same value of the z coordinate. We find

$$C(X_1, X_2) = \left(\frac{1}{2}(x_1^a + x_2^a), \frac{1}{2}\sqrt{|x_1 - x_2|^2 + 4\epsilon^2} \right) \quad (29)$$

i.e. we simply take the average of the spatial coordinates and move inwards in z by an amount which depends on the separation between the two points in the spatial direction. In this case the center of mass is actually the midpoint of the geodesic that connects the two points.

A.3 EQUIVARIANCE OF LIFTING PROCEDURE

Here we prove Proposition 4.1.1, i.e. that the lifting procedure is equivariant under translations, rotations and dilatations but not special conformal transformations. We denote a general element of the conformal group by G . G acts on points $x_i^a \in \mathbb{R}^d$ as in equation 2 and on points in $X_i^\mu \in \text{AdS}^{d+1}$ as in equation 8.

The lifting begins by embedding a set of points $\{x_i\} \in \mathbb{R}^d$ into AdS^d using a small regulator value $z = z_0$ as in equation 11:

$$X_{i,0}^\mu(x_i^a) := (x_i^a, z = z_0) \quad (30)$$

We then compute the AdS center of mass $C(\{X_{i,0}\})$ of the set of points using the algorithm above. The question is whether this approach is equivariant in the bulk when the conformal transformation acts *only* on the boundary coordinates x^a , i.e. is it true that

$$C(\{X_{i,0}(Gx_i^a)\}) \stackrel{?}{=} GC(\{X_{i,0}(x_i^a)\}) \quad (31)$$

(Note that if the conformal transformation also acted on z_0 then this would certainly be equivariant; the question at hand is how much the choice of fixed regulator z_0 spoils the equivariance).

It is immediately clear that the procedure is equivariant under translations and rotations, as in those cases we see from equation 8 that the z coordinate is left invariant under the transformation. Through direct computation using equation 2 and equation 8, we can further explicitly verify that the equality in equation 31 holds for dilatations but does *not* hold for special conformal transformations.

This may seem somewhat surprising, and we try to understand it more abstractly by studying the geometry of the embedded points $X_{i,0}$ in the Y^A coordinates defined in equation 6:

$$Y_i^0 := \frac{z_0^2 + \|x_i\|^2 + 1}{2z_0}, \quad Y_i^a := \frac{x_i^a}{z_0}, \quad Y_i^{d+1} := \frac{z_0^2 + \|x_i\|^2 - 1}{2z_0} \quad (32)$$

As we take $z_0 \rightarrow 0$ while holding x_i fixed, we can extract the dependence on z_0 by writing these points as $Y_i^A = \frac{1}{z_0} Y_{i,\text{null}}^A$. Explicitly we then have

$$Y_{i,\text{null}}^0 := \frac{\|x_i\|^2 + 1}{2}, \quad Y_{i,\text{null}}^a := x_i^a, \quad Y_{i,\text{null}}^{d+1} := \frac{\|x_i\|^2 - 1}{2} \quad (33)$$

Here the points $Y_{i,\text{null}}^A$ lie on the null cone defined by $\|Y\| = 0$, where the norm is taken with respect to the $\mathbb{R}^{d+1,1}$ metric defined in equation 4. The points on this cone no longer depend explicitly on z_0 , and furthermore the cone has a natural $O(d+1, 1)$ action on it given by regular matrix multiplication on Y^A as in equation 19.

Naively, one would now expect that one could just ignore the overall normalizing factor and sum up the points in the cone when computing the sum in equation 24. As $Y^0 > 0$ for all points this would lead to a point in the *interior* of the null cone (i.e. in AdS^{d+1} itself) and the procedure now appears to make no reference to z_0 and is manifestly equivariant under all of $O(d+1, 1)$.

Interestingly, this is not the case. It turns out that under $x^a \rightarrow Gx^a$ the points on the cone do not transform with the usual action of $O(d+1, 1)$; instead they acquire an overall scalar rescaling factor σ which depends on the parameters of G and the point in question, i.e.

$$GY_{i,\text{null}}(x_i^a) = \sigma(G, x^a) Y_{i,\text{null}}(Gx_i^a) \quad (34)$$

The scalar factor preserves the cone $\|Y\| = 0$, but it means that the sum in equation 24 no longer transforms equivariantly.

The existence of this factor can already be seen in the simplest case $d = 1$, where the embedding simply becomes

$$Y_i^0 := \frac{z_0^2 + x^2 + 1}{2z_0}, \quad Y_i^1 := \frac{x}{z_0}, \quad Y_i^2 := \frac{z_0^2 + x^2 - 1}{2z_0} \quad (35)$$

We can now compute the prefactor σ for various transformations by computing both sides of equation 34. Under a dilatation $x \rightarrow \lambda x$ the overall prefactor $\sigma(\lambda, x)$ is simply a constant λ^{-1} which can indeed be ignored when computing the center of mass; thus the lifting procedure is equivariant under dilatations. However under a special conformal transformation $x \rightarrow \frac{x}{bx+1}$ the prefactor is instead $\sigma(b, x) = (1 + bx)^2$ and thus depends on the point in question. This spoils the equivariance under special conformal transformations.

Our discussion here is somewhat formal but the conclusion is expected on physical grounds, as in general it is difficult to preserve full conformal invariance in the presence of a regulator such as z_0 .

A.4 VECTOR BULK-TO-BOUNDARY MAPPING

Given an $O(d+1, 1)$ vector feature v_A^{bulk} expressed in the Y^A coordinates in the bulk we may need to map it back to the boundary to interpret it as a neural network output. This is first done by using the Jacobian associated with equation 6 to map it to an orthonormal basis in the (x^a, z) coordinates:

$$v_{\hat{\mu}}^{\text{bulk}} = e_{\hat{\mu}}^{\mu} \frac{\partial Y^A}{\partial X^{\mu}} v_A^{\text{bulk}} \quad (36)$$

where we have used the Einstein summation convention, where $X^{\mu} = (x^a, z)$, we have denoted and where $e_{\hat{\mu}}^{\mu}$ is the inverse vielbein which satisfies:

$$e_{\hat{\mu}}^{\mu} e_{\hat{\nu}}^{\nu} g_{\mu\nu} = \delta_{\hat{\mu}\hat{\nu}} \quad (37)$$

where $g_{\mu\nu}$ is the AdS^{d+1} metric. In the coordinates used in equation 7 this vielbein is simply $e_{\hat{\mu}}^{\mu} = z \delta_{\hat{\mu}}^{\mu}$. Finally the analogue of equation 17 is

$$v_{\hat{\mu}}^{\text{output}} = z^{-\Delta} v_{\hat{\mu}}^{\text{bulk}} \quad (38)$$

Working out the Jacobian explicitly we find that the spatial components of the vector are

$$v_{\hat{a}}^{\text{output}} = z^{-\Delta} (v_a^{\text{bulk}} + x_a (v_{d+1}^{\text{bulk}} + v_0^{\text{bulk}})) \quad (39)$$

where we stress that the components on the right-hand side are of the vector v_A^{bulk} expressed in the Y^A coordinates, as is natural for the Clifford algebra construction.

B FURTHER EXPERIMENTAL DETAILS

B.1 PASCALVOC-SP

In a further experiment, we also compare AdS-GNN to EGNN on the LRGB data (Dwivedi et al., 2022), see Table 2. This is a pixel segmentation task; thus the output data at each node is conformally *invariant* and we take $\Delta = 0$ in equation 17, i.e. we read out the output from \mathbf{h}_i directly. The difference in performance between EGNN and AdS-GNN is statistically insignificant, which indicates that in this task conformal equivariance does not constrain the model significantly and still allows for high expressivity.

Table 2: Classification error on PascalVOC-SP.

Model	EGNN	AdS-GNN
Test F1 \uparrow	27.80 ± 0.74	28.07 ± 0.57

B.2 2D ISING GENERALIZABILITY

Here we discuss further the generalization properties of the 2d Ising task discussed in the main text. In particular, we provide information on how a model trained on a given number of nodes N_{train} performs when evaluated on a different number of nodes N_{test} . The connectivity on the graph on AdS is controlled by a parameter k_{con} , the number of nearest neighbours which are connected to each node.

For most of our results in the Ising section, we pick $k_{\text{con}} = N$ so that the graph is fully connected. This results in the best in-distribution performance. In that case note from Figures 8a and 8b that AdS-GNN generalizes dramatically better than EGNN. EGNN’s performance is particularly bad when taking a model trained on a larger system and evaluating it on a smaller one.

However if we reduce k_{con} – e.g. to 2 – then this greatly increases the generalization ability of both models. In particular, the gap between EGNN and AdS-GNN (as seen in Figures 8c and 8d) is much narrower, though still present. We see however that this choice hurts in-distribution performance; AdS-GNN is able to generalize reasonably well even in the fully connected case.

B.3 IMPLEMENTATION DETAILS

In every experiment, we use the AdamW optimizer (Loshchilov & Hutter, 2019) with a learning rate 10^{-3} . Every model is trained on a single Nvidia RTX6000 GPU. All models are implemented in JAX. All experiments are run 5 times with different seeds. Models are trained until convergence with early stopping. The number of layers is fixed to 4, with hidden dimension 32.

SuperPixel-MNIST The task is to predict a digit given a point cloud representation. We compare against MONET (Monti et al., 2017), SplineCNN (Fey et al., 2018), GCCP (Walker & Glocker, 2019), GAT (Velickovic et al., 2018), PNCNN (Finzi et al., 2021) and PΘNITA (Bekkers et al., 2024). Every model is trained with batch size 128, baseline results are taken from (Bekkers et al., 2024). k_{con} is set to 16, k_{lift} to 5. Training time is approximately 10 minutes.

Pascal-VOC The task is to predict a semantic segmentation label for each superpixel node (total of 21 classes). Each graph is embedded in 2D Euclidean space, each node is associated with 12 scalar features. We used the batch size of 96. $k_{\text{con}} = 16$, $k_{\text{lift}} = 5$. Training time is approximately 30 minutes.

System size	Δ_ϵ (energy)	Δ_σ (spin)
2	1.0000 ± 0.0000	0.1250 ± 0.000
4	0.9998 ± 0.0000	0.1250 ± 0.000
8	0.9924 ± 0.0010	0.1248 ± 0.000
12	0.9894 ± 0.0032	0.1247 ± 0.000
16	0.9893 ± 0.0007	0.1247 ± 0.000

Table 3: Learned value of Δ ’s from AdS-GNN. Ground truth values are $\Delta_\epsilon = 1$ and $\Delta_\sigma = 0.125$. Statistical uncertainties are standard deviations from 5 runs; for Δ_σ they are $O(10^{-5})$ and are not quoted.

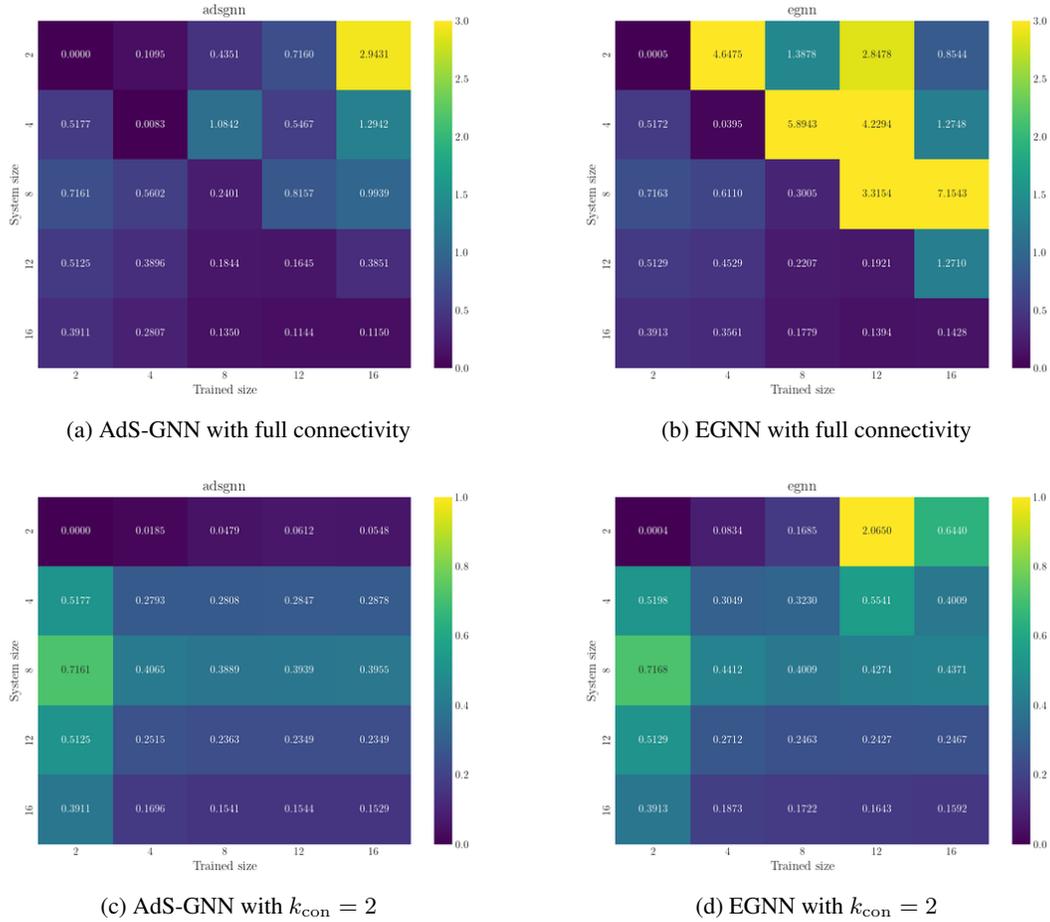


Figure 8: Generalization across system size; each square shows the relative L2 loss of a system trained on a system of size N_{train} (x -axis) when tested on a system of size N_{test} (y -axis) for different level of graph connectivity.

Shapes Given a point cloud, we predict for each point to each shape it belongs (circle, square, triangle, intersection). We train by minimizing cross-entropy loss. k_{con} is set to 16, $k_{\text{lift}} = 16$. The number of testing and validation points is set to 512, while the number of training points varies from 64 to 8192. Training time is approximately 5 minutes.

Ising Given a set of points, we predict two scalar values, one for energy correlation, one for spin correlation. We use relative L2 error as our training objective. k_{con} is equal to the number of points (i.e. fully connected system), $k_{\text{lift}} = 1$. The number of testing and validation points is set to 512, while the number of training points varies from 64 to 32768. Training time is approximately 3 minutes.

C ISING MODEL REVIEW

Here, for completeness, we provide an elementary review of the Ising model – which can be imagined as the simplest example of a model of magnetism – and the physics at its critical point. This material is standard; see e.g. (Kardar, 2007; Di Francesco et al., 1997) for textbook treatments. We begin by discussing the model in 2d.

C.1 2D ISING MODEL

The model is defined in terms of binary variables called *spins* $\sigma_i = \pm 1$ sitting on the sites i of a square lattice with L sites on a side. The model is defined in terms of an energy function:

$$E[\sigma] = - \sum_{\langle ij \rangle} \sigma_i \sigma_j \quad (40)$$

where the notation $\langle ij \rangle$ means that one sums over nearest-neighbour links connecting two adjacent sites i, j . We can see that the energy is minimized when spins on two adjacent sites have the same value, i.e. “spins want to align”. The energy is also invariant under a \mathbb{Z}_2 symmetry which acts by flipping all of the spins, $\sigma_i \rightarrow -\sigma_i$.

This energy defines a statistical physics model in which the probability of obtaining a given spin configuration $\{\sigma\}$ is given by

$$p_\beta[\sigma] = \frac{1}{Z} \exp(-\beta E[\sigma]) \quad (41)$$

where β is the inverse temperature and Z the usual normalizing constant. This model has two *phases*, which we now describe.

Consider first taking β very large; in that case any increase in energy will be heavily penalized, and the most likely configurations will be those that minimize the energy, i.e. where all spins have the same value, so either $\sigma_i = +1$ or $\sigma_i = -1$ for all i . This is called a phase with *spontaneous symmetry breaking*, as a choice of either of these configurations breaks the \mathbb{Z}_2 symmetry. It is also called the *ordered* or *ferromagnetic* phase.

Now consider taking β very small; in that case the system is very disordered, and all spins fluctuate strongly and randomly, and there is no sense in which a symmetry is spontaneously broken. This is called the *symmetry unbroken* or *disordered* phase or the *paramagnet*.

In the $L \rightarrow \infty$ limit there is a sharp distinction between the two phases. A quantitative way to understand it is to imagine taking the system size to infinity while computing e.g.

$$\langle \sigma \rangle = \lim_{L \rightarrow \infty} \left(\frac{1}{L^2} \mathbb{E}_{\sigma \sim p_\beta} \sum_i \sigma_i \right) \quad (42)$$

i.e. the expectation value of the spatial average of all the spins. This is called the order parameter. This is nonzero in the ordered phase (where all the spins are aligned, resulting in a net contribution to the expectation value) and zero in the disordered phase (where all the spins fluctuate strongly, resulting in a cancellation across sites). In the infinite- L limit there is a non-analyticity in the function $\langle \sigma(\beta) \rangle$ at a critical value of $\beta = \beta_c$ at the phase transition point. For the 2d Ising model the location of this point is known to be at $\beta_c = \frac{1}{2} \log(1 + \sqrt{2}) \approx 0.441$.

A great deal is known about this critical point. Here fluctuations of the spins take place over all scales, and do not decay exponentially with distance as one might normally expect. It is possible to capture the long-distance statistics of these fluctuations in a continuum limit where we formally take the lattice spacing to zero. The resulting structure is called the *2d Ising conformal field theory* (CFT), and is an example of a quantum field theory that exhibits conformal symmetry. In particular, one can define two operators in this continuum theory: the spin operator $\sigma(x)$ (which is the continuum limit of the spin operator σ_i defined on discrete lattice sites above) and the energy operator $\epsilon(x)$ (which can be thought of as a product of spins at adjacent sites). The 2d Ising CFT is completely solved and thus one can compute any arbitrary moments of any of these operators in closed form.

To get some intuition, the correlation function of two spins behaves as:

$$\langle \sigma(x)\sigma(y) \rangle = |x - y|^{-2\Delta_\sigma} \quad (43)$$

where $\Delta_\sigma = \frac{1}{8}$. The power-law functional form is completely fixed by scale-invariance, and the only input from the theory here is the value of the conformal dimension Δ_σ . A similar relation holds for the energy operator with $\Delta_\epsilon = 1$.

In this work we build a neural network to predict the N -point correlation functions of the spin and energy operators as a function of the positions of the operator insertions. To do this we use as ground-truth training data the following closed-form formulas from the theory of the 2d Ising CFT (Di Francesco et al., 1997):

$$\langle \epsilon(z_1)\epsilon(z_2)\cdots\epsilon(z_N) \rangle = \left| \text{Pf} \left[\frac{1}{z_i - z_j} \right]_{1 \leq i, j \leq N} \right|^2 \quad (44)$$

$$\langle \sigma(z_1)\sigma(z_2)\cdots\sigma(z_N) \rangle^2 = \frac{1}{2^{2N}} \sum_{\epsilon_i = \pm 1, \sum \epsilon_i = 0} \prod_{i < j} |z_i - z_j|^{\frac{\epsilon_i \epsilon_j}{2}} \quad (45)$$

where Pf denotes the matrix Pfaffian. Despite their compact presentation, these are rather complicated functions; e.g. combinatorially the Pfaffian can be viewed as a sum over all possible perfect matchings of N points. The number of these matchings grows factorially in N , and each of them contributes a (different) product of 2-body interactions. There are more efficient ways than this to actually *compute* the Pfaffian, but our neural networks will be unable to realize that structure.

C.2 3D ISING MODEL

The lattice construction equation 40 may easily be generalized to 3d. Unlike the 2d Ising model, the 3d Ising model has not been explicitly solved, and there are no simple exact formulas such as equation 45 for moments of the spin operator at the critical point. Nevertheless, a great deal is known about it, where the most precise information comes from the conformal bootstrap (El-Showk et al., 2012), i.e. a method which exploits the self-consistency of conformal field theory to solve the system. For reviews see (Simmons-Duffin, 2017; Poland et al., 2019).

Using these ideas, the 4-point function of the spins is given by

$$\langle \sigma(x_1)\sigma(x_2)\sigma(x_3)\sigma(x_4) \rangle = \frac{g(u, v)}{|x_1 - x_2|^{2\Delta_\sigma} |x_3 - x_4|^{2\Delta_\sigma}} \quad (46)$$

where the scaling dimension of the spin operator is $\Delta_\sigma = 0.5181489(10)$. u, v are conformally invariant functions of the four insertion points called *cross-ratios*, and are

$$u \equiv \frac{x_{12}^2 x_{34}^2}{x_{13}^2 x_{24}^2} \quad v \equiv \frac{x_{23}^2 x_{14}^2}{x_{13}^2 x_{24}^2} \quad x_{ij} \equiv x_i - x_j \quad (47)$$

The non-trivial information in this 4-point function is stored in the dependence on the cross-ratios $g(u, v)$. We compute this following (Rychkov et al., 2017), which we review briefly below.

This function $g(u, v)$ can be decomposed in the following fashion:

$$g(u, v) = \sum_{\mathbb{O} \in \sigma \times \sigma} C_{\sigma\mathbb{O}}^2 g_{\Delta_{\mathbb{O}}, l_{\mathbb{O}}}(u, v) \quad (48)$$

where the sum over \mathbb{O} runs over the conformal primaries of the theory, $\Delta_{\mathbb{O}}$ and $l_{\mathbb{O}}$ are the conformal dimension and spin of the operators respectively, and $g_{\Delta_{\mathbb{O}}, l_{\mathbb{O}}}(u, v)$ is a theory-independent function

called a *conformal block*. which These functions are the same for all theories, and thus the information characterizing this function is encoded in the set $\{\Delta_{\mathbb{O}}, l_{\mathbb{O}}, C_{\sigma\sigma\mathbb{O}}\}$, where $C_{\sigma\sigma\mathbb{O}}$ is a set of theory-specific numbers called *operator product expansion coefficients*. For the 3d Ising model this set of numbers is not known analytically, but they have been computed for the operators with lowest dimension to extremely high accuracy.

Following the method in (Rychkov et al., 2017), and using the lowest five operators and their associated $\{\Delta_{\mathbb{O}}, l_{\mathbb{O}}, C_{\sigma\sigma\mathbb{O}}\}$ from (Komargodski & Simmons-Duffin, 2017), we compute $g(u, v)$ using the Mathematica code of (Costa et al., 2016) to find the conformal blocks. We then create a dataset of 4-point functions precisely as for the 2d Ising model above. This is not an exact solution to the 3d Ising model, but the error from truncating the operator sum has been conservatively estimated to be one part in 10^{-2} in (Rychkov et al., 2017). (We also note that for benchmarking our results presumably *any* choice of $g(u, v)$ would suffice, but we think it is more physically reasonable to use the function that is relevant to a known physical problem).

Such methods would be much harder to implement for higher-point functions, and thus our test dataset for the 3d Ising model is restricted to the 4-point function.

D THE CONFORMAL GROUP OF THE PSEUDO-EUCLIDEAN SPACE

In this section we will provide a self-contained introduction to conformal geometry of the (pseudo-) Euclidean space in arbitrary signature and its conformal group. We provide definitions, examples and main theorems.

We follow the references of Schottenloher (2008) and McKay (2023).

D.1 CONFORMAL TRANSFORMATIONS OF THE PSEUDO-EUCLIDEAN SPACE

Roughly speaking, a conformal map is a map that preserves local angles. However, this definition is rather tedious to write down and work with. A technically more convenient definition is introduced below in Definition D.1.2. To motivate this definition, we first show in Lemma D.1.1 that in the (linear) Euclidean case these two definitions are equivalent:

Lemma D.1.1. *Let $A \in \mathbb{R}^{d \times d}$ be a real invertible $(d \times d)$ -matrix. Then the following statements are equivalent:*

1. *A is angle preserving, i.e. for all $v_1, v_2 \in \mathbb{R}^d \setminus \{0\}$ we have:*

$$\frac{\langle Av_1, Av_2 \rangle}{\|Av_1\| \cdot \|Av_2\|} = \frac{\langle v_1, v_2 \rangle}{\|v_1\| \cdot \|v_2\|}. \quad (49)$$

2. *A is a conformal matrix, i.e. $A = c\Lambda$ with a scalar $c > 0$ and an orthogonal matrix $\Lambda \in O(d)$.*

Proof. “ \Leftarrow ”: Let $A = c\Lambda$ with $c > 0$ and $\Lambda \in O(d)$. Then we get for all $v_1, v_2 \in \mathbb{R}^d$:

$$\langle Av_1, Av_2 \rangle = v_1^\top A^\top Av_2 \quad (50)$$

$$= c^2 \cdot v_1^\top \underbrace{\Lambda^\top \Lambda}_{=I} v_2 \quad (51)$$

$$= c^2 \cdot v_1^\top v_2 \quad (52)$$

$$= c^2 \cdot \langle v_1, v_2 \rangle. \quad (53)$$

This also implies:

$$\|Av_1\| = c \cdot \|v_1\|, \quad \|Av_2\| = c \cdot \|v_2\|. \quad (54)$$

Together this implies the claim:

$$\frac{\langle Av_1, Av_2 \rangle}{\|Av_1\| \cdot \|Av_2\|} = \frac{\langle v_1, v_2 \rangle}{\|v_1\| \cdot \|v_2\|}. \quad (55)$$

“ \implies ”: Assume that A is invertible and preserves angles. Then $C := A^\top A$ is symmetric and positive definite. By the spectral theorem we can diagonalize C , i.e. there exists an orthonormal basis $e_1, \dots, e_d \in \mathbb{R}^d$, $\langle e_i, e_j \rangle = \delta_{i,j}$ for all $i, j \in [d]$, and positive scalars $\lambda_1, \dots, \lambda_d > 0$ such that for all $i \in [d]$:

$$C e_i = \lambda_i \cdot e_i. \quad (56)$$

We now claim that: $\lambda_1 = \dots = \lambda_d$. By way of contradiction assume that there exists $\ell \in [d]$ such that $\lambda_1 \neq \lambda_\ell$. Then put: $v_1 := e_1$ and $v_2 := e_1 + e_\ell$. This gives:

$$\|v_1\|^2 = 1, \quad \|v_2\|^2 = \|e_1\|^2 + \|e_\ell\|^2 = 2, \quad (57)$$

$$\|A v_1\|^2 = \lambda_1, \quad A^\top A v_2 = \lambda_1 \cdot e_1 + \lambda_\ell \cdot e_\ell, \quad (58)$$

$$\langle A v_1, A v_2 \rangle = \lambda_1, \quad \|A v_2\|^2 = \lambda_1 + \lambda_\ell, \quad (59)$$

$$\langle v_1, v_2 \rangle = 1. \quad (60)$$

This implies:

$$\frac{\lambda_1}{\sqrt{\lambda_1} \cdot \sqrt{\lambda_1 + \lambda_\ell}} = \frac{\langle A v_1, A v_2 \rangle}{\|A v_1\| \cdot \|A v_2\|} = \frac{\langle v_1, v_2 \rangle}{\|v_1\| \cdot \|v_2\|} = \frac{1}{1 \cdot \sqrt{2}}. \quad (61)$$

Squaring and solving for λ_ℓ shows:

$$\lambda_\ell = \lambda_1, \quad (62)$$

which contradicts the assumption. So, we indeed have that:

$$c^2 := \lambda_1 = \dots = \lambda_d > 0. \quad (63)$$

This shows that: $C = c^2 I$. Putting $\Lambda := \frac{1}{c} A$ we get:

$$\Lambda^\top \Lambda = \frac{1}{c^2} C = I, \quad (64)$$

which shows that $\Lambda \in O(d)$ and thus $A = c\Lambda$ with $c > 0$ and $\Lambda \in O(d)$. \square

Definition D.1.2 (Conformal maps/conformal transformations). *Let (M, η^M) and (N, η^N) be two pseudo-Riemannian manifolds. A conformal map $f : M \rightarrow N$ is defined to be a smooth map such that there exists a smooth map $\omega : M \rightarrow \mathbb{R}_{>0}$ such that for all $x \in M$ and $v_1, v_2 \in T_x M$ we have:*

$$\eta_{f(x)}^N(df_x(v_1), df_x(v_2)) = \omega(x)^2 \cdot \eta_x^M(v_1, v_2), \quad (65)$$

where $df_x : T_x M \rightarrow T_{f(x)} N$ is the differential of f at $x \in M$. The above map ω is called the conformal factor of f .

In case the conformal factor ω of f equals the constant one, $\omega(x) = 1$, then we call f an isometric map or isometric transformation.

Definition D.1.3 (Conformal diffeomorphisms and isometries). *Let (M, η^M) be a pseudo-Riemannian manifold. A conformal diffeomorphism $f : M \rightarrow M$ is a conformal smooth map that has a conformal smooth inverse $f^{-1} : M \rightarrow M$:*

$$f \circ f^{-1} = f^{-1} \circ f = \text{id}_M. \quad (66)$$

If, in addition, its conformal factor ω of f is the constant 1, then we call f an isometry of (M, η^M) . The group of conformal diffeomorphisms of (M, η^M) is denoted as:

$$\text{ConfDiff}(M, \eta^M) := \{f : M \rightarrow M \text{ conformal map with conformal inverse}\}. \quad (67)$$

The isometry group of (M, η^M) is denoted as:

$$\text{Isom}(M, \eta^M) := \{f : M \rightarrow M \text{ isometric map with isometric inverse}\}. \quad (68)$$

Notation D.1.4. In the following we will denote by:

1. $(\mathbb{R}^{p,q}, \eta^{p,q})$ be the standard pseudo-Euclidean space of signature (p, q) :

$$\eta^{p,q}(v_1, v_2) := v_1^\top \Delta^{p,q} v_2, \quad \Delta^{p,q} := \text{diag}(\underbrace{+1, \dots, +1}_{\times p}, \underbrace{-1, \dots, -1}_{\times q}), \quad (69)$$

2. $O(p, q)$ the (pseudo-)orthogonal group of signature (p, q) :

$$O(p, q) := \{ \Lambda \in \text{GL}(\mathbb{R}^{p,q}) \mid \Lambda^\top \Delta^{p,q} \Lambda = \Delta^{p,q} \}, \quad (70)$$

3. $SO(p, q)$ the special (pseudo-)orthogonal group of signature (p, q) :

$$SO(p, q) := \{ \Lambda \in O(p, q) \mid \det \Lambda = 1 \}. \quad (71)$$

4. $O^0(p, q) = SO^0(p, q)$ the identity component (=connected component of the identity) of $O(p, q)$. Note that we have the inclusions of groups:

$$O^0(p, q) \subseteq SO(p, q) \subseteq O(p, q). \quad (72)$$

Lemma D.1.5 (The connected components of $O(p, q)$). *Let $p, q \geq 0$. Then we have the following:*

1. If $p = q = 0$ then $O(p, q) = \{0\}$ has only one connected component.

2. If $p + q \geq 1$ and $pq = 0$ then $O(p, q)$ has 2 connected components:

$$SO(p, q) = \{ \Lambda \in O(p, q) \mid \det \Lambda = 1 \} = O^0(p, q), \quad \{ \Lambda \in O(p, q) \mid \det \Lambda = -1 \}. \quad (73)$$

3. If $p, q \geq 1$ then $O(p, q)$ has 4 connected components:

$$O^{++}(p, q) := \left\{ \Lambda = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in O(p, q) \mid \det A \geq +1, \det D \geq +1 \right\} = O^0(p, q), \quad (74)$$

$$O^{+-}(p, q) := \left\{ \Lambda = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in O(p, q) \mid \det A \geq +1, \det D \leq -1 \right\}, \quad (75)$$

$$O^{-+}(p, q) := \left\{ \Lambda = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in O(p, q) \mid \det A \leq -1, \det D \geq +1 \right\}, \quad (76)$$

$$O^{--}(p, q) := \left\{ \Lambda = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in O(p, q) \mid \det A \leq -1, \det D \leq -1 \right\}. \quad (77)$$

We have the following open-and-closed subgroups of $O(p, q)$:

$$O^+(p, q) := \left\{ \Lambda = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in O(p, q) \mid \det A \geq +1 \right\} \quad (78)$$

$$= O^{++}(p, q) \dot{\cup} O^{+-}(p, q), \quad (79)$$

$$O^-(p, q) := \left\{ \Lambda = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in O(p, q) \mid \det D \geq +1 \right\} \quad (80)$$

$$= O^{++}(p, q) \dot{\cup} O^{-+}(p, q), \quad (81)$$

$$SO(p, q) = \{ \Lambda \in O(p, q) \mid \det \Lambda = 1 \} \quad (82)$$

$$= O^{++}(p, q) \dot{\cup} O^{--}(p, q), \quad (83)$$

$$O^0(p, q) = O^+(p, q) \cap O^-(p, q) \quad (84)$$

$$= O^+(p, q) \cap SO(p, q) \quad (85)$$

$$= O^-(p, q) \cap SO(p, q). \quad (86)$$

In particular, if $p + q \geq 1$ and p or q is an odd number then $-I \notin O^0(p, q)$.

Lemma D.1.6 (Polar representation). *Let $p, q \geq 1$. Then every $\Lambda \in O(p, q)$ can uniquely be written as:*

$$\Lambda = \begin{bmatrix} U & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} \sqrt{I_p + BB^\top} & B \\ B^\top & \sqrt{I_q + B^\top B} \end{bmatrix}, \quad (87)$$

with unique $U \in O(p)$, $V \in O(q)$, $B \in \mathbb{R}^{p \times q}$. Conversely, every such product of matrices lies in $O(p, q)$.

Example D.1.7 (Affine conformal diffeomorphisms). *Let $b \in \mathbb{R}^{p,q}$, $c > 0$ and $\Lambda \in O(p, q)$. Consider the affine map:*

$$f : \mathbb{R}^{p,q} \rightarrow \mathbb{R}^{p,q}, \quad f(x) := c\Lambda x + b. \quad (88)$$

Then f is a conformal diffeomorphism of $\mathbb{R}^{p,q}$ with constant conformal factor $\omega(x) = c$.

Proof. The differential at x is given by the matrix:

$$df_x = c\Lambda. \quad (89)$$

So for $v_1, v_2 \in T_x \mathbb{R}^{p,q} = \mathbb{R}^{p,q}$ we get:

$$\eta^{p,q}(df_x(v_1), df_x(v_2)) = (c\Lambda v_1)^\top \Delta^{p,q}(c\Lambda v_2) \quad (90)$$

$$= c^2 \cdot v_1^\top \Lambda^\top \Delta^{p,q} \Lambda v_2 \quad (91)$$

$$= c^2 \cdot v_1^\top \Delta^{p,q} v_2 \quad (92)$$

$$= c^2 \cdot \eta^{p,q}(v_1, v_2). \quad (93)$$

So, if we then define the conformal factor $\omega : \mathbb{R}^{p,q} \rightarrow \mathbb{R}_{>0}$ to be the constant map $\omega(x) := c$, we have shown that f is a conformal smooth map. Its inverse f^{-1} is given by:

$$f^{-1}(x) := c^{-1}\Lambda^{-1}x - c^{-1}\Lambda^{-1}b, \quad (94)$$

which is thus of the same form as f and thus also a conformal smooth map. This shows the claim. \square

Theorem D.1.8 (Affine conformal diffeomorphisms, see Amir-Moéz (1967)). *Consider the affine map on $\mathbb{R}^{p,q}$:*

$$f : \mathbb{R}^{p,q} \rightarrow \mathbb{R}^{p,q}, \quad f(x) := Ax + b, \quad (95)$$

with a square matrix A and translation vector $b \in \mathbb{R}^{p,q}$. Then f is a conformal map (w.r.t. $\eta^{p,q}$) iff there exists a $c \in \mathbb{R}_{>0}$ and a $\Lambda \in O(p, q)$ such that $A = c\Lambda$. If this is the case, f is a conformal diffeomorphism and both $c > 0$ and $\Lambda \in O(p, q)$ are uniquely determined by A as: $c = \sqrt[d]{|\det A|}$ with $d := p + q$, and, $\Lambda = \frac{1}{c}A$.

Proof. One direction is proven in Example D.1.7. For the other direction assume that f is a conformal map. Then we have $df_x = A$ and the conformal relation (in matrix form):

$$A^\top \Delta^{p,q} A = \omega(x)^2 \cdot \Delta^{p,q}. \quad (96)$$

Taking determinants on both sides gives:

$$(\det A)^2 = (\omega(x)^2)^d > 0, \quad (97)$$

showing that A is invertible and that $\omega(x) = \sqrt[d]{|\det A|} =: c$ is not dependent on x and thus equal to the constant $c > 0$. Dividing the first equation on both sides by c^2 and rearranging gives:

$$\left(\frac{1}{c}A\right)^\top \Delta^{p,q} \left(\frac{1}{c}A\right) = \Delta^{p,q}. \quad (98)$$

This shows that $\Lambda := \frac{1}{c}A \in O(p, q)$ and thus the claim: $A = c\Lambda$. The rest follows from Example D.1.7. \square

Definition D.1.9 (The linear and affine conformal group). *We define the linear and affine conformal group, resp., of signature (p, q) as follows:*

$$\text{CO}(p, q) := \mathbb{R}_{>0} \times O(p, q), \quad \text{CE}(p, q) := \text{CO}(p, q) \times \mathbb{R}^{p,q}. \quad (99)$$

Note that the entries correspond to scaling factor $c > 0$, reflection-rotation matrix $\Lambda \in O(p, q)$ and translation vector $b \in \mathbb{R}^{p,q}$ in the conformal affine maps from Example D.1.7 and Theorem D.1.8.

We also define their identity components:

$$\text{CO}^0(p, q) := \mathbb{R}_{>0} \times O^0(p, q), \quad \text{CE}^0(p, q) := \text{CO}^0(p, q) \times \mathbb{R}^{p,q}. \quad (100)$$

Example D.1.10 (The inversion at the pseudo-sphere). *The following (partial) map:*

$$\varsigma = \varsigma^{p,q} : \mathbb{R}^{p,q} \rightarrow \mathbb{R}^{p,q}, \quad \varsigma^{p,q}(x) := \frac{x}{\eta^{p,q}(x,x)}, \quad (101)$$

is called the inversion (at the pseudo-sphere) of signature (p, q) . Note that the inversion here is only defined for $x \in U := \{\tilde{x} \in \mathbb{R}^{p,q} \mid \eta^{p,q}(\tilde{x}, \tilde{x}) \neq 0\}$. It is an involution (self-invers) on U as (with $\eta := \eta^{p,q}$):

$$\eta(\varsigma(x), \varsigma(x)) = \frac{\eta(x,x)}{\eta(x,x)^2} = \frac{1}{\eta(x,x)} \neq 0, \quad (102)$$

$$\varsigma(\varsigma(x)) = \frac{\varsigma(x)}{\eta(\varsigma(x), \varsigma(x))} = \frac{\frac{x}{\eta(x,x)}}{\frac{1}{\eta(x,x)}} = x. \quad (103)$$

We now claim that $\varsigma : U \rightarrow U$ is a conformal diffeomorphism. For this we compute its Jacobian matrix (differential), which, in matrix form, is given as:

$$d\varsigma_x = \frac{1}{\eta(x,x)^2} (\eta(x,x)I - 2xx^\top \Delta^{p,q}). \quad (104)$$

To check the conformal relation we thus compute:

$$(d\varsigma_x)^\top \Delta^{p,q} (d\varsigma_x) = \frac{1}{\eta(x,x)^4} \left((\eta(x,x)I - 2xx^\top \Delta^{p,q})^\top \Delta^{p,q} (\eta(x,x)I - 2xx^\top \Delta^{p,q}) \right) \quad (105)$$

$$= \frac{1}{\eta(x,x)^4} \left((\eta(x,x)I - 2\Delta^{p,q}xx^\top) \Delta^{p,q} (\eta(x,x)I - 2xx^\top \Delta^{p,q}) \right) \quad (106)$$

$$= \frac{1}{\eta(x,x)^4} \left(\eta(x,x)^2 \Delta^{p,q} - 4\eta(x,x) \Delta^{p,q} xx^\top \Delta^{p,q} + 4\Delta^{p,q} xx^\top \Delta^{p,q} xx^\top \Delta^{p,q} \right) \quad (107)$$

$$= \frac{1}{\eta(x,x)^4} \left(\eta(x,x)^2 \Delta^{p,q} - 4\eta(x,x) \Delta^{p,q} xx^\top \Delta^{p,q} + 4\eta(x,x) \Delta^{p,q} xx^\top \Delta^{p,q} \right) \quad (108)$$

$$= \frac{1}{\eta(x,x)^2} \Delta^{p,q}. \quad (109)$$

This shows that the inversion $\varsigma : U \rightarrow U$ is a conformal diffeomorphism with conformal factor $\omega(x) = \frac{1}{|\eta(x,x)|}$. We continue the analysis in Example D.4.5.

Example D.1.11 (Special conformal transformations). *For $b \in \mathbb{R}^{p,q}$ we define the (partial) map:*

$$\sigma_b : \mathbb{R}^{p,q} \rightarrow \mathbb{R}^{p,q}, \quad \sigma_b(x) := \frac{x - \eta(x,x) \cdot b}{1 - 2 \cdot \eta(x,b) + \eta(b,b) \cdot \eta(x,x)} = \varsigma(\varsigma(x) - b), \quad (110)$$

where $\varsigma = \varsigma^{p,q} : \mathbb{R}^{p,q} \rightarrow \mathbb{R}^{p,q}$ denotes the inversion from Example D.1.10 of signature (p, q) . Maps of the form σ_b are called special conformal transformations of $\mathbb{R}^{p,q}$ and are conformal maps on their domain of definition:

$$U_b := \{x \in \mathbb{R}^{p,q} \mid \nu(x,b) \neq 0\}, \quad (111)$$

with conformal factor $\omega_b(x) := \frac{1}{|\nu(x,b)|}$, where we abbreviated the above denominator as:

$$\nu(x,b) := 1 - 2 \cdot \eta(x,b) + \eta(b,b) \cdot \eta(x,x). \quad (112)$$

Furthermore, $\sigma_{-b} : U_{-b} \rightarrow U_b$ is the inverse of $\sigma_b : U_b \rightarrow U_{-b}$.

Proof. We first check that $y := \sigma_b(x)$ lies in U_{-b} for $x \in U_b$. For this we compute:

$$\eta(x - \eta(x,x)b, x - \eta(x,x)b) = \eta(x,x) - 2\eta(x,x)\eta(x,b) + \eta(x,x)^2\eta(b,b) \quad (113)$$

$$= \eta(x,x)\nu(x,b), \quad (114)$$

$$\eta(x - \eta(x,x)b, -b) = -\eta(x,b) + \eta(x,x)\eta(b,b) \quad (115)$$

Dividing both sides by $\nu(x, b)^2$ and $\nu(x, b)$, resp., shows:

$$\eta(y, y) = \frac{\eta(x, x)}{\nu(x, b)}, \quad \eta(y, -b) = \frac{1}{\nu(x, b)} (-\eta(x, b) + \eta(x, x)\eta(b, b)). \quad (116)$$

With this we get:

$$\nu(y, -b) = 1 - 2 \cdot \eta(y, -b) + \eta(b, b) \cdot \eta(y, y) \quad (117)$$

$$= \frac{1}{\nu(x, b)} (\nu(x, b) + 2 \cdot \eta(x, b) - 2 \cdot \eta(x, x) \cdot \eta(b, b) + \eta(b, b) \cdot \eta(x, x)) \quad (118)$$

$$= \frac{1}{\nu(x, b)} (\nu(x, b) + 2 \cdot \eta(x, b) - \eta(b, b) \cdot \eta(x, x)) \quad (119)$$

$$= \frac{1}{\nu(x, b)} \quad (120)$$

$$\neq 0. \quad (121)$$

This shows: $y \in U_{-b}$ for $x \in U_b$. Plugging the relation $\nu(y, -b) = \frac{1}{\nu(x, b)}$ into Equation (116) shows that:

$$\sigma_{-b}(y) = \frac{y + \eta(y, y) \cdot b}{\nu(y, -b)} \quad (122)$$

$$= \frac{1}{\nu(y, -b)} \cdot y + \frac{\eta(y, y)}{\nu(y, -b)} \cdot b \quad (123)$$

$$= \nu(x, b) \cdot \left(\frac{x - \eta(x, x) \cdot b}{\nu(x, b)} \right) + \eta(x, x) \cdot b \quad (124)$$

$$= x. \quad (125)$$

This shows that $\sigma_{-b} : U_{-b} \rightarrow U_b$ is the inverse of $\sigma_b : U_b \rightarrow U_{-b}$.

To show that $\sigma_b : U_b \rightarrow \mathbb{R}^{p,q}$ is a conformal map we need to compute its differential. Computing the differential directly gives:

$$(d\sigma_b)_x = \frac{\nu(x, b) \cdot (I - 2 \cdot b \cdot x^\top \Delta^{p,q}) - (x - \eta(x, x) \cdot b) \cdot (-2 \cdot b^\top \Delta^{p,q} + 2 \cdot \eta(b, b) \cdot x^\top \Delta^{p,q})}{\nu(x, b)^2}, \quad (126)$$

which is difficult to work with. Instead, we use the fact that σ_b is a smooth extension of the map $\sigma_b(x) = \varsigma(\varsigma(x) - b)$. With the chain rule we then get:

$$(d\sigma_b)_x = (d\varsigma)_{(\varsigma(x)-b)} d\varsigma_x \quad (127)$$

With this and Example D.1.10 we get:

$$(d\sigma_b)_x^\top \Delta^{p,q} (d\sigma_b)_x = (d\varsigma_x)^\top (d\varsigma)_{(\varsigma(x)-b)}^\top \Delta^{p,q} (d\varsigma)_{(\varsigma(x)-b)} d\varsigma_x \quad (128)$$

$$= \frac{1}{\eta(\varsigma(x) - b, \varsigma(x) - b)^2} \cdot (d\varsigma_x)^\top \Delta^{p,q} d\varsigma_x \quad (129)$$

$$= \frac{1}{\eta(x, x)^2} \cdot \frac{1}{\eta(\varsigma(x) - b, \varsigma(x) - b)^2} \Delta^{p,q} \quad (130)$$

$$= \frac{1}{\nu(x, b)^2} \Delta^{p,q}. \quad (131)$$

For the last step note that:

$$\eta(\varsigma(x) - b, \varsigma(x) - b) = \frac{\nu(x, b)}{\eta(x, x)}. \quad (132)$$

So the conformal factor of σ_b at $x \in U_b$ is $\omega_b(x) := \frac{1}{|\nu(x, b)|}$. We continue the analysis in Example D.4.9. \square

Remark D.1.12 (Extending conformal maps to a conformal compactification). *The examples D.1.7, D.1.8, D.1.10, D.1.11 have given us several conformal transformations of $\mathbb{R}^{p,q}$. However, for us to be able to define the inversion $\varsigma : U \rightarrow \mathbb{R}^{p,q}$ and the special conformal transformations $\sigma_b : U_b \rightarrow \mathbb{R}^{p,q}$ we had to restrict their domain of definition to an open subset of $\mathbb{R}^{p,q}$. The question is now if we can actually extend those maps like ς and σ_b to a conformal map on the whole of $\mathbb{R}^{p,q}$ by possibly enlarging their codomain, e.g. by putting:*

$$\varsigma : \mathbb{R}^{p,q} \rightarrow \mathbb{R}^{p,q} \cup \{\infty\}, \quad \varsigma(x) := \begin{cases} \frac{x}{\eta(x,x)}, & \text{if } \eta(x,x) \neq 0, \\ \infty, & \text{if } \eta(x,x) = 0? \end{cases} \quad (133)$$

However, for this we would need to turn the extended codomain $\mathbb{R}^{p,q} \cup \{\infty\}$ into a proper pseudo-Riemannian manifold $\mathbb{M}^{p,q}$. Furthermore, we then would also like to properly define ς on the added new points $\mathbb{M}^{p,q} \setminus \mathbb{R}^{p,q}$ and turn it into a conformal diffeomorphism of $\mathbb{M}^{p,q}$. Note that in the non-Euclidean case, i.e. if $p, q > 0$, the space $\{x \in \mathbb{R}^{p,q} \mid \eta(x,x) = 0\}$ consists of more than one point and thus needs more consideration.

The above is the main question (of the existence) of a conformal compactification of $\mathbb{R}^{p,q}$. Below we follow an ad hoc approach to this question.

Remark D.1.13. *Even after the question of conformal compactification, as described in Remark D.1.12, is solved, we are still faced with the ambiguity of how to define the conformal group of $\mathbb{R}^{p,q}$. There are several non-equivalent options:*

1. *as the group $\text{ConfDiff}(\mathbb{R}^{p,q})$ of all conformal diffeomorphisms of $\mathbb{R}^{p,q}$, which would include all affine conformal transformations, but exclude the inversion and the special conformal transformations;*
2. *as the group $\text{ConfDiff}(\mathbb{M}^{p,q})$ of all conformal diffeomorphisms of a conformal compactification $\mathbb{M}^{p,q}$ of $\mathbb{R}^{p,q}$, which would include all affine and special conformal transformations and the inversion. However, one can show that for $p + q = 2$ some of the properties of the conformal compactification will break down and that $\text{ConfDiff}(\mathbb{M}^{p,q})$ can become pathologically big in the case $(p, q) = (1, 1)$, see Schottenloher (2008);*
3. *as the connected component of the identity of one of the above groups, in order to stick to the main conformal diffeomorphisms that have descriptions with help of a Lie algebra and Lie exponential map, etc.;*
4. *as the subgroup of (partially defined) conformal diffeomorphisms of $\mathbb{R}^{p,q}$ that is generated by all affine and special conformal transformations (excluding the inversion) of $\mathbb{R}^{p,q}$;*
5. *as the subgroup of conformal diffeomorphisms of $\mathbb{M}^{p,q}$ that is generated by all affine and special conformal transformations (excluding the inversion) of $\mathbb{R}^{p,q}$;*

For this draft we will settle with one of the last definitions, which are mostly equivalent, and call it the (restricted) conformal group of $\mathbb{R}^{p,q}$ and denote it by $\text{Conf}(\mathbb{R}^{p,q})$. For more details see further below.

To define the conformal compactification of $\mathbb{R}^{p,q}$ we first need to introduce the bigger pseudo-Euclidean space $\mathbb{R}^{p+1, q+1}$ and the projective space \mathbb{P}^{p+q+1} .

D.2 THE PROJECTIVE SPACE, DE SITTER SPACE AND ANTI-DE SITTER SPACE

Definition D.2.1 (The projective space). *Let V be a (real) finite dimensional vector space. We then define the projective space associated with V as the space of equivalence classes:*

$$\mathbb{P}(V) := (V \setminus \{0\}) / \sim, \quad (134)$$

where we define two non-zero vectors $x_1, x_2 \in V$ to be equivalent if they lie on the same straight line through the origin:

$$x_1 \sim x_2 : \iff \exists c \in \mathbb{R} \setminus \{0\}. x_1 = c \cdot x_2. \quad (135)$$

We then define the standard projective space of dimensions $d \geq 0$ as:

$$\mathbb{P}^d := \mathbb{P}(\mathbb{R}^{d+1}). \quad (136)$$

Note that d denotes the dimension of \mathbb{P}^d , which thus requires the Euclidean space \mathbb{R}^{d+1} to be of one dimension higher.

Remark D.2.2. *W.r.t. to our pseudo-Euclidean space $\mathbb{R}^{p,q}$ we will in the following mostly consider the projective space of dimension $d + 1$, where $d := p + q$:*

$$\mathbb{P}^{d+1} := \mathbb{P}(\mathbb{R}^{p+1,q+1}). \quad (137)$$

Elements of \mathbb{P}^{d+1} will be either denoted by $[z]$ with $z \in \mathbb{R}^{p+1,q+1}$ with coordinates $z = (z^0, z^1, \dots, z^d, z^{d+1})$ or directly as $[z] = [z^0 : z^1 : \dots : z^d : z^{d+1}]$. Note that we endow $\mathbb{R}^{p+1,q+1}$ with the standard metric $\eta^{p+1,q+1}$ of the pseudo-Euclidean space of dimension $d + 2$ and signature $(p + 1, q + 1)$ and that we consider the coordinates z^0, \dots, z^p to belong to the $(+1)$ -signature and the coordinates z^{p+1}, \dots, z^{d+1} to belong to the (-1) -signature.

Definition D.2.3 (de Sitter and anti-de Sitter space). *Let (V, η) be a non-degenerate finite dimensional real quadratic vector space. Then we define the de Sitter space and the anti-de Sitter space and the (projective) zero quadric associated to (V, η) as:*

$$\text{dS}(V, \eta) := \{[z] \in \mathbb{P}(V) \mid \eta(z, z) > 0\}, \quad (138)$$

$$\text{AdS}(V, \eta) := \{[z] \in \mathbb{P}(V) \mid \eta(z, z) < 0\}. \quad (139)$$

$$\mathbb{M}(V, \eta) := \{[z] \in \mathbb{P}(V) \mid \eta(z, z) = 0\}. \quad (140)$$

Note that these are all well-defined subsets of $\mathbb{P}(V)$. The pseudo-Riemannian metrics on the first two spaces are induced by the following subspaces of V :

$$Y_+ := \{y \in V \mid \eta(y, y) = 1\}, \quad Y_- := \{y \in V \mid \eta(y, y) = -1\}, \quad (141)$$

and the 2 : 1 locally diffeomorphic surjective maps:

$$\pi_+ : Y_+ \rightarrow \text{dS}(V, \eta) \subseteq \mathbb{P}(V), \quad \pm y \mapsto [y], \quad (142)$$

$$\pi_- : Y_- \rightarrow \text{AdS}(V, \eta) \subseteq \mathbb{P}(V), \quad \pm y \mapsto [y], \quad (143)$$

where both Y_+ and Y_- are endowed with the pull-back metrics η^+ and η^- from η of V to those subspaces. More explicitly, the tangent spaces and local metrics are given as follows:

$$\mathbb{T}_{[z]} \text{dS}(V, \eta) = \{v \in V \mid \eta(z, v) = 0\}, \quad \eta_{[z]}^{\text{dS}(V, \eta)}(v_1, v_2) = \eta(v_1, v_2), \quad (144)$$

$$\mathbb{T}_{[z]} \text{AdS}(V, \eta) = \{v \in V \mid \eta(z, v) = 0\}, \quad \eta_{[z]}^{\text{AdS}(V, \eta)}(v_1, v_2) = \eta(v_1, v_2). \quad (145)$$

We define the (standard) de Sitter space and the (standard) anti-de Sitter space and (standard) zero quadric of signatures (p, q) , resp., as follows:

$$\text{dS}^{p,q} := \text{dS}(\mathbb{R}^{p+1,q}, \eta^{p+1,q}), \quad (146)$$

$$\text{AdS}^{p,q} := \text{AdS}(\mathbb{R}^{p,q+1}, \eta^{p,q+1}), \quad (147)$$

$$\mathbb{M}^{p,q} := \mathbb{M}(\mathbb{R}^{p+1,q+1}, \eta^{p+1,q+1}). \quad (148)$$

The pseudo-Riemannian metric on $\mathbb{M}^{p,q}$ will be defined later in Definition D.3.5.

Definition D.2.4 (Projective orthogonal group). *Let (V, η) be a non-degenerate finite dimensional real quadratic vector space. Then the projective orthogonal group of (V, η) is defined to be:*

$$\text{PO}(V, \eta) := \text{O}(V, \eta) / \{\pm \text{id}_V\}. \quad (149)$$

The (standard) projective orthogonal group of signature (p, q) is:

$$\text{PO}(p, q) := \text{O}(p, q) / \{\pm I\}. \quad (150)$$

We denote the identity component of those groups as:

$$\text{PO}^0(V, \eta), \quad \text{and} \quad \text{PO}^0(p, q). \quad (151)$$

Lemma D.2.5. *Let (V, η) be a non-degenerate finite dimensional real quadratic vector space. Then the projective orthogonal group $\text{PO}(V, \eta)$ acts on the spaces $\mathbb{P}(V, \eta)$, $\text{dS}(V, \eta)$, $\text{AdS}(V, \eta)$ and on $\mathbb{M}(V, \eta)$ via matrix multiplication:*

$$[\Lambda][z] := [\Lambda z], \quad (152)$$

in a well-defined way.

Theorem D.2.6 (See McKay (2023) Thm. 7.5). *Let (V, η) be a non-degenerate finite dimensional real quadratic vector space. Then the standard action of $\text{PO}(V, \eta)$ on $\text{dS}(V, \eta)$ and $\text{AdS}(V, \eta)$ acts isometrically and we get the following identification with their isometry groups:*

$$\text{Isom}(\text{dS}(V, \eta)) \cong \text{PO}(V, \eta), \quad \text{Isom}(\text{AdS}(V, \eta)) \cong \text{PO}(V, \eta). \quad (153)$$

Remark D.2.7. *For our pseudo-Euclidean space $\mathbb{R}^{p,q}$ we will use the projective zero quadric $\mathbb{M}^{p,q}$ as its conformal compactification, see later in Definition D.3.5. For this we need to consider the bigger pseudo-Euclidean vector space $(\mathbb{R}^{p+1,q+1}, \eta^{p+1,q+1})$, and, the projective space $\mathbb{P}^{d+1} = \mathbb{P}(\mathbb{R}^{p+1,q+1})$. Here we get the disjoint decomposition of \mathbb{P}^{d+1} into anti-de Sitter space, conformal compactification and de Sitter space:*

$$\mathbb{P}^{d+1} = \text{AdS}^{p+1,q} \dot{\cup} \mathbb{M}^{p,q} \dot{\cup} \text{dS}^{p,q+1}. \quad (154)$$

Note the different signatures on each of the spaces. With Theorem D.2.6 we thus get the identifications:

$$\text{Isom}(\text{dS}^{p,q+1}) = \text{PO}(p+1, q+1) = \text{Isom}(\text{AdS}^{p+1,q}). \quad (155)$$

In the following, we further want to show that the action of $\text{PO}(p+1, q+1)$ on $\mathbb{M}^{p,q}$ induces conformal diffeomorphisms of $\mathbb{M}^{p,q}$. This identification via $\text{PO}(p+1, q+1)$ will thus induce a correspondence between (certain) conformal diffeomorphisms of the conformal compactification $\mathbb{M}^{p,q}$ of $\mathbb{R}^{p,q}$ and the isometries of the corresponding $\text{AdS}^{p+1,q}$ -space (or, also the $\text{dS}^{p,q+1}$ -space).

D.3 THE CONFORMAL COMPACTIFICATION OF THE PSEUDO-EUCLIDEAN SPACE

Notation/Lemma D.3.1 (The isometric embedding). *The following map:*

$$\iota : \mathbb{R}^{p,q} \rightarrow \mathbb{R}^{p+1,q+1}, \quad \iota(x) := \begin{bmatrix} \frac{1-\eta^{p,q}(x,x)}{2} \\ x \\ \frac{1+\eta^{p,q}(x,x)}{2} \end{bmatrix}. \quad (156)$$

is an isometric embedding and satisfies for all $x \in \mathbb{R}^{p,q}$:

$$\iota(x) \neq 0, \quad \text{and} \quad \eta^{p+1,q+1}(\iota(x), \iota(x)) = 0. \quad (157)$$

Proof. For the latter consider the computation for $x \in \mathbb{R}^{p,q}$:

$$4\eta^{p+1,q+1}(\iota(x), \iota(x)) \quad (158)$$

$$= (1 - \eta^{p,q}(x, x))^2 + \eta^{p,q}(2x, 2x) - (1 + \eta^{p,q}(x, x))^2 \quad (159)$$

$$= 1 - 2\eta^{p,q}(x, x) + \eta^{p,q}(x, x)^2 + 4\eta^{p,q}(x, x) - 1 - 2\eta^{p,q}(x, x) - \eta^{p,q}(x, x)^2 \quad (160)$$

$$= 0. \quad (161)$$

This shows: $\eta^{p+1,q+1}(\iota(x), \iota(x)) = 0$. $\iota(x) \neq 0$ is clear. To see that ι is an isometric embedding compute the differential (Jacobian matrix):

$$d\iota_x = \begin{bmatrix} -x^\top \Delta^{p,q} \\ I \\ x^\top \Delta^{p,q} \end{bmatrix}. \quad (162)$$

With this we get:

$$(d\iota_x)^\top \Delta^{p+1,q+1} d\iota_x = [-\Delta^{p,q}x \quad I \quad \Delta^{p,q}x] \Delta^{p+1,q+1} \begin{bmatrix} -x^\top \Delta^{p,q} \\ I \\ x^\top \Delta^{p,q} \end{bmatrix} \quad (163)$$

$$= [-\Delta^{p,q}x \quad \Delta^{p,q} \quad -\Delta^{p,q}x] \begin{bmatrix} -x^\top \Delta^{p,q} \\ I \\ x^\top \Delta^{p,q} \end{bmatrix} \quad (164)$$

$$= \Delta^{p,q}xx^\top \Delta^{p,q} + \Delta^{p,q} - \Delta^{p,q}xx^\top \Delta^{p,q} \quad (165)$$

$$= \Delta^{p,q}. \quad (166)$$

This shows that ι is an isometric embedding. \square

Definitions/Notations D.3.2. For $p, q \geq 0$, $d := p + q$, we introduce the following notations:

1. We introduce the affine zero quadric as the following subspace Y_0 of $\mathbb{R}^{p+1, q+1}$ via:

$$Y_0 := \{y \in \mathbb{R}^{p+1, q+1} \setminus \{0\} \mid \eta^{p+1, q+1}(y, y) = 0\}, \quad (167)$$

and endow the tangent spaces for $y \in Y_0$:

$$\mathbb{T}_y Y_0 = \{v \in \mathbb{R}^{p+1, q+1} \mid \eta^{p+1, q+1}(y, v) = 0\} \subseteq \mathbb{R}^{p+1, q+1}, \quad (168)$$

with the pull-back metric η^{Y_0} of $\eta^{p+1, q+1}$:

$$\eta_y^{Y_0}(v_1, v_2) = \eta^{p+1, q+1}(v_1, v_2) \quad \text{for } v_1, v_2 \in \mathbb{T}_y Y_0. \quad (169)$$

2. We also introduce the double sphere:

$$\mathbb{S}^{p, q} := \mathbb{S}^p \times \mathbb{S}^q := \left\{ y \in \mathbb{R}^{p+1, q+1} \mid \sum_{i=0}^p |y^i|^2 = \sum_{j=p+1}^{d+1} |y^j|^2 = 1 \right\} \subseteq Y_0, \quad (170)$$

and endow the tangent spaces for $y \in \mathbb{S}^{p, q}$:

$$\mathbb{T}_y \mathbb{S}^{p, q} = \left\{ v \in \mathbb{R}^{p+1, q+1} \mid \sum_{i=0}^p y^i \cdot v^i = \sum_{j=p+1}^{d+1} y^j \cdot v^j = 0 \right\} \subseteq \mathbb{R}^{p+1, q+1}. \quad (171)$$

with the pull-back metric $\eta^{\mathbb{S}^{p, q}}$ of $\eta^{p+1, q+1}$:

$$\eta_y^{\mathbb{S}^{p, q}}(v_1, v_2) = \eta^{p+1, q+1}(v_1, v_2) \quad \text{for } v_1, v_2 \in \mathbb{T}_y \mathbb{S}^{p, q}. \quad (172)$$

Note that we have the inclusions:

$$\iota(\mathbb{R}^{p, q}) \subseteq Y_0 \subseteq \mathbb{R}^{p+1, q+1}, \quad \mathbb{S}^{p, q} \subseteq Y_0 \subseteq \mathbb{R}^{p+1, q+1}. \quad (173)$$

Notation/Lemma D.3.3. Consider the maps:

$$\rho : Y_0 \rightarrow \mathbb{R}, \quad \rho(y) := \frac{1}{\sqrt{\sum_{i=0}^p |y^i|^2}}, \quad (174)$$

$$\psi : Y_0 \rightarrow \mathbb{S}^{p, q}, \quad \psi(y) := \rho(y) \cdot y. \quad (175)$$

Then ψ is a well-defined map and conformal with conformal factor $\omega_\psi(y) = \rho(y)$.

Proof. First note that for $y \in Y_0$ we have:

$$0 = \eta^{p+1, q+1}(y, y) = \sum_{i=0}^p |y^i|^2 - \sum_{j=p+1}^{d+1} |y^j|^2, \quad (176)$$

and thus:

$$\rho(y) = \frac{1}{\sqrt{\sum_{i=0}^p |y^i|^2}} \stackrel{!}{=} \frac{1}{\sqrt{\sum_{j=p+1}^{d+1} |y^j|^2}}. \quad (177)$$

This shows that for $y \in Y_0$ we have:

$$\psi(y) = \rho(y) \cdot y \in \mathbb{S}^p \times \mathbb{S}^q = \mathbb{S}^{p, q}. \quad (178)$$

The differential of ψ at $y \in Y_0$:

$$d\psi_y : \mathbb{T}_y Y_0 \rightarrow \mathbb{T}_{\psi(y)} \mathbb{S}^{p, q} \subseteq \mathbb{T}_{\psi(y)} \mathbb{R}^{p+1, q+1}. \quad (179)$$

is given with help of the product rule as the matrix:

$$d\psi_y = y \cdot \rho'(y) + \rho(y) \cdot I. \quad (180)$$

With this we get for all $y \in Y_0$ and $v_1, v_2 \in T_y Y_0$:

$$v_1^\top (d\psi_y)^\top \Delta^{p+1, q+1} (d\psi_y) v_2 \quad (181)$$

$$= v_1^\top (y \cdot \rho'(y) + \rho(y) \cdot I)^\top \Delta^{p+1, q+1} (y \cdot \rho'(y) + \rho(y) \cdot I) v_2 \quad (182)$$

$$= \rho(y)^2 \cdot v_1^\top \Delta^{p+1, q+1} v_2 + v_1^\top \rho'(y)^\top \underbrace{y^\top \Delta^{p+1, q+1} y \rho'(y)}_{=0, \text{ as } y \in Y_0} v_2 \quad (183)$$

$$+ \rho(y) \cdot v_1^\top \rho'(y)^\top \underbrace{y^\top \Delta^{p+1, q+1} v_2}_{=0, \text{ as } v_2 \in T_y Y_0} + v_1^\top \underbrace{\Delta^{p+1, q+1} y \cdot \rho(y)}_{=0, \text{ as } v_1 \in T_y Y_0} \cdot \rho'(y) v_2 \quad (184)$$

$$= \rho(y)^2 \cdot v_1^\top \Delta^{p+1, q+1} v_2. \quad (185)$$

This shows the conformal factor of $\rho(y)$. \square

Notation/Lemma D.3.4. Consider the following map:

$$\tau : \mathbb{R}^{p, q} \xrightarrow{\iota} Y_0 \xrightarrow{\psi} \mathbb{S}^{p, q}, \quad (186)$$

$$\tau(x) := \psi(\iota(x)) = \frac{1}{\sqrt{\frac{1}{4}|1 - \eta^{p, q}(x, x)|^2 + \sum_{i=1}^p |x^i|^2}} \cdot \begin{bmatrix} \frac{1 - \eta^{p, q}(x, x)}{2} \\ x \\ \frac{1 + \eta^{p, q}(x, x)}{2} \end{bmatrix}. \quad (187)$$

Then τ is a well-defined conformal map with conformal factor:

$$\rho(x) = \frac{1}{\sqrt{\frac{1}{4}|1 - \eta^{p, q}(x, x)|^2 + \sum_{i=1}^p |x^i|^2}}. \quad (188)$$

Definition D.3.5 (The standard conformal compactification of the pseudo-Euclidean space). For $p, q \geq 0$, $d := p + q$, the (standard) conformal compactification of $(\mathbb{R}^{p, q}, \eta^{p, q})$, is defined to be the following projective zero quadric subspace of the $(d + 1)$ -dimensional (real) projective space:

$$\mathbb{M}^{p, q} := \{[z] = [z^0 : z^1 : \dots : z^d : z^{d+1}] \in \mathbb{P}^{d+1} \mid \eta^{p+1, q+1}(z, z) = 0\} \subseteq \mathbb{P}^{d+1}. \quad (189)$$

Its metric $\eta^{\mathbb{M}^{p, q}}$ is induced with help of the double sphere:

$$\mathbb{S}^{p, q} := \mathbb{S}^p \times \mathbb{S}^q := \left\{ y \in \mathbb{R}^{p+1, q+1} \mid \sum_{i=0}^p |y^i|^2 = \sum_{j=p+1}^{d+1} |y^j|^2 = 1 \right\} \subseteq \mathbb{R}^{p+1, q+1}, \quad (190)$$

and via the 2 : 1 locally diffeomorphic surjective map:

$$\pi : \mathbb{S}^{p, q} \rightarrow \mathbb{M}^{p, q}, \quad \pm y \mapsto [y]. \quad (191)$$

Note that if $[z] \in \mathbb{M}^{p, q}$ then we can put:

$$y := \psi(z) = \rho(z) \cdot z \in \mathbb{R}^{p+1, q+1}, \quad (192)$$

leading to the following properties¹³:

$$y \in \mathbb{S}^{p, q}, \quad [y] = [z] \in \mathbb{M}^{p, q}. \quad (193)$$

More explicitly, the tangent space of $\mathbb{M}^{p, q}$ at $[z]$ is given by:

$$T_{[z]} \mathbb{M}^{p, q} = \left\{ v \in \mathbb{R}^{p+1, q+1} \mid \sum_{i=0}^p z^i \cdot v^i = \sum_{j=p+1}^{d+1} z^j \cdot v^j = 0 \right\} \quad (194)$$

$$= \left\{ v \in \mathbb{R}^{p+1, q+1} \mid \sum_{i=0}^p y^i \cdot v^i = \sum_{j=p+1}^{d+1} y^j \cdot v^j = 0 \right\} \quad (195)$$

$$= T_y \mathbb{S}^{p, q} \subseteq T_y \mathbb{R}^{p+1, q+1} = \mathbb{R}^{p+1, q+1}. \quad (196)$$

and the metric of $\mathbb{M}^{p, q}$ at $[z]$ for $v_1, v_2 \in T_{[z]} \mathbb{M}^{p, q}$ by:

$$\eta_{[z]}^{\mathbb{M}^{p, q}}(v_1, v_2) := \eta_y^{\mathbb{S}^{p, q}}(v_1, v_2) = \eta^{p+1, q+1}(v_1, v_2) = v_1^\top \Delta^{p+1, q+1} v_2. \quad (197)$$

¹³By also multiplying z with $\text{sgn}(z^j)$ for one index $j \in \{0, 1, \dots, d, d+1\}$, e.g. $j = d+1$, we can, in addition, arrange that $y^j \geq 0$ for this one fixed index j . For example, if $q = 0$ then $\mathbb{S}^{p, q} = \mathbb{S}^p \times \{\pm 1\}$, and we can always arrange the representative y to have: $y^{d+1} = +1$. In this case: $\mathbb{M}^{p, 0} \cong \mathbb{S}^p \times \{+1\}$.

Definition D.3.6 (The conformal embedding of the pseudo-Euclidean space into its conformal compactification). *The conformal embedding of $\mathbb{R}^{p,q}$ into $\mathbb{M}^{p,q}$ is defined to be:*

$$[\iota] : \mathbb{R}^{p,q} \rightarrow \mathbb{M}^{p,q}, \quad [\iota(x)] := \left[\frac{1 - \eta^{p,q}(x, x)}{2} : x : \frac{1 + \eta^{p,q}(x, x)}{2} \right] \quad (198)$$

$$= [1 - \eta^{p,q}(x, x) : 2x : 1 + \eta^{p,q}(x, x)]. \quad (199)$$

Note that this is a well-defined map, as by Notation/Lemma D.3.1 we have:

$$\iota(x) \neq 0, \quad \eta^{p+1, q+1}(\iota(x), \iota(x)) = 0. \quad (200)$$

Note that we can also factorize the conformal embedding as follows:

$$[\iota] : \mathbb{R}^{p,q} \xrightarrow{\tau} \mathbb{S}^{p,q} \xrightarrow{\pi} \mathbb{M}^{p,q}. \quad (201)$$

leading to the identity:

$$[\tau(x)] = [\iota(x)], \quad (202)$$

and showing that the conformal embedding $[\iota]$ is a conformal map with conformal factor:

$$\omega_{[\iota]}(x) = \omega_\tau(x) = \rho(\iota(x)) = \frac{1}{\sqrt{\frac{1}{4}|1 - \eta^{p,q}(x, x)|^2 + \sum_{i=1}^p |x^i|^2}}. \quad (203)$$

D.4 CONFORMAL TRANSFORMATIONS OF THE CONFORMAL COMPACTIFICATION OF THE PSEUDO-EUCLIDEAN SPACE

Proposition D.4.1 (See Schottenloher (2008) Thm. 2.6). *Let $p, q \geq 0$. Then every $[\Lambda] \in \text{PO}(p+1, q+1)$ acts as a conformal diffeomorphism on the conformal compactification $\mathbb{M}^{p,q}$ of $\mathbb{R}^{p,q}$ via matrix multiplication and with the conformal factor:*

$$\omega_{[\Lambda]}([z]) := \sqrt{\frac{\sum_{i=0}^p |z^i|^2}{\sum_{i=0}^p |(\Lambda z)^i|^2}}. \quad (204)$$

The inverse of $[\Lambda]$ is given by $[\Lambda^{-1}]$. Furthermore, exactly the two matrices $\pm\Lambda \in \text{O}(p+1, q+1)$ induce the same conformal diffeomorphism on $\mathbb{M}^{p,q}$, thus inducing an embedding/inclusion of groups:

$$\text{PO}(p+1, q+1) \subseteq \text{ConfDiff}(\mathbb{M}^{p,q}). \quad (205)$$

Proof. For every $\Lambda \in \text{O}(p+1, q+1)$ it is clear that $\Lambda|_{Y_0}$ is an isometric map from Y_0 to Y_0 . Now define the map:

$$\psi_\Lambda : \mathbb{S}^{p,q} \rightarrow \mathbb{S}^{p,q}, \quad \psi_\Lambda(y) := \psi(\Lambda y). \quad (206)$$

As the composition:

$$\psi_\Lambda : \mathbb{S}^{p,q} \subseteq Y_0 \xrightarrow{\Lambda} Y_0 \xrightarrow{\psi} \mathbb{S}^{p,q}, \quad (207)$$

of the isometry Λ (with conformal factor 1) and the conformal map ψ (with conformal factor $\rho(y)$) also ψ_Λ is a conformal map with the conformal factor $\rho(\Lambda y)$.

If now $[z] \in \mathbb{M}^{p,q}$ then $y := \rho(z) \cdot z \in \mathbb{S}^{p,q}$. This then shows that:

$$\psi_{[\Lambda]} : \mathbb{M}^{p,q} \rightarrow \mathbb{M}^{p,q}, \quad \psi_{[\Lambda]}([z]) := [\Lambda z] = [\psi_\Lambda(y)], \quad (208)$$

is a conformal map with conformal factor:

$$\omega_{[\Lambda]}([z]) = \rho(\Lambda y) = \rho(\Lambda(\rho(z)z)) = \frac{\rho(\Lambda z)}{\rho(z)} = \sqrt{\frac{\sum_{i=0}^p |z^i|^2}{\sum_{i=0}^p |(\Lambda z)^i|^2}}. \quad (209)$$

Since that action is induced through matrix multiplication it is clear that $[\Lambda^{-1}]$ induces the inverse map to $[\Lambda]$.

Now assume that Λ_1 and Λ_2 induce the same map on $\mathbb{M}^{p,q}$. Then for every $y \in \mathbb{S}^{p,q}$ we have: $[\psi_{\Lambda_1}(y)] = [\psi_{\Lambda_2}(y)]$. So $\psi_{\Lambda_1}(y) = \pm\psi_{\Lambda_2}(y)$. Since this holds for all $y \in \mathbb{S}^{p,q}$, we get: $\Lambda_1 = \pm\Lambda_2$. This shows the claim. \square

Theorem D.4.2 (See Schottenloher (2008) Thm. 2.6, Thm. 2.9, Thm. 2.11). *If either $p + q \geq 3$ or $(p, q) = (2, 0)$ then the inclusion from Proposition D.4.1 is already an isomorphism:*

$$\text{PO}(p + 1, q + 1) \cong \text{ConfDiff}(\mathbb{M}^{p,q}). \quad (210)$$

This means, in those (p, q) -cases, that every conformal diffeomorphism $\varphi : \mathbb{M}^{p,q} \rightarrow \mathbb{M}^{p,q}$ is given by the matrix multiplication with an (up to sign) unique matrix $\pm\Lambda \in \text{O}(p + 1, q + 1)$, i.e.: $\varphi([z]) = [\Lambda z]$ for all $[z] \in \mathbb{M}^{p,q}$.

Theorem D.4.3 (See Schottenloher (2008) Thm. 2.6, Thm. 2.9, Thm. 2.11). *In addition to the isomorphism in Equation (210) in Theorem D.4.2 we get the following stronger statements:*

1. *Let $p + q \geq 3$, then for every conformal map $\varphi : U \rightarrow \mathbb{R}^{p,q}$, defined on any connected open subset $U \subseteq \mathbb{R}^{p,q}$, there exists an (up to sign) unique matrix $\pm\Lambda \in \text{O}(p + 1, q + 1)$ such that the following diagram commutes:*

$$\begin{array}{ccc} U & \xrightarrow{\varphi} & \mathbb{R}^{p,q} \\ \downarrow [\iota] & & \downarrow [\iota] \\ \mathbb{M}^{p,q} & \xrightarrow{[\Lambda]} & \mathbb{M}^{p,q} \end{array} \quad (211)$$

i.e. for all $x \in U$ we have:

$$[\iota \circ \varphi(x)] = [\Lambda \iota(x)]. \quad (212)$$

In particular, φ is injective.

2. *Let $(p, q) = (2, 0)$, then for every injective conformal map $\varphi : U \rightarrow \mathbb{R}^{2,0}$, defined either on $U = \mathbb{R}^{2,0}$ or on any punctured plane $U = \mathbb{R}^{2,0} \setminus \{\hat{x}\}$, there exists an (up to sign) unique matrix $\pm\Lambda \in \text{O}(3, 1)$ such that the corresponding diagram from 211 commutes.¹⁴*

Remark D.4.4. *For $(p, q) = (1, 1)$ the group $\text{ConfDiff}(\mathbb{M}^{p,q})$ is much bigger than $\text{PO}(p + 1, q + 1)$. For details see Schottenloher (2008) section 2.5.*

Example D.4.5 (The inversion at the pseudo-sphere). *We continue the discussion about the inversion (at the pseudo-sphere) ς on $\mathbb{R}^{p,q}$ from Example D.1.10:*

$$\varsigma = \varsigma^{p,q} : U \rightarrow \mathbb{R}^{p,q}, \quad \varsigma^{p,q}(x) := \frac{x}{\eta^{p,q}(x, x)}. \quad (213)$$

defined on $U := \{x \in \mathbb{R}^{p,q} \mid \eta^{p,q}(x, x) \neq 0\}$. Consider the $(d + 2) \times (d + 2)$ -matrix:

$$\Lambda_\varsigma := \begin{bmatrix} -1 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (214)$$

We now claim that the map:

$$\bar{\varsigma} := [\Lambda_\varsigma] : \mathbb{M}^{p,q} \rightarrow \mathbb{M}^{p,q}, \quad [\Lambda_\varsigma][z] := [\Lambda_\varsigma z], \quad (215)$$

¹⁴Note that, for the case $(p, q) = (2, 0)$, the injectivity of φ needs to be assumed and we also can only allow for open subsets U of $\mathbb{R}^{2,0}$ where at most one point is removed from $\mathbb{R}^{2,0}$.

is the conformal extension of ς from U to whole $\mathbb{M}^{p,q}$. Indeed, first note that $\Lambda_\varsigma \in O(p+1, q+1)$ and $\det \Lambda_\varsigma = -1$ and $\det(-\Lambda_\varsigma) = (-1)^{d+1}$. Then compute:

$$\iota(\varsigma(x)) = \frac{1}{2} \begin{bmatrix} 1 - \eta(\varsigma(x), \varsigma(x)) \\ 2\varsigma(x) \\ 1 + \eta(\varsigma(x), \varsigma(x)) \end{bmatrix} \quad (216)$$

$$= \frac{1}{2} \begin{bmatrix} 1 - \frac{1}{\frac{\eta(x,x)}{2x}} \\ \frac{\eta(x,x)}{2x} \\ 1 + \frac{1}{\frac{\eta(x,x)}{2x}} \end{bmatrix} \quad (217)$$

$$= \frac{1}{2} \frac{1}{\eta(x,x)} \begin{bmatrix} \eta(x,x) - 1 \\ 2x \\ \eta(x,x) + 1 \end{bmatrix} \quad (218)$$

$$= \frac{1}{2} \frac{1}{\eta(x,x)} \begin{bmatrix} -1 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 - \eta(x,x) \\ 2x \\ 1 + \eta(x,x) \end{bmatrix} \quad (219)$$

$$= \frac{1}{\eta(x,x)} \cdot \Lambda_\varsigma \iota(x), \quad (220)$$

which shows the claim:

$$[\Lambda_\varsigma \iota(x)] = [\iota(\varsigma(x))] \in \mathbb{M}^{p,q}. \quad (221)$$

Note that for $x \in \mathbb{R}^{p,q}$ with $\eta(x,x) = 0$ we get:

$$\iota(x) = [1 : 2x : 1], \quad \bar{\varsigma}(\iota(x)) = [-1 : 2x : 1], \quad (222)$$

$$\iota(0) = [1 : 0 : 1], \quad \bar{\varsigma}(\iota(0)) = [-1 : 0 : 1]. \quad (223)$$

Example D.4.6 (Linear conformal transformations, see Schottenloher (2008) Thm. 2.9). We continue from Example D.1.7 and Theorem D.1.8 for the linear conformal map: $x \mapsto Ax$ with the matrix $A = c\Lambda$, where $A = c\Lambda \in CO(p,q)$ with $c > 0$ and $\Lambda \in O(p,q)$. Then we can define the $(d+2) \times (d+2)$ -matrix:

$$\Gamma_{c\Lambda} := \begin{bmatrix} \frac{1+c^2}{2c} & 0 & \frac{1-c^2}{2c} \\ 0 & \Lambda & 0 \\ \frac{1-c^2}{2c} & 0 & \frac{1+c^2}{2c} \end{bmatrix}. \quad (224)$$

Then $\Gamma_{c\Lambda} \in O(p+1, q+1)$ as:

$$\Gamma_{c\Lambda}^\top \Delta^{p+1, q+1} \Gamma_{c\Lambda} = \begin{bmatrix} \frac{1+c^2}{2c} & 0 & \frac{1-c^2}{2c} \\ 0 & \Lambda^\top & 0 \\ \frac{1-c^2}{2c} & 0 & \frac{1+c^2}{2c} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \Delta^{p,q} & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1+c^2}{2c} & 0 & \frac{1-c^2}{2c} \\ 0 & \Lambda & 0 \\ \frac{1-c^2}{2c} & 0 & \frac{1+c^2}{2c} \end{bmatrix} \quad (225)$$

$$= \begin{bmatrix} \frac{1+c^2}{2c} & 0 & \frac{1-c^2}{2c} \\ 0 & \Lambda^\top & 0 \\ \frac{1-c^2}{2c} & 0 & \frac{1+c^2}{2c} \end{bmatrix} \begin{bmatrix} \frac{1+c^2}{2c} & 0 & \frac{1-c^2}{2c} \\ 0 & \Delta^{p,q} \Lambda & 0 \\ -\frac{1-c^2}{2c} & 0 & -\frac{1+c^2}{2c} \end{bmatrix} \quad (226)$$

$$= \begin{bmatrix} \left(\frac{1+c^2}{2c}\right)^2 - \left(\frac{1-c^2}{2c}\right)^2 & 0 & 0 \\ 0 & \Lambda^\top \Delta^{p,q} \Lambda & 0 \\ 0 & 0 & \left(\frac{1-c^2}{2c}\right)^2 - \left(\frac{1+c^2}{2c}\right)^2 \end{bmatrix} \quad (227)$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \Delta^{p,q} & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (228)$$

$$= \Delta^{p+1, q+1}. \quad (229)$$

Also note that if $\Lambda \in O^0(p,q)$ then $\Gamma_{c\Lambda} \in O^0(p+1, q+1)$ as $c \rightarrow 1$ provides a path to the identity component.¹⁵ We now claim that:

$$[\Gamma_{c\Lambda}] : \mathbb{M}^{p,q} \rightarrow \mathbb{M}^{p,q}, \quad [\Gamma_{c\Lambda}][z] := [\Gamma_{c\Lambda}z], \quad (230)$$

¹⁵The reverse is also true: $\Gamma_{c\Lambda} \in O^0(p+1, q+1) \implies \Lambda \in O^0(p,q)$.

is the conformal extension of $c\Lambda$ from $\mathbb{R}^{p,q}$ to $\mathbb{M}^{p,q}$.

$$\Gamma_{c\Lambda}\iota(x) = \begin{bmatrix} \frac{1+c^2}{2c} & 0 & \frac{1-c^2}{2c} \\ 0 & \Lambda & 0 \\ \frac{1-c^2}{2c} & 0 & \frac{1+c^2}{2c} \end{bmatrix} \begin{bmatrix} \frac{1-\eta^{p,q}(x,x)}{2} \\ x \\ \frac{1+\eta^{p,q}(x,x)}{2} \end{bmatrix} \quad (231)$$

$$= \begin{bmatrix} \left(\frac{1+c^2}{2c}\right) \left(\frac{1-\eta^{p,q}(x,x)}{2}\right) + \left(\frac{1-c^2}{2c}\right) \left(\frac{1+\eta^{p,q}(x,x)}{2}\right) \\ \Lambda x \\ \left(\frac{1-c^2}{2c}\right) \left(\frac{1-\eta^{p,q}(x,x)}{2}\right) + \left(\frac{1+c^2}{2c}\right) \left(\frac{1+\eta^{p,q}(x,x)}{2}\right) \end{bmatrix} \quad (232)$$

$$= \frac{1}{4c} \begin{bmatrix} 1 + c^2 - \eta^{p,q}(x,x) - c^2 \cdot \eta^{p,q}(x,x) + 1 - c^2 + \eta^{p,q}(x,x) - c^2 \cdot \eta^{p,q}(x,x) \\ 4c\Lambda x \\ 1 - c^2 - \eta^{p,q}(x,x) + c^2 \cdot \eta^{p,q}(x,x) + 1 + c^2 + \eta^{p,q}(x,x) + c^2 \cdot \eta^{p,q}(x,x) \end{bmatrix} \quad (233)$$

$$= \frac{1}{4c} \begin{bmatrix} 2 - 2 \cdot c^2 \cdot \eta^{p,q}(x,x) \\ 4c\Lambda x \\ 2 + 2 \cdot c^2 \cdot \eta^{p,q}(x,x) \end{bmatrix} \quad (234)$$

$$= \frac{1}{c} \begin{bmatrix} \frac{1-\eta^{p,q}(c\Lambda x, c\Lambda x)}{2} \\ c\Lambda x \\ \frac{1+\eta^{p,q}(c\Lambda x, c\Lambda x)}{2} \end{bmatrix} \quad (235)$$

$$= \frac{1}{c} \cdot \iota(c\Lambda x). \quad (236)$$

This shows:

$$[\Gamma_{c\Lambda}\iota(x)] = [\iota(c\Lambda x)] \in \mathbb{M}^{p,q}, \quad (237)$$

and thus the claim.

Example D.4.7 (Translations, see Schottenloher (2008) Thm. 2.9). We now continue from Example D.1.7 and Theorem D.1.8 for case of translation: $x \mapsto x + b$ for $b \in \mathbb{R}^{p,q}$. Then we can define the $(d+2) \times (d+2)$ -matrix:

$$\Gamma_{I,b} := I + \Omega_b \quad (238)$$

$$:= I + \begin{bmatrix} -\frac{1}{2}\eta^{p,q}(b,b) & -b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b,b) \\ b & 0 & b \\ \frac{1}{2}\eta^{p,q}(b,b) & b^\top \Delta^{p,q} & \frac{1}{2}\eta^{p,q}(b,b) \end{bmatrix} \quad (239)$$

$$= \begin{bmatrix} 1 - \frac{1}{2}\eta^{p,q}(b,b) & -b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b,b) \\ b & I & b \\ \frac{1}{2}\eta^{p,q}(b,b) & b^\top \Delta^{p,q} & 1 + \frac{1}{2}\eta^{p,q}(b,b) \end{bmatrix}. \quad (240)$$

We first show that $\Gamma_{I,b} \in O(p+1, q+1)$. For this consider:

$$\Gamma_{I,b}^\top \Delta^{p+1, q+1} \Gamma_{I,b} = \Delta^{p+1, q+1} + \Omega_b^\top \Delta^{p+1, q+1} + \Delta^{p+1, q+1} \Omega_b + \Omega_b^\top \Delta^{p+1, q+1} \Omega_b. \quad (241)$$

We compute terms separately:

$$\Delta^{p+1, q+1} \Omega_b = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \Delta^{p,q} & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} -\frac{1}{2}\eta^{p,q}(b,b) & -b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b,b) \\ b & 0 & b \\ \frac{1}{2}\eta^{p,q}(b,b) & b^\top \Delta^{p,q} & \frac{1}{2}\eta^{p,q}(b,b) \end{bmatrix} \quad (242)$$

$$= \begin{bmatrix} -\frac{1}{2}\eta^{p,q}(b,b) & -b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b,b) \\ \Delta^{p,q} b & 0 & \Delta^{p,q} b \\ -\frac{1}{2}\eta^{p,q}(b,b) & -b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b,b) \end{bmatrix}, \quad (243)$$

$$\Omega_b^\top \Delta^{p+1, q+1} = (\Delta^{p+1, q+1} \Omega_b)^\top \quad (244)$$

$$= \begin{bmatrix} -\frac{1}{2}\eta^{p,q}(b,b) & b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b,b) \\ -\Delta^{p,q} b & 0 & -\Delta^{p,q} b \\ -\frac{1}{2}\eta^{p,q}(b,b) & b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b,b) \end{bmatrix}. \quad (245)$$

With this we compute:

$$\Omega_b^\top \Delta^{p+1,q+1} + \Delta^{p+1,q+1} \Omega_b = -\eta^{p,q}(b, b) \cdot \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad (246)$$

$$(247)$$

We also get:

$$\Omega_b^\top \Delta^{p+1,q+1} \Omega_b = \begin{bmatrix} -\frac{1}{2}\eta^{p,q}(b, b) & b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b, b) \\ -\Delta^{p,q}b & 0 & -\Delta^{p,q}b \\ -\frac{1}{2}\eta^{p,q}(b, b) & b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b, b) \end{bmatrix} \begin{bmatrix} -\frac{1}{2}\eta^{p,q}(b, b) & -b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b, b) \\ b & 0 & b \\ \frac{1}{2}\eta^{p,q}(b, b) & b^\top \Delta^{p,q} & \frac{1}{2}\eta^{p,q}(b, b) \end{bmatrix} \quad (248)$$

$$= \begin{bmatrix} b^\top \Delta^{p,q}b & 0 & b^\top \Delta^{p,q}b \\ 0 & 0 & 0 \\ b^\top \Delta^{p,q}b & 0 & b^\top \Delta^{p,q}b \end{bmatrix} \quad (249)$$

$$= \eta^{p,q}(b, b) \cdot \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}. \quad (250)$$

Together this shows:

$$\Gamma_{I,b}^\top \Delta^{p+1,q+1} \Gamma_{I,b} = \Delta^{p+1,q+1} + \Omega_b^\top \Delta^{p+1,q+1} + \Delta^{p+1,q+1} \Omega_b + \Omega_b^\top \Delta^{p+1,q+1} \Omega_b \quad (251)$$

$$= \Delta^{p+1,q+1} - \eta^{p,q}(b, b) \cdot \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} + \eta^{p,q}(b, b) \cdot \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad (252)$$

$$= \Delta^{p+1,q+1}. \quad (253)$$

This thus shows the claim: $\Gamma_{I,b} \in \mathcal{O}(p+1, q+1)$. Furthermore, for $b \rightarrow 0$ we see that $\Gamma_{I,b} \rightarrow I$.

This thus even shows that: $\Gamma_{I,b} \in \mathcal{O}^0(p+1, q+1)$.

We now claim that $[\Gamma_{I,b}]$ is the conformal extension of the translation map $x \mapsto x + b$ from $\mathbb{R}^{p,q}$ to $\mathbb{M}^{p,q}$. For this compute:

$$\Gamma_{I,b}\iota(x) = (I + \Omega_b)\iota(x) \quad (254)$$

$$= \iota(x) + \Omega_b\iota(x), \quad (255)$$

with:

$$\Omega_b\iota(x) = \frac{1}{2} \begin{bmatrix} -\frac{1}{2}\eta^{p,q}(b, b) & -b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b, b) \\ b & 0 & b \\ \frac{1}{2}\eta^{p,q}(b, b) & b^\top \Delta^{p,q} & \frac{1}{2}\eta^{p,q}(b, b) \end{bmatrix} \begin{bmatrix} 1 - \eta^{p,q}(x, x) \\ 2x \\ 1 + \eta^{p,q}(x, x) \end{bmatrix} \quad (256)$$

$$= \frac{1}{2} \begin{bmatrix} -\eta^{p,q}(b, b) - 2b^\top \Delta^{p,q}x \\ 2b \\ \eta^{p,q}(b, b) + 2b^\top \Delta^{p,q}x \end{bmatrix} \quad (257)$$

$$= \frac{1}{2} \begin{bmatrix} -\eta^{p,q}(b, b) - 2\eta^{p,q}(b, x) \\ 2b \\ \eta^{p,q}(b, b) + 2\eta^{p,q}(b, x) \end{bmatrix}. \quad (258)$$

With this we get:

$$\iota(x + b) = \frac{1}{2} \begin{bmatrix} 1 - \eta^{p,q}(x + b, x + b) \\ 2(x + b) \\ 1 + \eta^{p,q}(x + b, x + b) \end{bmatrix} \quad (259)$$

$$= \frac{1}{2} \begin{bmatrix} 1 - \eta^{p,q}(x, x) - \eta^{p,q}(b, b) - 2\eta^{p,q}(b, x) \\ 2x + 2b \\ 1 + \eta^{p,q}(x, x) + \eta^{p,q}(b, b) + 2\eta^{p,q}(b, x) \end{bmatrix} \quad (260)$$

$$= \frac{1}{2} \begin{bmatrix} 1 - \eta^{p,q}(x, x) \\ 2x \\ 1 + \eta^{p,q}(x, x) \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -\eta^{p,q}(b, b) - 2\eta^{p,q}(b, x) \\ 2b \\ \eta^{p,q}(b, b) + 2\eta^{p,q}(b, x) \end{bmatrix} \quad (261)$$

$$= \iota(x) + \Omega_b\iota(x) \quad (262)$$

$$= \Gamma_{I,b}\iota(x). \quad (263)$$

This then implies:

$$[\Gamma_{I,b}l(x)] = [l(x+b)] \in \mathbb{M}^{p,q}. \quad (264)$$

This shows that $[\Gamma_{I,b}]$ conformally extends the translation map $x \mapsto x+b$ from $\mathbb{R}^{p,q}$ to $\mathbb{M}^{p,q}$.

Example D.4.8 (Affine conformal transformations, see Schottenloher (2008) Thm. 2.9). We continue from Example D.1.7 and Theorem D.1.8 for the affine conformal map: $x \mapsto Ax+b$ with the matrix $A = c\Lambda$ and $b \in \mathbb{R}^{p,q}$, where $A = c\Lambda \in \text{CO}(p,q)$ with $c > 0$ and $\Lambda \in \text{O}(p,q)$. Then we can define the $(d+2) \times (d+2)$ -matrix:

$$\Gamma_{c\Lambda,b} := \Gamma_{I,b}\Gamma_{c\Lambda} \quad (265)$$

$$= \begin{bmatrix} 1 - \frac{1}{2}\eta^{p,q}(b,b) & -b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b,b) \\ b & I & b \\ \frac{1}{2}\eta^{p,q}(b,b) & b^\top \Delta^{p,q} & 1 + \frac{1}{2}\eta^{p,q}(b,b) \end{bmatrix} \begin{bmatrix} \frac{1+c^2}{2c} & 0 & \frac{1-c^2}{2c} \\ 0 & \Lambda & 0 \\ \frac{1-c^2}{2c} & 0 & \frac{1+c^2}{2c} \end{bmatrix}. \quad (266)$$

Then we have: $\Gamma_{c\Lambda} \in \text{O}(p+1, q+1)$. This is clear, as both matrices $\Gamma_{I,b}, \Gamma_{c\Lambda} \in \text{O}(p+1, q+1)$ by Example D.4.6 and Example D.4.7. It thus also clear that the matrix $[\Gamma_{c\Lambda,b}]$ extends the affine map $x \mapsto c\Lambda x + b$ from $\mathbb{R}^{p,q}$ to $\mathbb{M}^{p,q}$.

Example D.4.9 (Special conformal transformations, see Schottenloher (2008) Thm. 2.9). We continue from Example D.1.11 about the special conformal transformations. Recall that we defined the special conformal transformations for $b \in \mathbb{R}^{p,q}$ on an open subset $U_b \subseteq \mathbb{R}^{p,q}$ as:

$$\sigma_b : U_b \rightarrow \mathbb{R}^{p,q}, \quad \sigma_b(x) := \frac{x - \eta(x,x) \cdot b}{1 - 2 \cdot \eta(x,b) + \eta(b,b) \cdot \eta(x,x)} = \varsigma(\varsigma(x) - b), \quad (267)$$

where ς is the inversion at the pseudo-sphere from Example D.1.10 and Example D.4.5. With the latter representation and Example D.4.7 we immediately get that the $(d+2) \times (d+2)$ -matrix:

$$\Sigma_b := \Lambda_\varsigma \Gamma_{I,-b} \Lambda_\varsigma \quad (268)$$

$$= \begin{bmatrix} -1 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 - \frac{1}{2}\eta^{p,q}(b,b) & b^\top \Delta^{p,q} & -\frac{1}{2}\eta^{p,q}(b,b) \\ -b & I & -b \\ \frac{1}{2}\eta^{p,q}(b,b) & -b^\top \Delta^{p,q} & 1 + \frac{1}{2}\eta^{p,q}(b,b) \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (269)$$

$$= \begin{bmatrix} 1 - \frac{1}{2}\eta^{p,q}(b,b) & -b^\top \Delta^{p,q} & \frac{1}{2}\eta^{p,q}(b,b) \\ b & I & -b \\ -\frac{1}{2}\eta^{p,q}(b,b) & -b^\top \Delta^{p,q} & 1 + \frac{1}{2}\eta^{p,q}(b,b) \end{bmatrix} \quad (270)$$

lies in $\text{O}(p+1, q+1)$ and that $[\Sigma_b]$ extends the special conformal transformation σ_b from $\mathbb{R}^{p,q}$ to $\mathbb{M}^{p,q}$. Furthermore, for $b \rightarrow 0$ we see that $\Sigma_b \rightarrow I$. This then even shows that: $\Sigma_b \in \text{O}^0(p+1, q+1)$.

D.5 THE DEFINITION OF THE CONFORMAL GROUP OF THE PSEUDO-EUCLIDEAN SPACE

Definition D.5.1 (The global and restricted conformal group of the pseudo-Euclidean space). Based on Theorem D.4.2 we define for all $p, q \geq 0$ the global conformal group of $\mathbb{R}^{p,q}$ to be:

$$\text{Conf}_g(\mathbb{R}^{p,q}) := \text{PO}(p+1, q+1), \quad (271)$$

and the (restricted) conformal group $\text{Conf}(\mathbb{R}^{p,q})$ of $\mathbb{R}^{p,q}$ to be the connected component of the identity of $\text{PO}(p+1, q+1)$:

$$\text{Conf}(\mathbb{R}^{p,q}) := \text{Conf}_r(\mathbb{R}^{p,q}) := \text{Conf}_g^0(\mathbb{R}^{p,q}) = \text{PO}^0(p+1, q+1). \quad (272)$$

Note that by Proposition D.4.1 we always have the inclusions:

$$\text{Conf}_r(\mathbb{R}^{p,q}) \subseteq \text{Conf}_g(\mathbb{R}^{p,q}) \subseteq \text{ConfDiff}(\mathbb{M}^{p,q}). \quad (273)$$

Theorem D.5.2. $\text{Conf}(\mathbb{R}^{p,q})$ contains all affine conformal transformations $x \mapsto Ax+b$ of $\mathbb{R}^{p,q}$ where $A = c\Lambda \in \text{CO}^0(p,q)$ with $c > 0$ and $\Lambda \in \text{O}^0(p,q)$ and $b \in \mathbb{R}^{p,q}$ in form of the map $[\Gamma_{c\Lambda,b}] : \mathbb{M}^{p,q} \rightarrow \mathbb{M}^{p,q}$ from Example D.4.8, and, also all the special conformal transformations σ_b in form of the map $[\Sigma_b] : \mathbb{M}^{p,q} \rightarrow \mathbb{M}^{p,q}$ from Example D.4.9.

D.6 A (PARTIAL) PARAMETERIZATION OF THE ANTI-DE SITTER SPACE

Proposition D.6.1. *Let $p, q \geq 0$ and $d := p + q$. Consider the following space:*

$$\mathbb{A}^{p+1,q} := \mathbb{R}_{>0} \times \mathbb{R}^{p,q}, \quad (274)$$

which we endow with the metric $\eta^{\mathbb{A}^{p+1,q}}$, which is given at $x \in \mathbb{A}^{p+1,q}$ and with $v_1, v_2 \in \mathbb{T}_x \mathbb{A}^{p+1,q} = \mathbb{R}^{p+1,q}$ as follows:

$$\eta_x^{\mathbb{A}^{p+1,q}}(v_1, v_2) := \frac{\eta^{p+1,q}(v_1, v_2)}{|x^0|^2}. \quad (275)$$

Then consider the map:

$$\phi : \mathbb{A}^{p+1,q} \rightarrow \mathbb{R}^{p+1,q+1}, \quad x \mapsto \phi(x) =: y, \quad (276)$$

$$(277)$$

$$\phi^0(x) := \frac{1 - \eta^{p+1,q}(x, x)}{2x^0} = \frac{1}{2x^0} - \frac{x^0}{2} - \frac{1}{2x^0} \eta^{p,q}(x^{1:d}, x^{1:d}), \quad (278)$$

$$\phi^{1:d}(x) := \frac{x^{1:d}}{x^0}, \quad (279)$$

$$\phi^{d+1}(x) := \frac{1 + \eta^{p+1,q}(x, x)}{2x^0} = \frac{1}{2x^0} + \frac{x^0}{2} + \frac{1}{2x^0} \eta^{p,q}(x^{1:d}, x^{1:d}). \quad (280)$$

Then ϕ is a well-defined isometric embedding and its range coincides with the subspace:

$$\widetilde{\text{AdS}}^{p+1,q} := \{y \in \mathbb{R}^{p+1,q+1} \mid \eta^{p+1,q+1}(y, y) = -1 \wedge y^0 + y^{d+1} > 0\}. \quad (281)$$

An inverse $\phi^{-1} : \widetilde{\text{AdS}}^{p+1,q} \rightarrow \mathbb{A}^{p+1,q}$ of ϕ is given by:

$$x^0 := (\phi^{-1})^0(y) = \frac{1}{y^0 + y^{d+1}} > 0, \quad (282)$$

$$x^{1:d} := (\phi^{-1})^{1:d}(y) = \frac{y^{1:d}}{y^0 + y^{d+1}}. \quad (283)$$

Proof. We first show that ϕ is well-defined. For this we compute for $y := \phi(x)$:

$$4|x^0|^2 \cdot \eta^{p+1,q+1}(y, y) = (1 - \eta^{p+1,q}(x, x))^2 + 4\eta^{p,q}(x^{1:d}, x^{1:d}) - (1 + \eta^{p+1,q}(x, x))^2 \quad (284)$$

$$= -4\eta^{p+1,q}(x, x) + 4\eta^{p,q}(x^{1:d}, x^{1:d}) \quad (285)$$

$$= -4|x^0|^2, \quad (286)$$

implying: $\eta^{p+1,q+1}(y, y) = -1$. Furthermore:

$$y^0 + y^{d+1} = \frac{1 - \eta^{p+1,q}(x, x)}{2x^0} + \frac{1 + \eta^{p+1,q}(x, x)}{2x^0} = \frac{1}{x^0} > 0. \quad (287)$$

Together, this shows: $\phi(x) = y \in \widetilde{\text{AdS}}^{p+1,q} \subseteq \mathbb{R}^{p+1,q+1}$. It is easy to see that ϕ^{-1} is the left-inverse to ϕ , i.e.:

$$\phi^{-1} \circ \phi(x) = x \in \mathbb{A}^{p+1,q}. \quad (288)$$

To show that ϕ^{-1} is also a right-invers let $y \in \widetilde{\text{AdS}}^{p+1,q}$ and $x := \phi^{-1}(y)$. Then compute:

$$\phi^{1:d}(x) = \frac{1}{x^0} x^{1:d} = (y^0 + y^{d+1}) \cdot \frac{y^{1:d}}{y^0 + y^{d+1}} = y^{1:d}, \quad (289)$$

$$\phi^0(x) = \frac{1}{2x^0} - \frac{x^0}{2} - \frac{1}{2x^0} \eta^{p,q}(x^{1:d}, x^{1:d}) \quad (290)$$

$$= \frac{y^0 + y^{d+1}}{2} - \frac{1}{2(y^0 + y^{d+1})} - \frac{1}{2(y^0 + y^{d+1})} \eta^{p,q}(y^{1:d}, y^{1:d}) \quad (291)$$

$$= \frac{1}{2} \left((y^0 + y^{d+1}) - \frac{1}{y^0 + y^{d+1}} (1 + \eta^{p,q}(y^{1:d}, y^{1:d})) \right) \quad (292)$$

$$= \frac{1}{2} \left((y^0 + y^{d+1}) - \frac{1}{y^0 + y^{d+1}} (|y^{d+1}|^2 - |y^0|^2) \right) \quad (293)$$

$$= \frac{1}{2} ((y^0 + y^{d+1}) - (y^{d+1} - y^0)) \quad (294)$$

$$= y^0. \quad (295)$$

Similarly:

$$\phi^{d+1}(x) = \frac{1}{2x^0} + \frac{x^0}{2} + \frac{1}{2x^0} \eta^{p,q}(x^{1:d}, x^{1:d}) \quad (296)$$

$$= \frac{y^0 + y^{d+1}}{2} + \frac{1}{2(y^0 + y^{d+1})} + \frac{1}{2(y^0 + y^{d+1})} \eta^{p,q}(y^{1:d}, y^{1:d}) \quad (297)$$

$$= \frac{1}{2} \left((y^0 + y^{d+1}) + \frac{1}{y^0 + y^{d+1}} (1 + \eta^{p,q}(y^{1:d}, y^{1:d})) \right) \quad (298)$$

$$= \frac{1}{2} \left((y^0 + y^{d+1}) + \frac{1}{y^0 + y^{d+1}} (|y^{d+1}|^2 - |y^0|^2) \right) \quad (299)$$

$$= \frac{1}{2} ((y^0 + y^{d+1}) + (y^{d+1} - y^0)) \quad (300)$$

$$= y^{d+1}. \quad (301)$$

This shows that:

$$\phi \circ \phi^{-1}(y) = y \in \widetilde{\text{AdS}}^{p+1,q}. \quad (302)$$

We now compute the differential of ϕ :

$$d\phi_x = \begin{bmatrix} -\frac{1}{2|x^0|^2} - \frac{1}{2} + \frac{\eta^{p,q}(x^{1:d}, x^{1:d})}{2|x^0|^2} & -\frac{1}{x^0} (x^{1:d})^\top \Delta^{p,q} \\ -\frac{x^{1:d}}{|x^0|^2} & \frac{I_d}{x^0} \\ -\frac{1}{2|x^0|^2} + \frac{1}{2} - \frac{\eta^{p,q}(x^{1:d}, x^{1:d})}{2|x^0|^2} & \frac{1}{x^0} (x^{1:d})^\top \Delta^{p,q} \end{bmatrix} \quad (303)$$

and the metric at $x \in \mathbb{A}^{p+1,q}$:

$$(d\phi_x)^\top \Delta^{p+1,q+1} d\phi_x = (d\phi_x)^\top \begin{bmatrix} 1 & 0 & 0 \\ 0 & \Delta^{p,q} & 0 \\ 0 & 0 & -1 \end{bmatrix} d\phi_x. \quad (304)$$

We compute the entries of the last matrix separately. First, top left entry:

$$\frac{1}{4} \left(1 - \frac{\eta^{p,q}(x^{1:d}, x^{1:d})}{|x^0|^2} + \frac{1}{|x^0|^2} \right)^2 + \frac{\eta^{p,q}(x^{1:d}, x^{1:d})}{|x^0|^4} - \frac{1}{4} \left(1 - \frac{\eta^{p,q}(x^{1:d}, x^{1:d})}{|x^0|^2} - \frac{1}{|x^0|^2} \right)^2 \quad (305)$$

$$= \frac{1}{|x^0|^2} \left(1 - \frac{\eta^{p,q}(x^{1:d}, x^{1:d})}{|x^0|^2} \right) + \frac{\eta^{p,q}(x^{1:d}, x^{1:d})}{|x^0|^4} \quad (306)$$

$$= \frac{1}{|x^0|^2}. \quad (307)$$

Note that in the middle we used the formula: $(a + b)^2 - (a - b)^2 = 4ab$. Next, top right entry:

$$\frac{1}{2} \left(\left(1 - \frac{\eta^{p,q}(x^{1:d}, x^{1:d})}{|x^0|^2} \right) + \frac{1}{|x^0|^2} \right) \frac{(x^{1:p})^\top \Delta^{p,q}}{x^0} - \frac{(x^{1:d})^\top \Delta^{p,q}}{|x^0|^3} \quad (308)$$

$$- \frac{1}{2} \left(\left(1 - \frac{\eta^{p,q}(x^{1:d}, x^{1:d})}{|x^0|^2} \right) - \frac{1}{|x^0|^2} \right) \frac{(x^{1:d})^\top \Delta^{p,q}}{x^0} \quad (309)$$

$$= \frac{1}{2} \left(\frac{2}{|x^0|^2} \right) \frac{(x^{1:d})^\top \Delta^{p,q}}{x^0} - \frac{(x^{1:d})^\top \Delta^{p,q}}{|x^0|^3} \quad (310)$$

$$= 0. \quad (311)$$

By symmetry also the bottom left equals 0. Finally, bottom right entry:

$$\frac{\Delta^{p,q} x^{1:d}}{x^0} \frac{(x^{1:d})^\top \Delta^{p,q}}{x^0} + \frac{I_p}{x^0} \Delta^{p,q} \frac{I_p}{x^0} - \frac{\Delta^{p,q} x^{1:d}}{x^0} \frac{(x^{1:d})^\top \Delta^{p,q}}{x^0} = \frac{\Delta^{p,q}}{|x^0|^2}. \quad (312)$$

Together we get:

$$(d\phi_x)^\top \Delta^{p+1,q+1} d\phi_x = \frac{1}{|x^0|^2} \begin{bmatrix} 1 & 0 \\ 0 & \Delta^{p,q} \end{bmatrix} = \frac{1}{|x^0|^2} \Delta^{p+1,q}. \quad (313)$$

This shows the claim. \square

Theorem D.6.2. *The map:*

$$\mathbb{A}^{p+1,q} \cong \widetilde{\text{AdS}}^{p+1,q} \rightarrow \text{AdS}^{p+1,q}, \quad x \mapsto \phi(x) =: y \mapsto [y], \quad (314)$$

is an isometric embedding. The complement of its range is given by the closed subspace:

$$\mathcal{Z} := \{[z] \in \text{AdS}^{p+1,q} \mid z^0 + z^{d+1} = 0\}. \quad (315)$$

The left-inverse for $[z] \in \text{AdS}^{p+1,q}$ (with $z^0 + z^{d+1} \neq 0$) is given by:

$$y := \frac{\text{sgn}(z^0 + z^{d+1})}{\sqrt{|\eta^{p+1,q+1}(z, z)|}} \cdot z \in \widetilde{\text{AdS}}^{p+1,q}, \quad (316)$$

and to get $x \in \mathbb{A}^{p+1,q}$ we put:

$$x^0 := (\phi^{-1})^0(y) = \frac{\sqrt{|\eta^{p+1,q+1}(z, z)|}}{|z^{d+1} + z^0|} > 0, \quad (317)$$

$$x^{1:d} := (\phi^{-1})^{1:d}(y) = \frac{z^{1:d}}{z^{d+1} + z^0}. \quad (318)$$

Remark D.6.3. Note that in Theorem D.6.2 in the Euclidean case, i.e. if $q = 0$, the subspace \mathcal{Z} is empty, as $z^0 + z^{d+1} = 0$ would imply:

$$0 > \eta^{p+1,q+1}(z, z) \quad (319)$$

$$= |z^0|^2 + \eta^{p,q}(z^{1:d}, z^{1:d}) - |z^{d+1}|^2 \quad (320)$$

$$= (z^0 + z^{d+1})(z^0 - z^{d+1}) + \eta^{p,q}(z^{1:d}, z^{1:d}) \quad (321)$$

$$= \eta^{p,q}(z^{1:d}, z^{1:d}), \quad (322)$$

which is not possible for $q = 0$. It follows that for $q = 0$ we get isometries:

$$\mathbb{A}^{p+1,0} \cong \widetilde{\text{AdS}}^{p+1,0} \cong \text{AdS}^{p+1,0}. \quad (323)$$

Remark D.6.4. For $p, q \geq 0$, $d := p + q$, consider the map:

$$\bar{\phi} : \mathbb{R}_{\geq 0} \times \mathbb{R}^{p,q} \rightarrow \mathbb{P}^{d+1}, \quad x \mapsto [1 - \eta^{p+1,q}(x, x) : 2x^{1:d} : 1 + \eta^{p+1,q}(x, x)]. \quad (324)$$

Then $\bar{\phi}$ is well-defined. For $x^0 = 0$ the map $\bar{\phi}$ coincides with the conformal embedding ι of $\mathbb{R}^{p,q}$ into the conformal compactification $\mathbb{M}^{p,q} \subseteq \mathbb{P}^{d+1}$ of $\mathbb{R}^{p,q}$, and, for $x^0 > 0$, $\bar{\phi}$ coincides with the isometric embedding ϕ of $\mathbb{A}^{p+1,q}$ into $\text{AdS}^{p+1,q} \subseteq \mathbb{P}^{d+1}$.

D.7 THE EUCLIDEAN CASE ($p \geq 2$ AND $q = 0$)

In the following we now consider the Euclidean case, i.e. $q = 0$, for $p \geq 2$, $d = p$.

Remark D.7.1. Note that we have:

$$\mathbb{S}^0 = \{\pm 1\}, \quad \mathbb{S}^{d,0} = \mathbb{S}^d \times \{\pm 1\}. \quad (325)$$

So, for $q = 0$, This thus simplifies the description of the conformal compactification, the corresponding de Sitter and anti-de Sitter space.

Corollary D.7.2 (The conformal compactification of the Euclidean space). *We have the isometry for $d \geq 0$ and $q = 0$, $p = d$:*

$$\mathbb{M}^{d,0} \cong \mathbb{S}^d \times \{+1\} \cong \mathbb{S}^d, \quad (326)$$

$$[z] \mapsto \left(\frac{z^{0:d}}{z^{d+1}}, +1 \right) \mapsto \frac{z^{0:d}}{z^{d+1}}. \quad (327)$$

The conformal embedding is then given as:

$$\tau_+ : \mathbb{R}^{d,0} \rightarrow \mathbb{M}^{d,0} \cong \mathbb{S}^d \subseteq \mathbb{R}^{d+1}, \quad \tau_+(x) := \left(\frac{1 - \eta^{d,0}(x, x)}{1 + \eta^{d,0}(x, x)}, \frac{2x}{1 + \eta^{d,0}(x, x)} \right), \quad (328)$$

which coincides with the invers stereographic projection (from the ‘‘south pole’’).

The group $\text{PO}(d+1, 1)$ then acts on \mathbb{S}^d as follows, $[\Lambda] \in \text{PO}(d+1, 1)$, $y = y^{0:d} \in \mathbb{S}^d$:

$$[\Lambda] \cdot y = \frac{\left(\Lambda \begin{bmatrix} y \\ 1 \end{bmatrix} \right)^{0:d}}{\left(\Lambda \begin{bmatrix} y \\ 1 \end{bmatrix} \right)^{d+1}}, \quad (329)$$

where $\Lambda \begin{bmatrix} y \\ 1 \end{bmatrix}$ denotes the matrix vector product of $(d+2) \times (d+2)$ -matrix Λ and the vector y with a 1 appended as the last component. Note that the indices of the components of $y \in \mathbb{S}^d \subseteq \mathbb{R}^{d+1}$ range from $i = 0, 1, \dots, d$.

Remark D.7.3. Note that we always have:

$$\text{Isom}(\mathbb{S}^d) \cong \text{O}(d+1), \quad (330)$$

via the usual action and for $d \geq 2$:

$$\text{ConfDiff}(\mathbb{S}^d) \cong \text{PO}(d+1, 1), \quad (331)$$

via the action above. The latter is part of Theorem D.4.2.

Remark D.7.4. Recall that by Proposition D.6.1, Theorem D.6.2 and Remark D.6.3 for $q = 0$ we have isometries:

$$\mathbb{A}^{d+1,0} \cong \widetilde{\text{AdS}}^{d+1,0} \cong \text{AdS}^{d+1,0}, \quad (332)$$

where we can slightly re-write¹⁶ $\widetilde{\text{AdS}}^{d+1,0}$ as:

$$\widetilde{\text{AdS}}^{d+1,0} = \{y \in \mathbb{R}^{d+1,1} \mid \eta^{d+1,1}(y, y) = -1, y^{d+1} > 0\}. \quad (333)$$

With this we get the isometry:

$$\text{AdS}^{d+1,0} \cong \widetilde{\text{AdS}}^{d+1,0} \subseteq \mathbb{R}^{d+1,1}, \quad [z] \mapsto \frac{\text{sgn}(z^{d+1})}{\sqrt{|\eta^{d+1,1}(z, z)|}} \cdot z =: y. \quad (334)$$

Note that then:

$$[y] = [z] \in \text{AdS}^{d+1} \subseteq \mathbb{P}^{d+1}. \quad (335)$$

¹⁶Note that for $y \in \mathbb{R}^{d+1,1}$ with $\eta^{d+1,1}(y, y) < 0$ we have: $|y^{d+1}|^2 > |y^0|^2$. So, for those y , we have the equivalence: $y^{d+1} > 0 \iff y^0 + y^{d+1} > 0$.

Furthermore, for $[\Lambda] \in \text{PO}(d+1, 1)$ we define the action on $y \in \widetilde{\text{AdS}}^{d+1,0}$:

$$[\Lambda].y = \text{sgn}((\Lambda y)^{d+1}) \cdot (\Lambda y), \quad (336)$$

where Λy denotes the usual matrix product. Note that this is well-defined and that:

$$[\Lambda].y \in \widetilde{\text{AdS}}^{d+1,0}, \quad [[\Lambda].y] = [\Lambda y] = [\Lambda][y] \in \text{AdS}^{d+1,0}. \quad (337)$$

This shows that the above isometry $\widetilde{\text{AdS}}^{d+1,0} \cong \text{AdS}^{d+1,0}$ is also $\text{PO}(d+1, 1)$ -equivariant.

Definition/Lemma D.7.5 (Geodesic distance of $\text{AdS}^{d+1,0}$). *The geodesic distance on $\widetilde{\text{AdS}}^{d+1,0}$, $\text{AdS}^{d+1,0}$ and $\mathbb{A}^{d+1,0}$, resp., is given by:*

$$d^{\widetilde{\text{AdS}}^{d+1,0}}(y_1, y_2) := \text{arccosh}(|\eta^{d+1,1}(y_1, y_2)|), \quad (338)$$

$$d^{\text{AdS}^{d+1,0}}([z_1], [z_2]) := \text{arccosh}\left(\frac{|\eta^{d+1,1}(z_1, z_2)|}{\sqrt{|\eta^{d+1,1}(z_1, z_1)| \cdot |\eta^{d+1,1}(z_2, z_2)|}}\right), \quad (339)$$

$$d^{\mathbb{A}^{d+1,0}}(x_1, x_2) := \text{arccosh}\left(\frac{|x_1^0|^2 + |x_2^0|^2 + \eta^{d,0}(x_1^{1:d}, x_2^{1:d})}{2x_1^0 x_2^0}\right). \quad (340)$$