Alma Mater Studiorum
Università di Bologna

DOTTORATO DI RICERCA IN MATEMATICA
Ciclo XXXIII

Settore Concorsuale: 01/A2
Settore Scientifico Disciplinare: MAT 03

On the topological theory of Group Equivariant Non-Expansive Operators

Presentata da: Nicola Quercioli

Coordinatore Dottorato:
Chiar.mo Prof.
Fabrizio Caselli

Supervisore:
Chiar.mo Prof.
Patrizio Frosini

Esame finale anno 2021

Abstract

In this thesis we aim to provide a general topological and geometrical framework for group equivariance in the machine learning context. A crucial part of this framework is a synergy between persistent homology and the theory of group actions. In our approach, instead of focusing on data, we focus on suitable operators defined on the functions that represent the data. In particular, we define group equivariant non-expansive operators (GENEOs), which are maps between function spaces endowed with the actions of groups of transformations. We investigate the topological, geometric and metric properties of the space of GENEOs. We begin by defining suitable pseudo-metrics for the function spaces, the equivariance groups, and the set of GENEOs and proving some results about our model. Basing on these pseudo-metrics, we prove that the space of GENEOs is compact and convex, under the assumption that the function spaces are compact and convex. These results provide fundamental guarantees in a machine learning perspective. We show some new methods to build different classes of GENEOs in order to populate and approximate the space of GENEOs. Moreover, we define a suitable Riemannian structure on manifolds of GENEOs making available the use of gradient descent methods.

Contents

Introduction

The main aim of this thesis is to start building a bridge between Topological Data Analysis (TDA) and a new possible geometric approach to the theory of deep neural networks. In particular, we provide a general mathematical framework for group equivariance in the machine learning context.

In the last years the problem of data analysis has assumed a more and more relevant role in science, and many researchers have started to become interested in it from several different points of view. In many applications of the real world, data sets are represented by $\mathbb{R}^m$-valued continuous functions defined on a topological space X. As simple examples among many others, these functions can describe the coloring of 3D objects, the coordinates of the points in a planar curve, or the grey-levels in X-ray CT images. However, for the sake of simplicity, in this work we will focus on real-valued functions. Topological data analysis (TDA) has revealed important in managing the huge amount of data that surrounds us in the most varied contexts [11]. Persistent topology and homology are relevant mathematical tools in TDA, and many researchers have investigated these concepts both from a theoretical and an applicative point of view (see, e.g., [8, 13, 15, 17, 18, 21, 23, 28, 35, 38, 42, 45, 46]). In addition to this, other topological techniques have been introduced in TDA (cf., e.g., [24]). The continuity of the functions we consider enables us to apply persistent homology. In its one-parameter topological setting, this theory studies the birth and the death of k-dimensional holes when we move along the filtration defined by the sublevel sets of a continuous function φ from a topological space X to the real numbers. In particular, persistent homology has proven itself quite efficient both for qualitative and topological comparison of data. Interestingly, this procedure is invariant with respect to all homeomorphisms of X, that is if $g \in \mathrm{Homeo}(X)$, then φ and φg induce on X two filtrations which have exactly the same topological properties under the point of view of persistent homology. For further and more detailed information about persistent homology, we refer the reader to [27].

Deep learning-based algorithms reached human or superhuman performance in many real-world tasks. Beyond the extreme effectiveness of deep learning, one of the main reasons for its success is that raw data are sufficient—if not even more suitable than hand-crafted features—for these algorithms to learn a specific task. However, only few attempts have been made to create formal theories allowing for the creation of a controllable and interpretable framework, in which deep neural networks can be formally defined and studied. Furthermore, if learning directly from raw data allows one to outclass human feature engineering, the architectures of deep networks are growing more and more complex, and often are as task-specific as hand-crafted features used to be. Topology and TDA can play an important role in deep learning because they are able to give qualitative and concise descriptions of the data (ref. [7, 12]).

We aim at providing a general mathematical framework, where any agent capable of acting on a certain dataset (e.g., deep neural networks) can be formally described as a collection of operators acting on the data. To motivate our model, we assume that data cannot be studied directly, but only through the action of agents that measure and transform them. Consequently, our model stems from a functional viewpoint. By interpreting data as points of a function space, it is possible to learn and optimise operators defined on the data. In other words, we are interested in the space of transformations of the data, rather than the data themselves.

Albeit unformalised, this idea is not new in deep learning. For instance, one of the main features of convolutional neural networks [37] is the election of convolution as the operator of choice to act on the data. The convolutional kernels learned by optimising a loss function are operators that map an image to a new one that, for instance, is more easily classifiable. Moreover, convolutions are operators equivariant with respect to translations (at least in the ideal continuous case). An operator is called equivariant with respect to a group if the action of the group commutes with the operator. However, when working with images, volumes or even time series, oftentimes invariance with respect to transformations such as rotations, reflexions, or other deformations is fundamental to speed up the learning process, or even to reach satisfying accuracy. Currently, data augmentation, or heavy preprocessing (e.g., accurate image alignment) are the most common strategies used to produce networks resistant to even simple data transformations. We believe that the restriction to a specific family of operators and the equivariance with respect to interpretable transformations are key aspects of the success of this architecture. In our theory, operators are thought of as instruments allowing an agent to provide a measure of the world, as the kernels learned by a convolutional neural network allow a classifier to spot essential features to recognise objects belonging to the same category. The importance of group equivariance in machine learning is well-known (cf., e.g., [2, 20, 40, 41]). The reason is twofold. On one hand, the use of operators equivariant with respect to specific transformations allows one to inject pre-existing knowledge in the system, thus gaining control of the nature of the learned operators [4]. On the other hand, equivariance with respect to the action of a group (or a set) of transformations corresponds to the introduction of symmetries in the data space, hence drastically reducing the dimensionality of the space to be explored during optimisation, and opening the way to alternative kinds of abstract representation. Indeed, it is well known that incorporating prior domain knowledge helps machine learning.

We make use of topological data analysis to describe spaces of group equivariant non-expansive operators (GENEOs). GENEOs are maps between function spaces associated with groups of transformations. We study the topological and metric properties of the space of GENEOs to evaluate their approximating power and set the basis for general strategies to initialise, compose operators and eventually connect them hierarchically to form operator networks. Moreover, the use of GENEOs allows to formalize the role of the observer in data analysis and data comparison.

The research concerning the approximation and representation of observers could be crucial in understanding the role of conflicts and contradiction in the development of intelligence [30]. Our first contribution is to define suitable pseudo-metrics for the function spaces, the equivariance groups, and the set of non-expansive operators. Basing on these pseudo-metrics, we prove that the space of GENEOs is compact and convex, under the assumption that the function spaces are compact and convex. These results provide fundamental and provable guarantees for the goodness of this operator-based approach in a machine learning perspective: Compactness, for instance, guarantees that any operator belonging to a certain space can be approximated by a finite number of operators sampled in the same space. In order to proceed along this line of research we need general methods to build and populate the space of GENEOs. According to the goal of realizing those methods, we prove some new results about the algebra of GENEOs. We underline that, in some cases, in our methods we can treat the group of invariance as a variable. This is important because the change of the observer generally corresponds to a change of the invariance we want to analyze. This focus on spaces of group equivariant non-expansive operators stresses the need for a study of the topological and geometric structure of these spaces, in order to simplify their exploration and use. In particular, in this paper we show how a space of GENEOs can be endowed with the structure of a Riemannian manifold, making available the use of gradient descent methods for the minimization of cost functions. We hope that these results about the geometry and topology of the space of GENEOs can be of help in the construction of new kinds of explainable neural networks. Finally, we observe that GENEOs can be also seen as different ways of looking at data and producing new filtrations. In this sense, their study is related to the concept of filtration functor. For more details about this concept we refer the interested reader to [19].

The structure of the dissertation is the following one. Since most of the contents have been already published in some papers, we cite the corresponding references.

Chapter 1 is devoted to illustrating a new model for TDA. First, we start from a set Φ of suitable real-valued functions on a set X, endowed with the uniform convergence distance. This set Φ represents our data set. Then, we consider particular sets of bijections on X, which model symmetries in our framework. Finally, we define some suitable pseudo-metrics and prove some general results, most of which can be found in [5, 16].

In Chapter 2 we introduce the concept of Group Equivariant Non-Expansive Operator (GENEO). Then we prove, under suitable assumptions, compactness and convexity of the space of GENEOs. In the last part of the chapter we combine GENEOs and persistent homology to define a new strongly invariant pseudo-metric between data. Moreover, by means of persistent homology we define faster to compute and stable approximations of the pseudo-metrics defined in Chapter 1. Most of the contents of Chapter 2 can be found in [5].

In Chapter 3 we present some methods to build classes of GENEOs. In particular, we illustrate methods that combine a finite or infinite set of given GENEOs in order to produce a new one. Then, we construct an important class of linear GENEOs via permutants and permutant measures. Finally, we prove a result about representations of linear GENEOs. The majority of the results of this chapter appeared in [6, 10, 33, 44].

In Chapter 4 we start from defining a new slightly different framework which heavily involves measure and probability theory. Then we define a natural Hilbert space where GENEOs live. This enables us to induce a Riemannian structure on manifolds of GENEOs. The contents of this chapter are a joint work with Pasquale Cascarano, Patrizio Frosini and Amir Saki. Most of the contents of Chapter 4 can be found in [14].

We conclude the dissertation with some final remarks and a short list of open problems in our line of research.

Epistemological assumptions

Our mathematical model is justified by an epistemological background which revolves around the following assumptions:

1. Data are represented as functions defined on topological spaces, since only data that are stable with respect to a certain criterion (e.g., with respect to some kind of measurement) can be considered for applications, and stability requires a topological structure.

2. Data cannot be studied in a direct and absolute way. They are only knowable through acts of transformation made by an agent. From the point of view of data analysis, only the pair (data, agent) matters. In general terms, agents are not endowed with purposes or goals: they are just ways and methods to transform data. Acts of measurement are a particular class of acts of transformation.

3. Agents are described by the way they transform data while respecting some kind of invariance. In other words, any agent can be seen as a group equivariant operator acting on a function space.

4. Data similarity depends on the output of the considered agent.

In other words, in our framework we assume that the analysis of data is replaced by the analysis of the pair (data, agent). Since an agent can be seen as a group equivariant operator, from the mathematical viewpoint our purpose consists of presenting a good topological theory of suitable operators of this kind, representing agents. For more details, we refer the interested reader to [31].

Chapter 1

A new mathematical model for TDA

This chapter is devoted to illustrating our new model for topological data analysis and prove some results about the topology and geometry of the introduced concepts. First, for us a data set is given by a set of bounded real-valued measurements on some set X:

$$\Phi = \{\varphi\colon X \to \mathbb{R}\} \subseteq \mathbb{R}_b^X,$$

where $\mathbb{R}_b^X$ is the set of all bounded real-valued functions on X. We can think of X as the space where one makes measurements, and of Φ as the set of admissible measurements. For example, an image can be represented as a function φ from the real plane X to the real numbers. Furthermore, the set X is called the **domain** of the data set Φ and is denoted by $\mathrm{dom}(\Phi)$.

1.1 Transformations on data and actions

In our model, agents transform data preserving important invariances and symmetries. In order to describe those symmetries of a data set Φ, we consider operations on X that convert measurements into measurements.

Definition 1.1.1. A **Φ-operation** is a function $g\colon X \to X$ such that, for every measurement φ in Φ, the composition φg also belongs to Φ.

If $g\colon X \to X$ is such an operation, then, for all φ and ψ in Φ:

$$\|\varphi - \psi\|_\infty = \max_{x \in X}|\varphi(x) - \psi(x)| \geq \max_{x \in \mathrm{im}(g)}|\varphi(x) - \psi(x)| = \|\varphi g - \psi g\|_\infty.$$

Thus the function $R_g\colon \Phi \to \Phi$ that maps φ to φg is non-expansive. If g is bijective, $R_g\colon \Phi \to \Phi$ is an isometry. The composition of Φ-operations is again a Φ-operation, and the identity function id_X is always a Φ-operation.

A Φ-operation g is invertible if there is a Φ-operation h such that $gh = hg = \mathrm{id}_X$. Moreover, we can consider the collection of all invertible Φ-operations that is denoted by

$$\mathrm{Aut}_\Phi(X) = \{g : X \to X \mid \ g \text{ is a bijection, and } \varphi g, \varphi g^{-1} \in \Phi \text{ for every } \varphi \in \Phi\}.$$

With the composition operation, $\mathrm{Aut}_\Phi(X)$ becomes a group.

A data set Φ is naturally equipped with an associative right action:

$$\rho\colon \Phi \times \mathrm{Aut}_\Phi(X) \to \Phi, \qquad (\varphi, g) \mapsto \varphi g \tag{1.1}$$

where φg is the usual function composition. Thus Φ is not just a set, but a set with an action of the group $\mathrm{Aut}_\Phi(X)$. In other words, Φ is a $\mathrm{Aut}_\Phi(X)$-set. To encode the symmetries of Φ induced by this action, we consider its perception pairs.

Definition 1.1.2. A **perception pair** is a couple (Φ, G) with $\Phi \subseteq \mathbb{R}_b^X$ and $G \subseteq \mathrm{Aut}_\Phi(X)$.

The choice of G encodes certain symmetries of Φ. The pair $(\Phi, \mathrm{Aut}_\Phi(X))$ is an example of a perception pair, and we call it universal.

Example 1.1.3. Let $X = S^1$ be the unit circle and Φ be the set of all non-expansive functions from S^1 to the unit interval $[0, 1]$, where both S^1 and $[0, 1]$ are endowed with the Euclidean metric. Consider the group G of all rotations of S^1. One could easily check that $G \subseteq \mathrm{Aut}_\Phi(X)$. Then (Φ, G) is a perception pair.

Example 1.1.4. Let us consider $X = \{1, \ldots, n\}$ and the space $\mathbb{R}^X$ of real-valued functions on X. We observe that $\mathrm{Aut}_{\mathbb{R}^X}(X) = S_n$, where S_n is the set of permutations of X. Then $(\mathbb{R}^X, S_n)$ is a perception pair.

1.2 Topologies on data sets

In our mathematical model, data are represented as function spaces, that is, as sets Φ of bounded real-valued functions $\varphi : X \to \mathbb{R}$, with X a set. A natural choice for the topology on Φ is the topology of uniform convergence that is induced by the distance

$$D_\Phi(\varphi_1, \varphi_2) := \|\varphi_1 - \varphi_2\|_\infty . \tag{1.2}$$

The topological structure of the set X is inherited by the extended pseudo-metric D_X defined as follows:

$$D_X(x_1, x_2) = \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)| \tag{1.3}$$

for every $x_1, x_2 \in X$.

Proposition 1.2.1 *D_X is an extended pseudo-metric.*

Proof. D_X is obviously symmetrical. The definition of D_X immediately implies that $D_X(x, x) = 0$ for any $x \in X$. The triangle inequality holds, since

$$\begin{aligned} D_X(x_1, x_2) &= \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)| \\ &\leq \sup_{\varphi \in \Phi} (|\varphi(x_1) - \varphi(x_3)| + |\varphi(x_3) - \varphi(x_2)|) \\ &\leq \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_3)| + \sup_{\varphi \in \Phi} |\varphi(x_3) - \varphi(x_2)| \\ &= D_X(x_1, x_3) + D_X(x_3, x_2) \end{aligned} \tag{1.4}$$

for any $x_1, x_2, x_3 \in X$. □

Remark 1.2.2. The assumption behind the definition of D_X is that two points can be distinguished only if they are mapped to different values for some measurements. As an example, if Φ contains only constant functions and X contains at least two points, no discrimination can be made between points in X and hence $D_X(x_1, x_2)$ vanishes for every $x_1, x_2 \in X$. Note that in this case X is not even a T_0-space.

The extended pseudo-metric space (X, D_X) can be considered as a topological space by choosing as a base $\mathcal{B}_{D_X}$ the collection of all the sets

$$B_X(x, \varepsilon) = \{x' \in X : D_X(x, x') < \varepsilon\} \tag{1.5}$$

where $0 < \varepsilon < \infty$, and $x \in X$ (see [34]). We recall that a pseudo-metric is just a distance d without the property that $d(a, b) = 0$ implies $a = b$. An extended pseudo-metric is simply a pseudo-metric that could take an infinite value.

Remark 1.2.3. If Φ is bounded, then the value $D_X(x_1, x_2)$ is finite for every $x_1, x_2 \in X$. Indeed, a finite constant L exists such that $\|\varphi\|_\infty \leq L$ for every $\varphi \in \Phi$. Hence,

$$|\varphi(x_1) - \varphi(x_2)| \leq \|\varphi\|_\infty + \|\varphi\|_\infty \leq 2L \tag{1.6}$$

for any $\varphi \in \Phi$ and any $x_1, x_2 \in X$. This implies that $D_X(x_1, x_2) \leq 2L$ for every $x_1, x_2 \in X$. Thus, if Φ is bounded, D_X is a pseudo-metric.

Remark 1.2.4. Suppose that X is a subspace of $\mathbb{R}$ endowed with the Euclidean topology. The reason to consider a topological space X rather than considering just a set follows from the need of formalising the assumption that data are stable. Note that our choice of topology allows to deal with non-continuous functions, with respect to the Euclidean topology.

Remark 1.2.5. In general X is not compact with respect to the topology induced by D_X, even if Φ is compact. For example, if X is the open interval $]0, 1[$ and Φ contains only the identity from $]0, 1[$ to $]0, 1[$, the topology induced by D_X is simply the Euclidean topology and hence X is not compact.

Before proceeding, we would like to recall the following Lemma (see [34]):

Lemma 1.2.6 *Let (P, d) be a pseudo-metric space. The following conditions are equivalent:*

1. *P is totally bounded;*
2. *every sequence in P admits a Cauchy subsequence.*

We have seen that X can be non-compact. However, the following statement holds.

Theorem 1 *If Φ is totally bounded, then (X, D_X) is totally bounded.*

Proof. First of all we want to prove that every sequence $(x_i)_{i\in\mathbb{N}}$ in X admits a Cauchy subsequence in X. Let us consider an arbitrary sequence $(x_i)_{i\in\mathbb{N}}$ in X and an arbitrarily small $\varepsilon > 0$. Since Φ is totally bounded, we can find a finite subset $\Phi_\varepsilon = \{\varphi_1, \ldots, \varphi_n\}$ such that $\Phi = \bigcup_{i=1}^n B_\Phi(\varphi_i, \varepsilon)$, where $B_\Phi(\varphi, \varepsilon) = \{\varphi' \in \Phi : D_\Phi(\varphi', \varphi) < \varepsilon\}$. In particular, we can say that for any $\varphi \in \Phi$ there exists $\varphi_{\bar{k}} \in \Phi_\varepsilon$ such that $\|\varphi - \varphi_{\bar{k}}\|_\infty < \varepsilon$. Now, we consider the real sequence $(\varphi_1(x_i))_{i\in\mathbb{N}}$ that is bounded because all the functions in Φ are bounded. From Bolzano-Weierstrass Theorem it follows that we can extract a convergent subsequence $(\varphi_1(x_{i_h}))_{h\in\mathbb{N}}$. Then we consider the sequence $(\varphi_2(x_{i_h}))_{h\in\mathbb{N}}$. Since φ_2 is bounded, we can extract a convergent subsequence $(\varphi_2(x_{i_{h_t}}))_{t\in\mathbb{N}}$. We can repeat the same argument for any $\varphi_k \in \Phi_\varepsilon$. Thus, we obtain a subsequence $(x_{p_j})_{j\in\mathbb{N}}$ of $(x_i)_{i\in\mathbb{N}}$, such that $(\varphi_k(x_{p_j}))_{j\in\mathbb{N}}$ is a real convergent sequence for any $k \in \{1,\ldots,n\}$, and hence a Cauchy sequence in $\mathbb{R}$. Moreover, since Φ_ε is a finite set, there exists an index $\bar{\jmath}$ such that for any $k \in \{1,\ldots,n\}$ we have that

$$|\varphi_k(x_{p_r}) - \varphi_k(x_{p_s})| < \varepsilon, \quad \text{for all } r,s \geq \bar{\jmath}. \tag{1.7}$$

We observe that $\bar{\jmath}$ does not depend on k, but only on ε and Φ_ε.

In order to prove that $(x_{p_j})_{j\in\mathbb{N}}$ is a Cauchy sequence in X, we observe that for any $r,s \in \mathbb{N}$ and any $\varphi \in \Phi$, by choosing a k such that $\|\varphi - \varphi_k\|_\infty < \varepsilon$ we have:

$$\begin{aligned}|\varphi(x_{p_r}) - \varphi(x_{p_s})| &= |\varphi(x_{p_r}) - \varphi_k(x_{p_r}) + \varphi_k(x_{p_r}) - \varphi_k(x_{p_s}) + \varphi_k(x_{p_s}) - \varphi(x_{p_s})| \\ &\leq |\varphi(x_{p_r}) - \varphi_k(x_{p_r})| + |\varphi_k(x_{p_r}) - \varphi_k(x_{p_s})| + |\varphi_k(x_{p_s}) - \varphi(x_{p_s})| \\ &\leq \|\varphi - \varphi_k\|_\infty + |\varphi_k(x_{p_r}) - \varphi_k(x_{p_s})| + \|\varphi_k - \varphi\|_\infty. \end{aligned} \tag{1.8}$$

It follows that $|\varphi(x_{p_r}) - \varphi(x_{p_s})| < 3\varepsilon$ for every $\varphi \in \Phi$ and every $r,s \geq \bar{\jmath}$. Thus, $\sup_{\varphi\in\Phi} |\varphi(x_{p_r}) - \varphi(x_{p_s})| = D_X(x_{p_r}, x_{p_s}) \leq 3\varepsilon$. Hence, the sequence $(x_{p_j})_{j\in\mathbb{N}}$ is a Cauchy sequence in X. By Lemma 1.2.6, the statement of Theorem 1 is true. □

Corollary 1.2.7 *If Φ is totally bounded and (X, D_X) is complete, then (X, D_X) is compact.*

Example 1.2.8. Let Φ be the set containing all the non-expansive functions from

$$X = \{(\cos 2\pi p, \sin 2\pi p) \in \mathbb{R}^2 : p \in \mathbb{Q}\}$$

to $[0,1]$, and G be the group of all rotations $\rho_{2\pi q}$ of $2\pi q$ radians with $q \in \mathbb{Q}$. While Φ is compact, the topological space X is neither complete nor compact.

It is interesting to stress the link between the topology τ_{D_X} associated with D_X and the initial topology τ_{in} on X with respect to Φ, when we take the Euclidean topology τ_e on $\mathbb{R}$. We recall that τ_{in} is the coarsest topology on X such that each function $\varphi \in \Phi$ is continuous. Before proceeding, we need to prove a technical lemma about the approximation of the pseudo-distance D_X.

Lemma 1.2.9 *If Φ is totally bounded, then for any $\delta > 0$ there exists a finite subset Φ_δ of Φ such that*

$$\left| \sup_{\varphi\in\Phi} |\varphi(x_1) - \varphi(x_2)| - \max_{\varphi\in\Phi_\delta} |\varphi(x_1) - \varphi(x_2)| \right| \leq 2\delta \tag{1.9}$$

for every $x_1, x_2 \in X$.

Proof. Since Φ is totally bounded, we can find a finite subset $\Phi_\delta = \{\varphi_1, \ldots, \varphi_n\}$ such that for each $\varphi \in \Phi$ there exists $\varphi_i \in \Phi_\delta$, for which $\|\varphi - \varphi_i\|_\infty < \delta$. It follows that for any $x \in X$, $|\varphi(x) - \varphi_i(x)| < \delta$. Without loss of generality, let us fix

$x_1, x_2 \in X$. Because of the definition of supremum of a subset of the set $\mathbb{R}^+$ of all positive real numbers, for any $\varepsilon > 0$ we can choose a $\varphi_\varepsilon \in \Phi$ such that

$$\sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)| - |\varphi_\varepsilon(x_1) - \varphi_\varepsilon(x_2)| \le \varepsilon. \tag{1.10}$$

For any $\varphi \in \Phi$, if we take an index i, for which $\|\varphi - \varphi_i\|_\infty < \delta$, we have that:

$$\begin{aligned} |\varphi(x_1) - \varphi(x_2)| &= |\varphi(x_1) - \varphi_i(x_1) + \varphi_i(x_1) - \varphi_i(x_2) + \varphi_i(x_2) - \varphi(x_2)| \\ &\le |\varphi(x_1) - \varphi_i(x_1)| + |\varphi_i(x_1) - \varphi_i(x_2)| + |\varphi_i(x_2) - \varphi(x_2)| \\ &< |\varphi_i(x_1) - \varphi_i(x_2)| + 2\delta \\ &\le \max_{\varphi_j \in \Phi_\delta} |\varphi_j(x_1) - \varphi_j(x_2)| + 2\delta. \end{aligned} \tag{1.11}$$

We note that the above inequality

$$|\varphi(x_1) - \varphi(x_2)| < \max_{\varphi_j \in \Phi_\delta} |\varphi_j(x_1) - \varphi_j(x_2)| + 2\delta$$

does not depend on the choice of the index i. Hence,

$$\sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)| - \varepsilon \le |\varphi_\varepsilon(x_1) - \varphi_\varepsilon(x_2)| < \max_{\varphi_j \in \Phi_\delta} |\varphi_j(x_1) - \varphi_j(x_2)| + 2\delta. \tag{1.12}$$

Finally, as ε goes to zero, we have that

$$\sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)| \le \max_{\varphi_j \in \Phi_\delta} |\varphi_j(x_1) - \varphi_j(x_2)| + 2\delta. \tag{1.13}$$

On the other hand, since $\Phi_\delta \subseteq \Phi$:

$$\sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)| > \max_{\varphi_j \in \Phi_\delta} |\varphi_j(x_1) - \varphi_j(x_2)| - 2\delta. \tag{1.14}$$

Therefore we proved the statement. □

The choice of the topology on X induced by D_X naturally makes our signals continuous.

Proposition 1.2.10 *Each element φ of Φ is a non-expansive map, and hence it is continuous with respect to D_X.*

Proof. Let us fix two point x_1, x_2 in X. For any $\varphi \in \Phi$ we have that

$$|\varphi(x_1) - \varphi(x_2)| \le \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)| = D_X(x_1, x_2). \tag{1.15}$$

Hence the statement is proved. □

We denote with τ_{in} the initial topology with rispect to Φ. By definition, τ_{in} is the coarsest topology on X such that each function $\varphi \in \Phi$ is continuous.

Theorem 2 *If Φ is totally bounded, then the topology τ_{D_X} coincides with τ_{in}.*

Proof. We recall each element φ of Φ is a continuous function with respect to τ_{D_X}. Thus we can immediately state that τ_{D_X} is finer than τ_{in}. We know that the set $\mathcal{B}_{D_X} = \{B_X(x,\varepsilon) : x \in X, \varepsilon > 0\}$ is a base for the topology τ_{D_X} and the set $\mathcal{B}_{\text{in}} = \left\{\bigcap_{i\in I} \varphi_i^{-1}(U_i) : |I| < \infty, U_i \in \tau_e, \varphi_i \in \Phi \text{ for every } i \in I\right\}$ is a base for the topology τ_{in}.

If Φ is totally bounded, Lemma 1.2.9 guarantees that for every $\delta > 0$ a finite subset Φ_δ of Φ exists such that

$$\left|\sup_{\varphi\in\Phi} |\varphi(x_1) - \varphi(x_2)| - \max_{\varphi\in\Phi_\delta} |\varphi(x_1) - \varphi(x_2)|\right| \leq 2\delta \tag{1.16}$$

for every $x_1, x_2 \in X$. Let us now set

$$B_\delta(x,r) := \left\{x' \in X \middle| \max_{\varphi_i\in\Phi_\delta} |\varphi_i(x) - \varphi_i(x')| < r\right\} \tag{1.17}$$

for every $x \in X$ and every $r > 0$. We have to prove that the initial topology τ_{in} is finer than the topology τ_{D_X}. In order to do this, it will be sufficient to show that for every $y \in B_X(x,\varepsilon) \in \mathcal{B}_{D_X}$ a set $\bigcap_{i\in I} \varphi_i^{-1}(U_i) \in \mathcal{B}_{\text{in}}$ exists, such that $y \in \bigcap_{i\in I} \varphi_i^{-1}(U_i) \subseteq B_X(x,\varepsilon)$.

Let us choose a positive δ such that $2\delta < \varepsilon$. Inequality (1.16) implies that $B_\delta(y,\varepsilon-2\delta) \subseteq B_X(y,\varepsilon)$. We now set $U_i :=]\varphi_i(y) - \varepsilon + 2\delta, \varphi_i(y) + \varepsilon - 2\delta[$ for $i \in I$. Obviously, $y \in \bigcap_{\varphi_i\in\Phi_\delta} \varphi_i^{-1}(U_i)$. If $z \in \bigcap_{\varphi_i\in\Phi_\delta} \varphi_i^{-1}(U_i)$, then $|\varphi_i(z) - \varphi_i(y)| < \varepsilon - 2\delta$ for every $\varphi_i \in \Phi_\delta$. Hence, $z \in B_\delta(y,\varepsilon-2\delta)$. It follows that $\bigcap_{\varphi_i\in\Phi_\delta} \varphi_i^{-1}(U_i) \subseteq B_\delta(y,\varepsilon-2\delta)$. Therefore, $y \in \bigcap_{\varphi_i\in\Phi_\delta} \varphi_i^{-1}(U_i) \subseteq B_X(x,\varepsilon)$ because of the inclusion $B_\delta(y,\varepsilon-2\delta) \subseteq B_X(y,\varepsilon)$. This means that τ_{in} is finer than τ_{D_X}. Since we already know that τ_{D_X} is finer than τ_{in}, it follows that τ_{D_X} coincides with τ_{in}. □

Remark 1.2.11. The statement of Theorem 2 becomes false if Φ is not totally bounded. For example, assume Φ equal to the set of all functions from $X = [0,1]$ to $\mathbb{R}$ that are continuous with respect to the Euclidean topologies on $[0,1]$ and $\mathbb{R}$. Indeed, it is easy to check that in this case τ_{D_X} is the discrete topology, while the initial topology τ_{in} is the Euclidean topology on $[0,1]$.

1.2.1 Transformations on data and topological actions

We recall that the definition of isometry between extended pseudo-metric spaces can be derived by generalising the concept of isometry between metric spaces. Let (X_1, d_1) and (X_2, d_2) be two extended pseudo-metric spaces. It is easy to check that if $f : X_1 \to X_2$ is a function verifying the equality $d_1(x,y) = d_2(f(x), f(y))$ for every $x, y \in X_1$, then f is continuous with respect to the topologies induced by d_1 and d_2. If f verifies the previous equality and is bijective, we say that it is an *isometry* between the considered pseudo-metric spaces. Let $\text{Iso}(X)$ the set of all isometries from X to itself. The following Proposition 1.2.12 implies that $\text{Aut}_\Phi(X)$ is exactly the set of all isometries $g\colon X \to X$ that are also invertible Φ-operations.

Proposition 1.2.12 $\text{Aut}_\Phi(X) \subseteq \text{Iso}(X)$.

Proof. Let us fix two arbitrary points x, x' in X. We know that if g is in $\mathrm{Aut}_\Phi(X)$, then $R_g\colon \Phi \to \Phi$ is a bijection. Therefore, g preserves the extended pseudo-distance D_X:

$$\begin{aligned} D_X(g(x), g(x')) &= \sup_{\varphi\in\Phi} |\varphi(g(x)) - \varphi(g(x'))| \\ &= \sup_{\varphi\in\Phi} |(\varphi g)(x) - (\varphi g)(x')| \qquad (1.18) \\ &= \sup_{\varphi\in\Phi g} |\varphi(x) - \varphi(x')| \qquad (1.19) \\ &= \sup_{\varphi\in\Phi} |\varphi(x) - \varphi(x')| = D_X(x, x'). \end{aligned}$$

Since g is bijective, it follows that g is an isometry with respect to D_X. □

Corollary 1.2.13 *The invertible Φ-operations are exactly the isometries of X that preserve Φ by composition on the right.*

Remark 1.2.14. In general, $\mathrm{Aut}_\Phi(X) \neq \mathrm{Iso}(X)$. As an example, take $X = [-1, 1]$ and $\Phi = \{\mathrm{id}_X\}$. In this case $D_X(x_1, x_2) = |x_1 - x_2|$ and $\mathrm{Aut}_\Phi(X) = \{\mathrm{id}_X\}$, while $\mathrm{Iso}(X) = \{\mathrm{id}_X, r\colon x \mapsto -x\}$.

Now we are ready to put more structure on $\mathrm{Aut}_\Phi(X)$. In our model every comparison must be based on the max-norm distance between admissible acts of measurement. As a consequence, we define the distance between two invertible Φ-operations as the difference of their actions on the set Φ of possible measurements. Hence, we consider the pseudo-distance D_{Aut} on $\mathrm{Aut}_\Phi(X)$, i.e.

$$D_{\mathrm{Aut}}(g_1, g_2) := \sup_{\varphi\in\Phi} D_\Phi(\varphi g_1, \varphi g_2).$$

for any g_1, g_2 in $\mathrm{Aut}_\Phi(X)$. We observe that

$$\begin{aligned} D_{\mathrm{Aut}}(g_1, g_2) &:= \sup_{\varphi\in\Phi} D_\Phi(\varphi g_1, \varphi g_2) \\ &= \sup_{x\in X} \sup_{\varphi\in\Phi} |\varphi(g_1(x)) - \varphi(g_2(x))| \\ &= \sup_{x\in X} D_X(g_1(x), g_2(x)) \qquad (1.20) \end{aligned}$$

for any g_1, g_2 in $\mathrm{Aut}_\Phi(X)$. Thus, D_{Aut} coincides with the pseudo-metric induced by the uniform convergence on $\mathrm{Aut}_\Phi(X)$.

In our model every comparison must be based on the max-norm distance between admissible acts of measurement. As a consequence, we define the distance between two homeomorphisms as the difference of their actions on the set Φ of possible measurements.

Lemma 1.2.15 *Let $f, g, h \in \mathrm{Aut}_\Phi(X)$. It holds that*

$$D_{\mathrm{Aut}}(f, g) = D_{\mathrm{Aut}}(hf, hg) = D_{\mathrm{Aut}}(fh, gh).$$

Proof. Consider $f, g, h \in \mathrm{Aut}_\Phi(X)$. Since R_h is an isometry, we have that:

$$\begin{aligned} D_{\mathrm{Aut}}(fh, gh) &:= \sup_{\varphi\in\Phi} \|\varphi f h - \varphi g h\|_\infty \\ &= \sup_{\varphi\in\Phi} \|\varphi f - \varphi g\|_\infty \\ &= D_{\mathrm{Aut}}(f, g). \end{aligned}$$

Since h is an isometry, it follows that:

$$\begin{aligned} D_{\mathrm{Aut}}(hf, hg) &= \sup_{x\in X} D_X(h(f(x)), h(g(x)) \\ &= \sup_{x\in X} D_X(f(x), g(x)) \\ &= D_{\mathrm{Aut}}(hf, hg). \end{aligned}$$

Hence, the claim immediately follows. □

We can now state the following theorem:

Theorem 3 *The following statements hold:*

- $\mathrm{Aut}_\Phi(X)$ *is a topological group with respect to the topology induced by* D_{Aut}*;*
- *the action of* $\mathrm{Aut}_\Phi(X)$ *on* Φ *is continuous.*

Proof. First we will prove that the binary operation is continuous. We consider the sum pseudo-distance on $\mathrm{Aut}_\Phi(X) \times \mathrm{Aut}_\Phi(X)$, which induces the product topology. Consider $(g_1, g_2), (g_1', g_2') \in \mathrm{Aut}_\Phi(X) \times \mathrm{Aut}_\Phi(X)$. Using Lemma 1.2.15, we have that

$$\begin{aligned} D_{\mathrm{Aut}}(g_1 g_2, g_1' g_2') &= D_{\mathrm{Aut}}(g_1, g_1' g_2' g_2^{-1}) \\ &\leq D_{\mathrm{Aut}}(g_1, g_1') + D_{\mathrm{Aut}}(g_1', g_1' g_2' g_2^{-1}) \\ &= D_{\mathrm{Aut}}(g_1, g_1') + D_{\mathrm{Aut}}(id_X, g_2' g_2^{-1}) \\ &= D_{\mathrm{Aut}}(g_1, g_1') + D_{\mathrm{Aut}}(g_2, g_2'). \end{aligned} \tag{1.21}$$

It follows that the binary operation in $\mathrm{Aut}_\Phi(X)$ is non-expansive, and hence it is continuous. We also have to prove that the inverse operation is continuous. Note that the inverse operation is a bijection from $\mathrm{Aut}_\Phi(X)$ to itself. Consider $h_1, h_2 \in \mathrm{Aut}_\Phi(X)$. Because of Lemma 1.2.15, we obtain that

$$\begin{aligned} D_{\mathrm{Aut}}(h_1^{-1}, h_2^{-1}) &= D_{\mathrm{Aut}}(h_1^{-1} h_2, h_2^{-1} h_2) \\ &= D_{\mathrm{Aut}}(h_1^{-1} h_2, id_X) \\ &= D_{\mathrm{Aut}}(h_1^{-1} h_2, h_1^{-1} h_1) \\ &= D_{\mathrm{Aut}}(h_1, h_2). \end{aligned} \tag{1.22}$$

Thus, the function mapping the elements of $\mathrm{Aut}_\Phi(X)$ to their respective inverses is an isometry. We have now to prove that the action of G on Φ through right composition is continuous. We have that

$$\begin{aligned} D_\Phi(\varphi f, \psi g) &\le D_\Phi(\varphi f, \varphi g) + D_\Phi(\varphi g, \psi g) \\ &= D_\Phi(\varphi f, \varphi g) + D_\Phi(\varphi, \psi) \\ &\le D_{\mathrm{Aut}}(f, g) + D_\Phi(\varphi, \psi) \end{aligned} \quad (1.23)$$

This proves that the action of G on Φ through right composition is continuous.

□

We can now give a result about the compactness of a group $G \subseteq \mathrm{Aut}_\Phi(X)$ by means of a suitable hypothesis on the space Φ.

Theorem 4 *If Φ is totally bounded, then (G, D_{Aut}) is totally bounded.*

Proof. Let $(g_i)_{i\in\mathbb{N}}$ be a sequence in G and take a real number $\varepsilon > 0$. Given that Φ is totally bounded, we can find a finite subset $\Phi_\varepsilon = \{\varphi_1, \ldots, \varphi_n\}$ such that for every $\varphi \in \Phi$ there exists $\varphi_h \in \Phi_\varepsilon$ for which $D_\Phi(\varphi_h, \varphi) < \varepsilon$.

Let us consider the sequence $(\varphi_1 g_i)_{i\in\mathbb{N}}$ in Φ. Since Φ is totally bounded, from Lemma 1.2.6 it follows that we can extract a Cauchy subsequence $(\varphi_1 g_{i_h})_{h\in\mathbb{N}}$. Then we consider the sequence $(\varphi_2 g_{i_h})_{h\in\mathbb{N}}$. Again, we can extract a Cauchy subsequence $(\varphi_2 g_{i_{h_t}})_{t\in\mathbb{N}}$. We can repeat the same argument for any $\varphi_k \in \Phi_\varepsilon$. Thus, we are able to extract a subsequence $(g_{i_j})_{j\in\mathbb{N}}$ of $(g_i)_{i\in\mathbb{N}}$ such that $(\varphi_k g_{i_j})_{j\in\mathbb{N}}$ is a Cauchy sequence for any $k \in \{1, \ldots, n\}$. For the finiteness of set Φ_ε, we can find an index $\bar{\jmath}$ such that for any $k \in \{1, \ldots, n\}$

$$D_\Phi(\varphi_k g_{i_r}, \varphi_k g_{i_s}) < \varepsilon, \text{ for every } s, r \ge \bar{\jmath}. \quad (1.24)$$

In order to prove that $(g_{i_j})_{j\in\mathbb{N}}$ is a Cauchy sequence, we observe that for any $\varphi \in \Phi$, any $\varphi_k \in \Phi_\varepsilon$, and any $r, s \in \mathbb{N}$ we have

$$\begin{aligned} D_\Phi(\varphi g_{i_r}, \varphi g_{i_s}) &\le D_\Phi(\varphi g_{i_r}, \varphi_k g_{i_r}) + D_\Phi(\varphi_k g_{i_r}, \varphi_k g_{i_s}) + D_\Phi(\varphi_k g_{i_s}, \varphi g_{i_s}) \\ &= D_\Phi(\varphi, \varphi_k) + D_\Phi(\varphi_k g_{i_r}, \varphi_k g_{i_s}) + D_\Phi(\varphi_k, \varphi). \end{aligned} \quad (1.25)$$

We observe that $\bar{\jmath}$ does not depend on φ, but only on ε and Φ_ε. By choosing a $\varphi_k \in \Phi_\varepsilon$ such that $D_\Phi(\varphi_k, \varphi) < \varepsilon$, we get $D_\Phi(\varphi g_{i_r}, \varphi g_{i_s}) < 3\varepsilon$ for every $\varphi \in \Phi$ and every $r, s \ge \bar{\jmath}$. Thus, $D_{\mathrm{Aut}}(g_{i_r}, g_{i_s}) \le 3\varepsilon$. Hence, the sequence $(g_{i_j})_{j\in\mathbb{N}}$ is a Cauchy sequence. Therefore, by Lemma 1.2.6, G is totally bounded. □

Corollary 1.2.16 *If Φ is totally bounded and (G, D_{Aut}) is complete, then (G, D_{Aut}) is compact.*

Remark 1.2.17. Let Φ be the set containing all the non-expansive functions from $X = S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ to $[0, 1]$, and G be the group of all rotations $\rho_{2\pi q}$ of X of $2\pi q$ radians with q rational number. The space (G, D_{Aut}) is neither complete nor compact.

For the rest of the section, we set a compact space $\Phi \subseteq \mathbb{R}^X_b$. We also assume that (X, D_X) is a complete metric space. By Corollary 1.2.7, X is compact. With these assumptions we can say something more about the compactness of $\mathrm{Aut}_\Phi(X)$. Consider the group $\mathrm{Iso}(X)$ of all isometries from X to itself. Before stating the compactness of $\mathrm{Iso}(X)$, we have to recall the following lemma about isometries in metric spaces:

Lemma 1.2.18 *Let (X, D_X) be a compact metric space. Consider a function $f\colon X \to X$. f preserves distances if and only if f is an isometry.*

Proof. Preserving distances directly implies that f is continuous and injective. It will suffice to prove that f is surjective. Let x_0 be a point in X. Consider the sequence $(x_n)_{n\in\mathbb{N}}$ given by $x_{n+1} = f(x_n)$. By compactness of $f(X)$, there exists a subsequence $(x_{n_i})_{i\in\mathbb{N}}$ that is convergent and, hence, it is a Cauchy sequence.

Let ε be a positive real number. There exists a natural number N such that $D_X(x_{n_i}, x_{n_j}) < \varepsilon$ for every i, j greater than N. If $n_j \geq n_i$, then $D_X(x_{n_i}, x_{n_j}) = D_X(x_0, x_{n_j - n_i})$, since f preserves distances. Then

$$D_X(x_0, f(X)) := \inf_{x\in f(X)} D_X(x_0, x) \leq D_X(x_0, x_{n_j-n_i}) = D_X(x_{n_i}, x_{n_j}) < \varepsilon.$$

Since ε was arbitrary, we conclude $D_X(x_0, f(X)) = 0$. Since $f(X)$ is closed, then x_0 belongs to $f(X)$. □

Consider the group $\mathrm{Iso}(X)$ of isometries from (X, D_X) to itself. On $\mathrm{Iso}(X)$ we consider the metric structure induced by d_∞ defined as

$$d_\infty(f, g) := \sup_{x\in X} D_X\left(f(x), g(x)\right),$$

for any f and g belonging to $\mathrm{Iso}(X)$. We recall that on $\mathrm{Aut}_\Phi(X)$ the metric d_∞ coincides with D_{Aut}.

Remark 1.2.19. Let (X, D_X) be a compact metric space. Assume $(f_i)_{i\in\mathbb{N}}$ is a sequence in the space $C(X, X)$ of all continuous functions from X to itself. If $(f_i)_{i\in\mathbb{N}}$ converges to f with respect to d_∞, then for any x in X, $f(x) = \lim_{i\to\infty} f_i(x)$ with respect to D_X. Indeed, we have that

$$D_X\left(f(x), \lim_{i\to\infty}(f_i(x))\right) = \lim_{i\to\infty} D_X(f(x), f_i(x)) \leq \lim_{i\to\infty} d_\infty(f, f_i) = 0.$$

Proposition 1.2.20 *Let (X, D_X) be a compact metric space. Then $\mathrm{Iso}(X)$ is compact with respect to the topology induced by the uniform distance d_∞.*

Proof. Ascoli-Arzelà Theorem (see [39]) implies that $\mathrm{Iso}(X)$ is a relatively compact set. In order to state compactness it will suffice to show that $\mathrm{Iso}(X)$ is a closed subset of $C(X, X)$. Let be $(f_i)_{i\in\mathbb{N}}$ a convergent sequence of isometries. Assume that the limit point is $f \in C(X, X)$. It will suffice to show that f preserves distances. For any x and y in X, by recalling Remark 1.2.19, we have that:

$$\begin{aligned} D_X(f(x), f(y)) &= D_X(\lim_{i\to\infty} f_i(x), \lim_{i\to\infty} f_i(y)) \\ &= \lim_{i\to\infty} D_X(f_i(x), f_i(y)) \\ &= \lim_{i\to\infty} D_X(x, y) \\ &= D_X(x, y). \end{aligned}$$

□

Let us consider the set $\mathcal{H}$ of all non-empty compact subsets of the space $\mathbf{NE}(X,\mathbb{R})$ of all real-valued non-expansive functions on (X, D_X). We have that $(\mathcal{H}, d_{\mathcal{H}})$ is a metric space, where $d_{\mathcal{H}}$ is the usual Hausdorff distance [29].

Remark 1.2.21. Take $g \in \mathrm{Iso}(X)$. Since $R_g\colon \varphi \mapsto \varphi g$ is continuous, we have that, if $\Omega \in \mathcal{H}$, then $\Omega g := \{\varphi g, \varphi \in \Omega\} \in \mathcal{H}$. Hence, the function $\mathcal{R}_g\colon \mathcal{H} \to \mathcal{H}$, such that $\mathcal{R}_g(\Omega) := \Omega g$ for any $\Omega \in \mathcal{H}$, is well defined. Since $R_g\colon \varphi \mapsto \varphi g$ and $g\colon X \to X$ are bijective, $\mathcal{R}_g$ is an isometry:

$$\begin{aligned}
d_{\mathcal{H}}(\mathcal{R}_g(\Omega), \mathcal{R}_g(\Psi)) &:= \max\left\{\sup_{\varphi\in\mathcal{R}_g(\Omega)} \inf_{\psi\in\mathcal{R}_g(\Psi)} \|\varphi-\psi\|_\infty, \sup_{\psi\in\mathcal{R}_g(\Psi)} \inf_{\varphi\in\mathcal{R}_g(\Omega)} \|\varphi-\psi\|_\infty\right\}\\
&= \max\left\{\sup_{\varphi\in\Omega g} \inf_{\psi\in\Psi g} \|\varphi-\psi\|_\infty, \sup_{\psi\in\Psi g} \inf_{\varphi\in\Omega g} \|\varphi-\psi\|_\infty\right\}\\
&= \max\left\{\sup_{\varphi\in\Omega} \inf_{\psi\in\Psi} \|\varphi g-\psi g\|_\infty, \sup_{\psi\in\Psi} \inf_{\varphi\in\Omega} \|\varphi g-\psi g\|_\infty\right\}\\
&= \max\left\{\sup_{\varphi\in\Omega} \inf_{\psi\in\Psi} \|\varphi-\psi\|_\infty, \sup_{\psi\in\Psi} \inf_{\varphi\in\Omega} \|\varphi-\psi\|_\infty\right\}\\
&= d_{\mathcal{H}}(\Omega, \Psi)
\end{aligned}$$

for any $\Omega, \Psi \in \mathcal{H}$.

Let us consider the action $\rho_{\mathcal{H}}$ of $\mathrm{Iso}(X)$ on $\mathcal{H}$, defined as follows:

$$\rho_{\mathcal{H}}\colon \begin{array}{ccccc} \mathcal{H} & \times & \mathrm{Iso}(X) & \to & \mathcal{H}\\ \Omega & , & g & \mapsto & \Omega g. \end{array}$$

Remark 1.2.22. Consider $\Omega \in \mathcal{H}$ and $g, h \in \mathrm{Iso}(X)$. We have that:

$$\begin{aligned}
d_{\mathcal{H}}(\Omega g, \Omega h) &:= \max\left\{\sup_{\varphi\in\Omega g} \inf_{\psi\in\Omega h} \|\varphi-\psi\|_\infty, \sup_{\psi\in\Omega h} \inf_{\varphi\in\Omega g} \|\varphi-\psi\|_\infty\right\}\\
&= \max\left\{\sup_{\varphi\in\Omega} \inf_{\psi\in\Omega} \|\varphi g-\psi h\|_\infty, \sup_{\psi\in\Omega} \inf_{\varphi\in\Omega} \|\varphi g-\psi h\|_\infty\right\}\\
&\le \max\left\{\sup_{\varphi\in\Omega} \|\varphi g-\varphi h\|_\infty, \sup_{\psi\in\Omega} \|\psi g-\psi h\|_\infty\right\}\\
&= \sup_{\varphi\in\Omega} \|\varphi g-\varphi h\|_\infty\\
&= \sup_{\varphi\in\Omega} \sup_{x\in X} |\varphi g(x)-\varphi h(x)|\\
&\le \sup_{\varphi\in\Omega} \sup_{x\in X} D_X(g(x), h(x))\\
&= d_\infty(g, h).
\end{aligned}$$

Proposition 1.2.23 *$\rho_{\mathcal{H}}$ is a topological action.*

Proof. It will suffice to prove that $\rho_{\mathcal{H}}$ is continuous.We consider the sum pseudo-distance on $\mathcal{H} \times \mathrm{Iso}(X)$, which induces the product topology. By Remark 1.2.22

and the isometry property of $\mathcal{R}_g$ for any $g \in \mathrm{Iso}(X)$, we have that:

$$\begin{aligned} d_{\mathcal{H}}(\rho_{\mathcal{H}}(\Omega, g), \rho_{\mathcal{H}}(\Psi, h)) &= d_{\mathcal{H}}(\Omega g, \Psi h) \\ &\leq d_{\mathcal{H}}(\Omega g, \Omega h) + d_{\mathcal{H}}(\Omega h, \Psi h) \\ &= d_{\mathcal{H}}(\Omega g, \Omega h) + d_{\mathcal{H}}(\Omega, \Psi) \\ &\leq d_{\infty}(g, h) + d_{\mathcal{H}}(\Omega, \Psi) \end{aligned}$$

for any $\Omega, \Psi \in \mathcal{H}$ and any $g, h \in \mathrm{Iso}(X)$. Non-expansiveness of $\rho_{\mathcal{H}}$ ensures us that it is continuous. □

We note that $\mathrm{Aut}_{\Phi}(X)$ is exactly the isotropy group at Φ with respect to the action $\rho_{\mathcal{H}}$.

Theorem 5 *Consider a compact subspace $\Phi \subseteq \mathbb{R}^X_b$. Assume that (X, D_X) is a compact metric space. Then $\mathrm{Aut}_{\Phi}(X)$ is a closed subgroup of $\mathrm{Iso}(X)$. Hence, $\mathrm{Aut}_{\Phi}(X)$ is compact.*

Proof. First, we consider the map $\rho_{\mathcal{H},\Phi} \colon \mathrm{Iso}(X) \to \mathcal{H}$, such that $\rho_{\mathcal{H},\Phi}(g) := \Phi g$ for any $g \in \mathrm{Iso}(X)$. Remark 1.2.22 implies non-expansivity, and hence continuity, of $\rho_{\mathcal{H},\Phi}$:

$$d_{\mathcal{H}}(\rho_{\mathcal{H},\Phi}(g), \rho_{\mathcal{H},\Phi}(h)) = d_{\mathcal{H}}(\Phi g, \Phi h) \leq d_{\infty}(g, h)$$

for any $g, h \in \mathrm{Iso}(X)$. It would be easy to check that $\rho^{-1}_{\mathcal{H},\Phi}(\Phi) = \mathrm{Aut}_{\Phi}(X)$. Since $\mathrm{Aut}_{\Phi}(X)$ is the preimage of a closed set by a continuous function, $\mathrm{Aut}_{\Phi}(X)$ is closed. The compactness of $\mathrm{Iso}(X)$ implies that $\mathrm{Aut}_{\Phi}(X)$ is compact. □

After Corollary 1.2.7 and Corollary 1.2.16, it seems that compactness of X and compactness of a group G of invertible Φ-operations are not correlated. However, Theorem 5 tells us that completeness, and hence compactness, of X implies at least the compactness of the largest group that preserves Φ.

1.3 The natural pseudo-distance.

In addition to the pseudo-metric D_{Φ}, we define another pseudo-distance d_G on the space Φ (see also [32]). It represents the ground truth in our model. Indeed, it allows for comparison between functions and it vanishes for pairs of functions that are equivalent with respect to the action of our group of isometries $G \subseteq \mathrm{Aut}_{\Phi}(X)$, which expresses the data equivalences relevant for the observer.

Definition 1.3.1. The pseudo-distance d_G on Φ is defined by setting

$$d_G(\varphi_1, \varphi_2) := \inf_{g \in G} D_{\Phi}(\varphi_1, \varphi_2 g).$$

It is called the **natural pseudo-distance** associated with the group G acting on Φ.

In other terms, the natural pseudo-distance d_G between two signals φ_1 and φ_2 can be seen as the distance between the orbits $\varphi_1 G$ and $\varphi_2 G$ with respect to the action of G on Φ (cf. [9]). If $G = \{\mathrm{Id} : x \mapsto x\}$, then d_G equals the sup-norm distance D_Φ on Φ. If G_1 and G_2 are subgroups of $\mathrm{Aut}_\Phi(X)$ and $G_1 \subseteq G_2$, then the definition of d_G implies that

$$d_{\mathrm{Aut}_\Phi(X)}(\varphi_1, \varphi_2) \le d_{G_2}(\varphi_1, \varphi_2) \le d_{G_1}(\varphi_1, \varphi_2) \le D_\Phi(\varphi_1, \varphi_2) \tag{1.26}$$

for every $\varphi_1,\ \varphi_2 \in \mathit{\Phi}$.

Though d_G represents the ground truth for data similarity in our model, unfortunately it is difficult to compute. This is also a consequence of the fact that we can easily find subgroups G of $\mathrm{Aut}_\Phi(X)$ that cannot be approximated with arbitrary precision by smaller finite subgroups of G (e.g., when G is the group of rigid motions of $X = \mathbb{R}^3$; see, e.g., Section 3.1 in [32]).

In the following sections, we will also show how d_G can be approximated with arbitrary precision by means of a dual approach based on group equivariant non-expansive operators (GENEOs) and persistent homology.

Chapter 2

The space of GENEOs

This chapter is devoted to introducing the concept of Group Equivariant Non-Expansive Operator (GENEO) and illustrate some results about its topological structure. GENEOs enable us to transform data sets (or perception pairs), preserving symmetries and distances. In other words, we can say that in our framework agents are modelled as sets of GENEOs. Moreover, we define a new strongly-invariant pseudo-distance on Φ that involves both persistent homology (cf. Appendix A) and GENEOs. This pseudo-metric will be of use to approximate the natural pseudo-distance.

Definition 2.0.1. Consider two perception pairs (Φ, G), (Ψ, H). Each map (F, T) : $(\Phi, G) \to (\Psi, H)$ such that F is a continuous map, T is an homomorphism and F is T-equivariant (i.e., $F(\varphi g) = F(\varphi)T(g)$ for every $\varphi \in \Phi,\ g \in G$) is said to be a **Group Equivariant Operator (GEO)** from (Φ, G) to (Ψ, H).

We observe that the functions in Φ and the functions in Ψ are defined on spaces that are generally different from each other, and the groups of invariance can be different as well. This is important, as it allows one to compose operators hierarchically.

Definition 2.0.2. Assume that (Φ, G), (Ψ, H) are two perception pairs. If (F, T) is a GEO from (Φ, G) to (Ψ, H) and F is non-expansive, then F is called a **Group Equivariant Non-Expansive Operator (GENEO)**.

In many situations it could be important to fix the homomorphism T. In this case we call GENEO (or GEO) with respect to T a map $F : \Phi \to \Psi$ such that (F, T) is a GENEO (or GEO).

2.1 Compactness and convexity of the space of GENEOs

Let us consider two perception pairs (Φ, G) and (Ψ, H). Under suitable assumptions on the data sets, we can prove two important properties of the space of GENEOs from (Φ, G) to (Ψ, H) with respect to a fixed homomorphism $T \colon G \to H$: compactness and convexity. The compactness guarantees that the space of GENEOs can be approximated by a finite set. The convexity implies that new GENEOs can be obtained by convex combinations of pre-existing GENEOs. If $\mathcal{F}_T^{\mathrm{all}}$ denotes the set of GENEOs between two perception pairs (Φ, G), (Ψ, H) associated with the homomorphism $T : G \to H$, then the following theorem holds:

Theorem 6 *If Φ and Ψ are compact with respect to D_Φ and D_Ψ, respectively, then $\mathcal{F}_T^{\text{all}}$ is compact with respect to the uniform convergence distance set by*

$$D_{\text{GENEO}}(F_1, F_2) := \sup_{\varphi \in \Phi} D_\Psi(F_1(\varphi), F_2(\varphi))$$

for any $F_1, F_2 \in \mathcal{F}_T^{\text{all}}$.

Proof. It is easy to check that the space $\mathbf{NE}(\Phi, \Psi)$ of non-expansive maps from Φ to Ψ is uniformly equicontinuous. The compactness of Φ and Ψ implies the compactness of the space $\mathbf{NE}(\Phi, \Psi)$ by Arzelà-Ascoli Theorem (see [39]). To prove our statement it will suffice to show that $\mathcal{F}_T^{\text{all}}$ is closed as a subset of $\mathbf{NE}(\Phi, \Psi)$. Let us consider a sequence $(F_i)_{i\in\mathbb{N}}$ of GENEOs convergent to a non-expansive operator F. Since the action of the groups G and H are continuous and F_i is equivariant for every i, we have that for any $g \in G$ and any $\varphi \in \Phi$:

$$F(\varphi g) = \lim_{i\to\infty} F_i(\varphi g) = \lim_{i\to\infty} F_i(\varphi)T(g) = F(\varphi)T(g).$$

This proves that F is a GENEO, and $\mathcal{F}_T^{\text{all}}$ is closed. □

Now let $F_1, F_2, \ldots, F_n \in \mathcal{F}_T^{\text{all}}$. Consider an n-tuple $\Sigma = (a_1, a_2, \ldots, a_n)$ of real numbers with $\sum_{i=1}^n |a_i| \leq 1$ and assume Ψ is convex. We can define an operator $F_\Sigma \colon \Phi \to \Psi$:

$$F_\Sigma(\varphi) := \sum_{i=1}^n a_i F_i(\varphi) \tag{2.1}$$

for any $\varphi \in \Phi$. Note that the convexity of Ψ ensures us that F_Σ is well defined.

Lemma 2.1.1 *Under the above assumptions, F_Σ belongs to $\mathcal{F}_T^{\text{all}}$.*

Proof. First we prove that F_Σ is a GEO with respect to T. Since every F_i is T-equivariant, we have that:

$$F_\Sigma(\varphi g) = \sum_{i=1}^n a_i F_i(\varphi g) = \sum_{i=1}^n a_i (F_i(\varphi)T(g)) = \left(\sum_{i=1}^n a_i F_i(\varphi)\right) T(g) = F_\Sigma(\varphi)T(g).$$

Since every F_i is non-expansive, F_Σ is non-expansive:

$$\begin{aligned} D_\Psi(F_\Sigma(\varphi_1), F_\Sigma(\varphi_2)) &= \left\| \sum_{i=1}^n a_i F_i(\varphi_1) - \sum_{i=1}^n a_i F_i(\varphi_2) \right\|_\infty \\ &= \left\| \sum_{i=1}^n a_i (F_i(\varphi_1) - F_i(\varphi_2)) \right\|_\infty \\ &\leq \sum_{i=1}^n |a_i| \, \|(F_i(\varphi_1) - F_i(\varphi_2))\|_\infty \\ &\leq \sum_{i=1}^n |a_i| \, \|\varphi_1 - \varphi_2\|_\infty \leq D_\Phi(\varphi_1, \varphi_2). \end{aligned}$$

Therefore F_Σ is a GENEO. □

Therefore, the following theorem holds:

Theorem 7 *If Ψ is convex, then the set $\mathcal{F}_T^{\text{all}}$ is convex.*

Proof. It is sufficient to apply Lemma 2.1.1 for $n = 2$, by setting $a_1 = t$, $a_2 = 1 - t$ for $0 \leq t \leq 1$. □

2.2 A strongly invariant pseudo-metric induced by Persistent Homology

Let us consider two perception pairs (Φ, G) and (Ψ, H). To compare data under the action of a set $\mathcal{F}$ of GENEOs from (Φ, G) to (Ψ, H) with respect to T, one could simply define, for any $\varphi_1, \varphi_2 \in \Phi$,

$$D_{\mathcal{F},\Phi}(\varphi_1, \varphi_2) := \sup_{F \in \mathcal{F}} \|F(\varphi_1) - F(\varphi_2)\|_\infty.$$

Though, the computation of $D_{\mathcal{F},\Phi}(\varphi_1, \varphi_2)$ for every pair (φ_1, φ_2) of admissible functions is computationally expensive. Persistent homology allows us to replace $D_{\mathcal{F},\Phi}$ with a pseudo-metric $\mathcal{D}_{\text{match}}^{\mathcal{F},k}$ computationally more efficient, but still stable and strongly invariant. Where, a pseudo-metric $\hat{d}$ on Φ is **strongly G-invariant** if it is invariant under the action of G with respect to each variable, that is, if

$$\hat{d}(\varphi_1, \varphi_2) = \hat{d}(\varphi_1 g, \varphi_2) = \hat{d}(\varphi_1, \varphi_2 g) = \hat{d}(\varphi_1 g, \varphi_2 g)$$

for every $\varphi_1, \varphi_2 \in \varphi$ and every $g \in G$.

Remark 2.2.1. It is easily seen that the natural pseudo-distance d_G is strongly G-invariant.

From now on we will freely use persistent homology theory. The interested reader can find more information about persistence in Appendix A.

Remark 2.2.2. In our setting, Persistent Betti numbers functions (PBNs) are not necessarily finite. For example, let us consider the set $X = \{0\} \cup \{\frac{1}{n}$, with $n \in \mathbb{N}^+\}$ and $\Phi = \{i \colon X \hookrightarrow \mathbb{R}\}$, where i is the natural inclusion. Even if X is compact, every sublevel set $X_u = \{x \in X : x \leq u\}$ with $u > 0$ has infinitely many connected components, and hence the 0th persistent Betti numbers function takes infinite value at every point (u, v) with $0 < u < v$.

We add the assumption that the persistent Betti numbers function of every element φ of the considered perception pairs takes a finite value at each point $(u, v) \in \Delta^+$ to get stability and preclude pathological cases (for example the case that the set φ of admissible functions is the set of all maps from X to $\mathbb{R}$).

Remark 2.2.3. Since the PBNs of the pseudo-metric space (X, D_X) coincide with the persistent Betti numbers functions of its Kolmogorov quotient $\bar{X}$, the finiteness of the persistent Betti numbers functions can be obtained when $\bar{X}$ is finitely triangulable (cf. [15]).

Before proceeding, we recall the stability of the classical pseudo-distance d_{match} between persistent Betti numbers functions (cf. definitions in Appendix A) with respect to the pseudo-metrics D_Φ and $d_{\mathrm{Aut}_\Phi(X)}$. In other terms, the distance d_{match} can be used as an efficient proxy for the max-norm distance between real-valued functions.

Theorem 8 *If k is a natural number, $G \subseteq \mathrm{Aut}_\Phi(X)$ and $\varphi_1, \varphi_2 \in \Phi$, then*

$$d_{\mathrm{match}}(r_k(\varphi_1), r_k(\varphi_2)) \le d_{\mathrm{Aut}_\Phi(X)}(\varphi_1, \varphi_2) \le d_G(\varphi_1, \varphi_2) \le D_\Phi(\varphi_1, \varphi_2),$$

where $r_k(\varphi_i)$ is the persistent Betti numbers function of φ_i for $i = 1, 2$.

The proof of the first inequality $d_{\mathrm{match}}(r_k(\varphi_1), r_k(\varphi_2)) \le d_{\mathrm{Aut}_\Phi(X)}(\varphi_1, \varphi_2)$ in Theorem 8 is based on the stability of d_{match} with respect to D_Φ and can be found in [15]. The other inequalities follow from the definition of the natural pseudo-distance. that the distance d_{match} can be used as an efficient proxy for the max-norm distance between real-valued functions.

Let us consider a non-empty subset $\mathcal{F}$ of $\mathcal{F}_T^{\mathrm{all}}$. For every fixed k, we can consider the following extended pseudo-metric $\mathcal{D}_{\mathrm{match}}^{\mathcal{F},k}$ on Φ:

$$\mathcal{D}_{\mathrm{match}}^{\mathcal{F},k}(\varphi_1, \varphi_2) := \sup_{F \in \mathcal{F}} d_{\mathrm{match}}(r_k(F(\varphi_1)), r_k(F(\varphi_2))) \tag{2.2}$$

for every $\varphi_1, \varphi_2 \in \Phi$, where $r_k(\varphi)$ denotes the kth persistent Betti numbers function with respect to the function $\varphi : X \to \mathbb{R}$.

Proposition 2.2.4 *$\mathcal{D}_{\mathrm{match}}^{\mathcal{F},k}$ is a strongly G-invariant pseudo-metric on Φ.*

Proof. Theorem 8 and the non-expansivity of every $F \in \mathcal{F}$ imply that

$$\begin{aligned} d_{\mathrm{match}}(r_k(F(\varphi_1)), r_k(F(\varphi_2))) &\le D_\Psi\left(F(\varphi_1), F(\varphi_2)\right) \\ &\le D_\Phi\left(\varphi_1, \varphi_2\right). \end{aligned}$$

Therefore $\mathcal{D}_{\mathrm{match}}^{\mathcal{F},k}$ is a pseudo-metric, since it is the supremum of a family of pseudo-metrics that are bounded at each pair (φ_1, φ_2). Moreover, for every $\varphi_1, \varphi_2 \in \Phi$ and every $g \in G$

$$\begin{aligned} \mathcal{D}_{\mathrm{match}}^{\mathcal{F},k}(\varphi_1, \varphi_2 g) &:= \sup_{F \in \mathcal{F}} d_{\mathrm{match}}(r_k(F(\varphi_1)), r_k(F(\varphi_2 g))) \\ &= \sup_{F \in \mathcal{F}} d_{\mathrm{match}}(r_k(F(\varphi_1)), r_k(F(\varphi_2)T(g))) \\ &= \sup_{F \in \mathcal{F}} d_{\mathrm{match}}(r_k(F(\varphi_1)), r_k(F(\varphi_2)) \\ &= \mathcal{D}_{\mathrm{match}}^{\mathcal{F},k}(\varphi_1, \varphi_2) \end{aligned}$$

because of the equality $F(\varphi g) = F(\varphi)T(g)$ for every $\varphi \in \Phi$ and every $g \in G$ and the invariance of persistent homology under the action of the homeomorphisms. Since the function $\mathcal{D}_{\mathrm{match}}^{\mathcal{F},k}$ is symmetric, this is sufficient to guarantee that $\mathcal{D}_{\mathrm{match}}^{\mathcal{F},k}$ is strongly G-invariant. □

2.2.1 Some theoretical results on the pseudo-metric $\mathcal{D}_{\text{match}}^{\mathcal{F},k}$

On one hand, we prove that the pseudo-metric $\mathcal{D}_{\text{match}}^{\mathcal{F},k}$ is stable with respect to both the natural pseudo-distance d_G associated with the group G and the distance D_Φ. On the other hand, we will also show that by means of $D_{\mathcal{F},\Phi}$ we could approximate the natural pseudo-distance d_G, under suitable assumptions.

Theorem 9 *Let $\mathcal{F}^{\text{all}}$ the space of all GENEOs from (Φ, G) to (Ψ, H). If $\mathcal{F}$ is a non-empty subset of $\mathcal{F}^{\text{all}}$, then*

$$\mathcal{D}_{\text{match}}^{\mathcal{F},k} \leq d_G \leq D_\Phi. \tag{2.3}$$

Proof. For every $(F, T) \in \mathcal{F}$, every $g \in G$ and every $\varphi_1, \varphi_2 \in \Phi$, we have that

$$\begin{aligned} d_{\text{match}}(r_k(F(\varphi_1)), r_k(F(\varphi_2))) &= d_{\text{match}}(r_k(F(\varphi_1)), r_k(F(\varphi_2)T(g))) \\ &= d_{\text{match}}(r_k(F(\varphi_1)), r_k(F(\varphi_2 g))) \\ &\leq D_\Psi\left(F(\varphi_1), F(\varphi_2 g)\right) \leq D_\Phi\left(\varphi_1, \varphi_2 g\right). \end{aligned}$$

The first equality follows from the invariance of persistent homology under action of the group $\text{Homeo}(X)$ of all homeomorphisms from X to itself (see Remark A.0.2), and the second equality follows from the fact F is T-equivariant. The first inequality follows from the stability of persistent homology (Theorem 8), while the second inequality follows from the non-expansivity of F. It follows that, if $\mathcal{F} \subseteq \mathcal{F}^{\text{all}}$, then for every $g \in G$ and every $\varphi_1, \varphi_2 \in \Phi$

$$\mathcal{D}_{\text{match}}^{\mathcal{F},k}(\varphi_1, \varphi_2) \leq D_\Phi\left(\varphi_1, \varphi_2 g\right). \tag{2.4}$$

Hence, the inequality $\mathcal{D}_{\text{match}}^{\mathcal{F},k} \leq d_G$ follows, while $d_G \leq D_\Phi$ is stated in Theorem 8. □

The definitions of the natural pseudo-distance d_G and the pseudo-distance $\mathcal{D}_{\text{match}}^{\mathcal{F},k}$ come from different theoretical concepts. The former relies on finding a homeomorphism in G that minimizes the L^∞-distance between data, while the latter refers only to a comparison of persistent homologies depending on a family of GENEOs. Given those comments, the next result may appear unexpected.

Theorem 10 *Let $\mathcal{F}^{\text{all}}$ the space of all GENEOs from (Φ, G) to itself. Let us assume that every function in φ is non-negative, the kth Betti number of X does not vanish, and Φ contains each constant function c for which a function $\varphi \in \Phi$ exists such that $0 \leq c \leq \|\varphi\|_\infty$. Then $\mathcal{D}_{\text{match}}^{\mathcal{F}^{\text{all}},k} = d_G$.*

Proof. For every $\varphi' \in \Phi$ let us consider the operator $(F_{\varphi'}, \text{id}_G) \colon (\Phi, G) \to (\Phi, G)$ defined by setting $F_{\varphi'}(\varphi)$ equal to the constant function taking everywhere the value $d_G(\varphi, \varphi')$ for every $\varphi \in \Phi$ (i.e., $F_{\varphi'}(\varphi)(x) = d_G(\varphi, \varphi')$ for any $x \in X$). Our assumptions guarantee that such a constant function belongs to Φ.

We observe that

1. $F_{\varphi'}$ is a GEO on Φ, because the strong invariance of the natural pseudo-distance d_G with respect to the group G (Remark 2.2.1) implies that if $\varphi \in \varphi$ and $g \in G$, then $F_{\varphi'}(\varphi g)(x) = d_G(\varphi g, \varphi') = F_{\varphi'}(\varphi)(g(x)) = (F_{\varphi'}(\varphi)g)(x) = (F_{\varphi'}(\varphi)\text{id}_G(g))(x)$, for every $x \in X$.

2. $F_{\varphi'}$ is non-expansive on φ, because for every $\varphi_1, \varphi_2 \in \varphi$

$$D_\Psi\left(F_{\varphi'}(\varphi_1), F_{\varphi'}(\varphi_2)\right) = |d_G(\varphi_1, \varphi') - d_G(\varphi_2, \varphi')| \\ \leq d_G(\varphi_1, \varphi_2) \leq D_\Phi(\varphi_1, \varphi_2).$$

Therefore, $(F_{\varphi'}, \mathrm{id}_G)$ is a GENEO.

For every $\varphi_1, \varphi_2, \varphi' \in \Phi$ we have that

$$d_{\text{match}}(r_k(F_{\varphi'}(\varphi_1)), r_k(F_{\varphi'}(\varphi_2))) = |d_G(\varphi_1, \varphi') - d_G(\varphi_2, \varphi')|. \tag{2.5}$$

Indeed, apart from the trivial points on the line $\{(u, v) \in \mathbb{R}^2 \ : \ u = v\}$, the persistence diagram associated with $r_k(F_{\varphi'}(\varphi_1))$ contains only the point $(d_G(\varphi_1, \varphi'), \infty)$, while the persistence diagram associated with $r_k(F_{\varphi'}(\varphi_2))$ contains only the point $(d_G(\varphi_2, \varphi'), \infty)$. Both the points have the same multiplicity, which equals the (non-null) k-th Betti number of X.

Setting $\varphi' = \varphi_2$, we have that

$$d_{\text{match}}(r_k(F_{\varphi'}(\varphi_1)), r_k(F_{\varphi'}(\varphi_2))) = d_G(\varphi_1, \varphi_2). \tag{2.6}$$

As a consequence, we have that

$$\mathcal{D}_{\text{match}}^{\mathcal{F}^{\text{all}},k}(\varphi_1, \varphi_2) \geq d_G(\varphi_1, \varphi_2). \tag{2.7}$$

By applying Theorem 9, we get

$$\mathcal{D}_{\text{match}}^{\mathcal{F}^{\text{all}},k}(\varphi_1, \varphi_2) = d_G(\varphi_1, \varphi_2) \tag{2.8}$$

for every φ_1, φ_2. □

Remark 2.2.5. We observe that if Φ is bounded, the assumption that every function in Φ is non-negative is not quite restrictive. Indeed, we can obtain it by adding a suitable constant value to every admissible function.

The next result will be of use for the approximation of $\mathcal{D}_{\text{match}}^{\mathcal{F},k}$. Before proceeding we have to define a new pseudo-distance of the space of GENEOs, based on the natural pseudo-distance d_H, as follows: we set

$$D_{\text{GENEO},H}(F_1, F_2) := \sup_{\varphi \in \Phi} d_H(F_1(\varphi), F_2(\varphi))$$

for any F_1 and F_2 be two GENEOs from (Φ, G) to (Ψ, H). We note that $D_{\text{GENEO},H} \leq D_{\text{GENEO}}$.

Proposition 2.2.6 *Let $\mathcal{F}^{\text{all}}$ the space of all GENEOs from (Φ, G) to (Ψ, H). Assume that $\mathcal{F}, \mathcal{F}' \subseteq \mathcal{F}^{\text{all}}$. If the Hausdorff distance*

$$HD(\mathcal{F}, \mathcal{F}') := \max\left\{\sup_{F \in \mathcal{F}} \inf_{F' \in \mathcal{F}'} D_{\text{GENEO,H}}(F, F'), \sup_{F' \in \mathcal{F}'} \inf_{F \in \mathcal{F}} D_{\text{GENEO,H}}(F, F')\right\}$$

is not larger than ε, then

$$\left|\mathcal{D}_{\text{match}}^{\mathcal{F},k}(\varphi_1, \varphi_2) - \mathcal{D}_{\text{match}}^{\mathcal{F}',k}(\varphi_1, \varphi_2)\right| \leq 2\varepsilon \tag{2.9}$$

for every $\varphi_1, \varphi_2 \in \Phi$.

Proof. Since $HD(\mathcal{F}, \mathcal{F}') \leq \varepsilon$, for every $F \in \mathcal{F}$ a $F' \in \mathcal{F}'$ and an $\eta > 0$ exist such that $D_{\mathrm{GENEO,H}}(F, F') \leq \varepsilon + \eta$. The definition of $D_{\mathrm{GENEO,H}}$ implies that $d_H(F(\varphi), F'(\varphi)) \leq \varepsilon + \eta$ for every $\varphi \in \varphi$. From Theorem 8 it follows that

$$d_{\mathrm{match}}(r_k(F(\varphi_1)), r_k(F'(\varphi_1)) \leq \varepsilon + \eta \tag{2.10}$$

and

$$d_{\mathrm{match}}(r_k(F(\varphi_2)), r_k(F'(\varphi_2)) \leq \varepsilon + \eta \tag{2.11}$$

for every $\varphi_1, \varphi_2 \in \Phi$.

Therefore,

$$\left| d_{\mathrm{match}}(r_k(F(\varphi_1)), r_k(F(\varphi_2)) - d_{\mathrm{match}}(r_k(F'(\varphi_1)), r_k(F'(\varphi_2)) \right| \leq 2(\varepsilon + \eta). \tag{2.12}$$

As a consequence, $\mathcal{D}^{\mathcal{F},k}_{\mathrm{match}}(\varphi_1, \varphi_2) \leq \mathcal{D}^{\mathcal{F}',k}_{\mathrm{match}}(\varphi_1, \varphi_2) + 2(\varepsilon + \eta)$. We can show analogously that $\mathcal{D}^{\mathcal{F}',k}_{\mathrm{match}}(\varphi_1, \varphi_2) \leq \mathcal{D}^{\mathcal{F},k}_{\mathrm{match}}(\varphi_1, \varphi_2) + 2(\varepsilon + \eta)$. Since η can be chosen arbitrarily small, from the previous two inequalities the proof of our statement follows. □

Since the compactness of the space $\mathcal{F}^{\mathrm{all}}_T \subseteq \mathcal{F}^{\mathrm{all}}$ guarantees we can cover $\mathcal{F}$ by a finite set of balls in $\mathcal{F}^{\mathrm{all}}_T$ of radius ε, centered at points of a finite set $\mathcal{F}' \subseteq \mathcal{F}$, the following proposition states that the approximation of $\mathcal{D}^{\mathcal{F},k}_{\mathrm{match}}(\varphi_1, \varphi_2)$ can be reduced to the computation of $\mathcal{D}^{\mathcal{F}',k}_{\mathrm{match}}(\varphi_1, \varphi_2)$, i.e. the maximum of a finite set of bottleneck distances between persistence diagrams, which are well-known to be computable by means of efficient algorithms.

Proposition 2.2.7 *Let $\mathcal{F}$ be a non-empty subset of $\mathcal{F}^{\mathrm{all}}_T$. For every $\varepsilon > 0$, a finite subset $\mathcal{F}^*$ of $\mathcal{F}$ exists, such that*

$$|\mathcal{D}^{\mathcal{F}^*,k}_{\mathrm{match}}(\varphi_1, \varphi_2) - \mathcal{D}^{\mathcal{F},k}_{\mathrm{match}}(\varphi_1, \varphi_2)| \leq \varepsilon \tag{2.13}$$

for every $\varphi_1, \varphi_2 \in \Phi$.

Proof. Let us consider the closure $\bar{\mathcal{F}}$ of $\mathcal{F}$ in $\mathcal{F}^{\mathrm{all}}_T$. Let us also consider the covering $\mathcal{U}$ of $\bar{\mathcal{F}}$ obtained by taking all the open balls of radius $\frac{\varepsilon}{2}$ centered at points of $\mathcal{F}$, with respect to D_{GENEO}. Theorem 6 guarantees that $\mathcal{F}^{\mathrm{all}}_T$ is compact, hence also $\bar{\mathcal{F}}$ is compact. Therefore we can extract a finite covering $\{B_1, \ldots, B_m\}$ of $\bar{\mathcal{F}}$ from $\mathcal{U}$. We can set $\mathcal{F}^*$ equal to the set of centers of the balls $B_1, \ldots, B_m$. The statement of the proposition immediately follows from Proposition 2.2.6, by recalling that $D_{\mathrm{GENEO,H}} \leq D_{\mathrm{GENEO}}$ and hence $HD\left(\bar{\mathcal{F}}, \mathcal{F}^*\right) \leq \varepsilon/2$. □

2.3 Other pseudo-metrics induced by persistent homology

Persistent homology can be seen as a topological method to build new and easily computable pseudo-metrics for the sets Φ, $\mathrm{Aut}_\Phi(X)$ and $\mathcal{F}^{\mathrm{all}}_T$. These new pseudo-metrics $\mathfrak{D}_\Phi$, $\mathfrak{D}_{\mathrm{Aut}}$, $\mathfrak{D}_{\mathrm{GENEO}}$ can be used as proxies for $d_{\mathrm{Aut}_\Phi(X)}$ (and hence D_Φ), D_{Aut}, D_{GENEO}, respectively:

- If $\varphi_1, \varphi_2 \in \Phi$, we can set $\mathfrak{D}_\Phi(\varphi_1, \varphi_2) := d_{\text{match}}(r_k(\varphi_1), r_k(\varphi_2))$. The stability theorem for persistence diagrams (Theorem 8) can be reformulated as the inequalities $\mathfrak{D}_\Phi \leq d_{\text{Aut}_\Phi(X)} \leq D_\Phi$.

- If $g_1, g_2 \in \text{Aut}_\Phi(X)$, we can set

$$\mathfrak{D}_{\text{Aut}}(g_1, g_2) := \sup_{\varphi \in \Phi} d_{\text{match}}(r_k(\varphi g_1), r_k(\varphi g_2)).$$

From Theorem 8 the inequality $\mathfrak{D}_{\text{Aut}} \leq D_{\text{Aut}}$ follows.

- If $F_1, F_2 \in \mathcal{F}_T^{\text{all}}$, we can set

$$\mathfrak{D}_{\text{GENEO}}(F_1, F_2) := \sup_{\varphi \in \Phi} d_{\text{match}}(r_k(F_1(\varphi)), r_k(F_2(\varphi))).$$

From Theorem 8 the inequality $\mathfrak{D}_{\text{GENEO}} \leq D_{\text{GENEO}}$ follows.

Remark 2.3.1. Theorem 6 and the inequality $\mathfrak{D}_{\text{GENEO}} \leq D_{\text{GENEO}}$ immediately imply that $\mathcal{F}_T^{\text{all}}$ is compact also with respect to the topologies induced by $\mathfrak{D}_{\text{GENEO}}$.

We underline that the use of persistent homology is a key tool in our approach: it allows for a fast comparison between functions and between GENEOs. Without persistent homology, this comparison would be much more computationally expensive.

Chapter 3

Methods for building GENEOs

In Theorem 6 it has been proved that $\mathcal{F}_T^{\text{all}}$ is compact, if we assume that Φ is compact. This guarantees that, in principle, $\mathcal{F}_T^{\text{all}}$ can be approximated by a finite subset. In order to proceed along this line of research we need general methods for building GENEOs. According to the goal of realizing those methods, this chapter is devoted to proving some new results about the construction of GENEOs.

3.1 Building GENEOs by means of a finite set of known GENEOs

A first simple method to build new GENEOs is simply the composition of two GENEOs.

Proposition 3.1.1 *If $(F_1, T_1)\colon (\Phi, G) \to (\Psi, H)$ and $(F_2, T_2)\colon (\Psi, H) \to (\Omega, M)$ are GENEOs, then $(F, T) := (F_2 \circ F_1, T_2 \circ T_1)\colon (\Phi, G) \to (\Omega, M)$ is a GENEO.*

Proof. It will suffice to verify that $(F, T) := (F_2 \circ F_1, T_2 \circ T_1)$ is a GENEO.

1. Since F_1 is T_1-equivariant and F_2 is T_2-equivariant, F is $(T_2 \circ T_1)$-equivariant:

$$\begin{aligned} F(\varphi g) &= (F_2 \circ F_1)(\varphi g) = F_2(F_1(\varphi g)) \\ &= F_2(F_1(\varphi)T_1(g)) = F_2(F_1(\varphi))T_2(T_1(g)) \\ &= F(\varphi)T(g) \end{aligned} \tag{3.1}$$

   for every $\varphi \in \Phi$ and every $g \in G$.

2. Since F_1, F_2 are non-expansive, F is non-expansive:

$$\begin{aligned} \|F(\varphi_1) - F(\varphi_2)\|_\infty &= \|(F_2 \circ F_1)(\varphi_1) - (F_2 \circ F_1)(\varphi_2)\|_\infty \\ &= \|F_2(F_1(\varphi_1)) - F_2(F_1(\varphi_2))\|_\infty \\ &\leq \|F_1(\varphi_1) - F_1(\varphi_2)\|_\infty \\ &\leq \|\varphi_1 - \varphi_2\|_\infty \end{aligned} \tag{3.2}$$

   for any $\varphi_1, \varphi_2 \in \Phi$.

□

In general, we are often interested in building GENEOs with respect to a fixed group homomorphism. In the sequel, we present a method that combines a given finite set of GENEOs in order to produce a new operator.

Let $F_1, \ldots, F_n$ be GENEOs from (Φ, G) to (Ψ, H) with respect to $T\colon G \to H$, where $\mathrm{dom}(\Psi) = Y$. Assume that $\mathcal{L}\colon \mathbb{R}^n \to \mathbb{R}$ is a non-expansive map, where $\mathbb{R}^n$ is endowed with the usual norm $\|(x_1, \ldots, x_n)\|_\infty = \max_{1\le i\le n} |x_i|$. Now we consider the function

$$\mathcal{L}^*(F_1, \ldots, F_n)(\varphi) := [\mathcal{L}(F_1(\varphi), \ldots, F_n(\varphi))]$$

from Φ to the space $C^0(Y, \mathbb{R})$ of real-valued continuous functions on Y, where $[\mathcal{L}(F_1(\varphi), \ldots, F_n(\varphi))]$ is defined by setting

$$\mathcal{L}(F_1(\varphi), \ldots, F_n(\varphi))](x) := \mathcal{L}(F_1(\varphi)(x), \ldots, F_n(\varphi)(x))$$

for any $x \in X$.

Proposition 3.1.2 *Assume that $F_1, \ldots, F_n$ are GENEOs from (Φ, G) to (Ψ, H) with respect to $T\colon G \to H$ and $\mathcal{L}$ is a non-expansive map from $\mathbb{R}^n$ to $\mathbb{R}$. If $\mathcal{L}^*(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$, then $\mathcal{L}^*(F_1, \ldots, F_n)$ is a GENEO from (Φ, G) to (Ψ, H) with respect to T.*

Proof.

1. The T-equivariance of $F_1, \ldots, F_n$ implies that $\mathcal{L}^*(F_1, \ldots, F_n)$ is T-equivariant:

$$\begin{aligned}\mathcal{L}^*(F_1, \ldots, F_n)(\varphi g) &= [\mathcal{L}(F_1(\varphi g), \ldots, F_n(\varphi g))]\\ &= [\mathcal{L}(F_1(\varphi)T(g), \ldots, F_n(\varphi)T(g))]\\ &= [\mathcal{L}(F_1(\varphi), \ldots, F_n(\varphi))]T(g)\\ &= \mathcal{L}^*(F_1, \ldots, F_n)(\varphi)T(g) \qquad (3.3)\end{aligned}$$

   for every $\varphi \in \Phi$ and every $g \in G$.

2. Since $F_1, \ldots, F_n$ are non-expansive and $\mathcal{L}$ is non-expansive, for every $x \in X$ and every $\varphi_1, \varphi_2 \in \Phi$ we have that

$$\begin{aligned}&|\mathcal{L}(F_1(\varphi_1)(x), \ldots, F_n(\varphi_1)(x)) - \mathcal{L}(F_1(\varphi_2)(x), \ldots, F_n(\varphi_2)(x))|\\ &\le \|(F_1(\varphi_1(x)) - F_1(\varphi_2(x)), \ldots, F_n(\varphi_1(x)) - F_n(\varphi_2(x)))\|_\infty\\ &= \max_{1\le i\le n} |F_i(\varphi_1(x)) - F_i(\varphi_2(x))|\\ &\le \max_{1\le i\le n} \|F_i(\varphi_1) - F_i(\varphi_2)\|_\infty\\ &\le \|\varphi_1 - \varphi_2\|_\infty. \qquad (3.4)\end{aligned}$$

   In conclusion,

$$\|\mathcal{L}^*(F_1, \ldots, F_n)(\varphi_1) - \mathcal{L}^*(F_1, \ldots, F_n)(\varphi_2)\|_\infty \le \|\varphi_1 - \varphi_2\|_\infty.$$

   Therefore $\mathcal{L}^*(F_1, \ldots, F_n)$ is non-expansive.

□

The above proposition exposes a quite general method to build news GENEOs from a finite set of known GENEOs. In the following, we show some examples.

Maximum operator Let $F_1, \ldots, F_n$ be GENEOs from (Φ, G) to (Ψ, H) with respect to $T\colon G \to H$, where $\mathrm{dom}(\Psi) = Y$. Consider the function

$$\max(F_1, \ldots, F_n)(\varphi) := [\max(F_1(\varphi), \ldots, F_n(\varphi))]$$

from Φ to $C^0(Y, \mathbb{R})$, where $[\max(F_1(\varphi), \ldots, F_n(\varphi))]$ is defined by setting

$$[\max(F_1(\varphi), \ldots, F_n(\varphi))](x) := \max\{F_1(\varphi)(x), \ldots, F_n(\varphi)(x)\}.$$

Proposition 3.1.3 *If* $\max(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$*, then* $\max(F_1, \ldots, F_n)$ *is a GENEO from* (Φ, G) *to* (Ψ, H) *with respect to* T*.*

In order to proceed, we recall the proof of the following lemma:

Lemma 3.1.4 *For every* $u_1, \ldots, u_n, v_1, \ldots, v_n \in \mathbb{R}$ *it holds that*

$$|\max\{u_1, \ldots, u_n\} - \max\{v_1, \ldots, v_n\}| \leq \max\{|u_1 - v_1|, \ldots, |u_n - v_n|\}.$$

Proof. Without loss of generality we can suppose that $\max\{u_1, \ldots, u_n\} = u_1$. If $\max\{v_1, \ldots, v_n\} = v_1$ the claim trivially follows. It only remains to check the case $\max\{v_1, \ldots, v_n\} = v_i,\ i \neq 1$. We have that

$$\begin{aligned}
\max\{u_1, \ldots, u_n\} - \max\{v_1, \ldots, v_n\} &= u_1 - v_i \\
&\leq u_1 - v_1 \\
&\leq |u_1 - v_1| \\
&\leq \max\{|u_1 - v_1|, \ldots, |u_n - v_n|\}.
\end{aligned}$$

Similarly, we obtain

$$\begin{aligned}
\max\{v_1, \ldots, v_n\} - \max\{u_1, \ldots, u_n\} &= v_i - u_1 \\
&\leq v_i - u_i \\
&\leq |u_i - v_i| \\
&\leq \max\{|u_1 - v_1|, \ldots, |u_n - v_n|\}.
\end{aligned}$$

This proves the statement. □

Proof. Because of Proposition 3.1.2, it will suffice to prove that the maximum function $\max : \mathbb{R}^n \to \mathbb{R}$, that maps $x = (x_1, \ldots, x_n)$ to $\max\{x_1, \ldots, x_n\}$, is a non-expansive function. Let us consider two n-tuples $x = (x_1, \ldots, x_n)$ and $y = (x_1, \ldots, x_n)$ of real numbers. Lemma 3.1.4 implies that:

$$\begin{aligned}
|\max(x) - \max(y)|| &= |\max\{x_1, \ldots, x_n\} - \max\{y_1, \ldots, y_n\}| \\
&\leq \max|x_1 - y_1|, \ldots, |x_n - y_n| \\
&= \|x - y\|_\infty.
\end{aligned}$$

Hence, the maximum function is a non-expansive function. □

Translation operator Let F be a GENEO from (Φ, G) to (Ψ, H) with respect to $T\colon G \to H$ and $b \in \mathbb{R}$. Assume that $\mathrm{dom}(\Psi) = Y$. We can consider the function

$$F_b(\varphi) := F(\varphi) - b$$

from Φ to $C^0(Y, \mathbb{R})$.

Proposition 3.1.5 *If $F_b(\Phi) \subseteq \Psi$ then the operator F_b is a GENEO from (Φ, G) to (Ψ, H) with respect to T.*

Proof. One could easily check that the function $S_b\colon \mathbb{R} \to \mathbb{R}$, such that $S_b(x) := x - b$, is an isometry. Then, because of Proposition 3.1.2, F_b is a GENEO from (Φ, G) to (Ψ, H) with respect to T. □

Affine combination operator We can restate Lemma 2.1.1 as a corollary of Proposition 3.1.2. Let $F_1, \dots, F_n$ be GENEOs from (Φ, G) to (Ψ, H) with respect to $T\colon G \to H$, where $\mathrm{dom}(\Psi) = Y$. Consider $(a_1, \dots, a_n) \in \mathbb{R}^n$ with $\sum_{i=1}^n |a_i| \leq 1$. We can consider the function

$$F_\Sigma(\varphi) := \sum_{i=1}^n a_i F_i(\varphi)$$

from Φ to $C^0(Y, \mathbb{R})$.

Proposition 3.1.6 *If $F_\Sigma(\Phi) \subseteq \Psi$, then F_Σ is a GENEO from (Φ, G) to (Ψ, H) with respect to T.*

Proof. Fix $(a_1, \dots, a_n) \in \mathbb{R}^n$ with $\sum_{i=1}^n |a_i| \leq 1$. Because of Proposition 3.1.2, it will suffice to prove that the function $\Sigma : \mathbb{R}^n \to \mathbb{R}$, that maps $x = (x_1, \dots, x_n)$ to $\sum_{i=1}^n a_i x_i$, is a non-expansive function. Let us consider two n-tuples $x = (x_1, \dots, x_n)$ and $y = (x_1, \dots, x_n)$ of real numbers. We have that:

$$\begin{aligned} |\Sigma(x) - \Sigma(y)| &= \left| \sum_{i=1}^n a_i x_i - \sum_{i=1}^n a_i y_i \right| \\ &= \left| \sum_{i=1}^n a_i (x_i - y_i) \right| \\ &\leq \sum_{i=1}^n |a_i| |x_i - y_i| \\ &\leq \sum_{i=1}^n |a_i| \|x - y\|_\infty \\ &\leq \|x - y\|_\infty . \end{aligned}$$

Hence, Σ is a non-expansive function. □

Power mean operator In order to apply Proposition 3.1.2, we recall some definitions and properties about power means and p-norms. Let us consider a sample of real numbers $x_1, \dots, x_n$ and a real number $p > 0$. As well known, the power mean $M_p(x_1, \dots, x_n)$ of $x_1, \dots, x_n$ is defined by setting

$$M_p(x_1, \dots, x_n) := \left(\frac{1}{n} \sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}.$$

In order to proceed, we consider the function $\| \cdot \|_p : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by setting

$$\|x\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$$

where $x = (x_1, \dots, x_n)$ is a point of $\mathbb{R}^n$. It is well know that, for $p \geq 1$, $\| \cdot \|_p$ is a norm and that for any $x \in \mathbb{R}^n$, we have $\lim_{p \to \infty} \|x\|_p = \|x\|_\infty$. Finally, it is easy to check that if $x \in \mathbb{R}^n$ and $0 < p < q < \infty$, it holds that

$$\|x\|_q \leq \|x\|_p \leq n^{\frac{1}{p} - \frac{1}{q}} \|x\|_q. \tag{3.5}$$

For q tending to infinity, we obtain a similar inequality:

$$\|x\|_\infty \leq \|x\|_p \leq n^{\frac{1}{p}} \|x\|_\infty. \tag{3.6}$$

Now we can define a new class of GENEOs. Let us consider $F_1, \dots, F_n$ GENEOs from (Φ, G) to (Ψ, H) with respect to $T \colon G \to H$, where $\mathrm{dom}(\Psi) = Y$. Consider also a real number $p > 0$. Let us define the operator $M_p(F_1, \dots, F_n) : \Phi \longrightarrow C_b^0(Y, \mathbb{R})$ by setting

$$M_p(F_1, \dots, F_n)(\varphi)(x) := M_p(F_1(\varphi)(x), \dots, F_n(\varphi)(x)).$$

Proposition 3.1.7 *If $p \geq 1$ and $M_p(F_1, \dots, F_n)(\Phi) \subseteq \Psi$, $M_p(F_1, \dots, F_n)$ is a GENEO from (Φ, G) to (Ψ, H) with respect to T.*

Proof. If we show that M_p is a non-expansive function for $p \geq 1$, Proposition 3.1.2 will ensure us that $M_p(F_1, \dots, F_n)$ is a GENEO.

Let $p \geq 1$ and $x, y \in \mathbb{R}^n$. Since $\| \cdot \|_p$ is a norm, the reverse triangle inequality holds. Therefore, because of inequality (3.6) we have that:

$$\begin{aligned}
\left| \left(\frac{1}{n} \sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} - \left(\frac{1}{n} \sum_{i=1}^{n} |y_i|^p \right)^{\frac{1}{p}} \right| &= \left(\frac{1}{n} \right)^{\frac{1}{p}} \left| \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} - \left(\sum_{i=1}^{n} |y_i|^p \right)^{\frac{1}{p}} \right| \\
&= \left(\frac{1}{n} \right)^{\frac{1}{p}} \left| \|x\|_p - \|y\|_p \right| \\
&\leq \left(\frac{1}{n} \right)^{\frac{1}{p}} \|x - y\|_p \\
&\leq \left(\frac{1}{n} \right)^{\frac{1}{p}} n^{\frac{1}{p}} \|x - y\|_\infty = \|x - y\|_\infty.
\end{aligned}$$

Hence, for $p \geq 1$ M_p is non-expansive and the statement of our theorem is proved. □

Remark 3.1.8. If $0 < p < 1$ and $n > 1$, M_p is not a 1-Lipschitz function. This can be easily proved by showing that for $x_2 = x_3 = \cdots = x_n = 1$ the derivative $\frac{\partial M_p}{\partial x_1}$ is not bounded.

3.2 Series of GENEOs

In this subsection we will investigate a method to build GENEOs starting from an infinite set of known GENEOs. A first natural approach to this problem is to study series of GENEOs. First we recall some well-known results about series of functions.

Theorem 11 *Let (a_k) be a positive real sequence such that (a_k) is decreasing and $\lim_{k\to\infty} a_k = 0$. Let (g_k) be a sequence of bounded functions from the topological space X to $\mathbb{C}$. If there exists a real number $M > 0$ such that*

$$\left|\sum_{k=1}^{n} g_k(x)\right| \leq M \tag{3.7}$$

for every $x \in X$ and every $n \in \mathbb{N}$, then the series $\sum_{k=1}^{\infty} a_k g_k$ is uniformly convergent on X.

The second result ensures us that a uniformly convergent series of continuous functions is a continuous function.

Theorem 12 *Let (f_n) be a sequence of continuous function from a compact topological space X to $\mathbb{R}$. If the series $\sum_{k=1}^{\infty} f_k$ is uniformly convergent, then $\sum_{k=1}^{\infty} f_k$ is continuous from X to $\mathbb{R}$.*

Now we can define a series of GENEOs. Consider two perception pairs (Φ, G) and (Ψ, H). Assume that the domain (Y, D_Y) of (Ψ, H) is a compact pseudo-metric space. Let (a_k) be a positive real sequence such that (a_k) is decreasing and $\sum_{k=1}^{\infty} a_k \leq 1$. Suppose that a sequence (F_k) of GENEOs from (Φ, G) to (Ψ, H) with respect to $T\colon G \to H$ is given and that for any $\varphi \in \Phi$ there exists $M(\varphi) > 0$ such that

$$\left|\sum_{k=1}^{n} F_k(\varphi)(x)\right| \leq M(\varphi) \tag{3.8}$$

for every $x \in X$ and every $n \in \mathbb{N}$. These assumptions fulfill the hypotheses of the previous theorems and ensure that the following operator is well-defined. Let us consider the operator $F : \Phi \longrightarrow C_b^0(Y, \mathbb{R})$ defined by setting

$$F(\varphi) := \sum_{k=1}^{\infty} a_k F_k(\varphi). \tag{3.9}$$

Proposition 3.2.1 *If $F(\Phi) \subseteq \Psi$, then F is a GENEO from (Φ, G) to (Ψ, H) with respect to T.*

Proof. Let $g \in G$. Since F_k is T-equivariant for any k and g is uniformly continuous (because it is an isometry), F is T-equivariant:

$$\begin{aligned} F(\varphi g) &= \sum_{k=1}^{\infty} a_k F_k(\varphi g) \\ &= \sum_{k=1}^{\infty} a_k (F_k(\varphi) g) \\ &= \left(\sum_{k=1}^{\infty} a_k F_k(\varphi) \right) g \\ &= F(\varphi) g \end{aligned}$$

for any $\varphi \in \Phi$. Since F_k is non-expansive for any k and $\sum_{k=1}^{\infty} a_k \leq 1$, F is non-expansive:

$$\begin{aligned} \|F(\varphi_1) - F(\varphi_2)\|_\infty &= \left\| \sum_{k=1}^{\infty} a_k F_k(\varphi_1) - \sum_{k=1}^{\infty} a_k F_k(\varphi_2) \right\|_\infty \\ &= \left\| \lim_{n \to \infty} \left(\sum_{k=1}^{n} a_k F_k(\varphi_1) - \sum_{k=1}^{n} a_k F_k(\varphi_2) \right) \right\|_\infty \\ &= \lim_{n \to \infty} \left\| \sum_{k=1}^{n} a_k (F_k(\varphi_1) - F_k(\varphi_2)) \right\|_\infty \\ &\leq \lim_{n \to \infty} \sum_{k=1}^{n} (a_k \|F_k(\varphi_1) - F_k(\varphi_2)\|_\infty) \\ &\leq \lim_{n \to \infty} \sum_{k=1}^{n} (a_k \|\varphi_1 - \varphi_2\|_\infty) \\ &= \sum_{k=1}^{\infty} a_k \|\varphi_1 - \varphi_2\|_\infty \\ &\leq \|\varphi_1 - \varphi_2\|_\infty \end{aligned}$$

for any $\varphi_1, \varphi_2 \in \Phi$.

□

3.3 Building GENEOs via Permutants and Permutant Measures

This section is devoted to introducing a new method to construct GENEOs by means of particular subsets of $\mathrm{Aut}_\Phi(X)$, called permutants, and their natural generalisation, the concept of permutant measure. We underline that in those methods we can treat the group of invariance as a variable of the problem. This is important because the change of the observer generally corresponds to a change of the invariance we want to analyze. In this section, for the sake of simplicity we fix a perception pair (Φ, G), where $\Phi \subseteq \mathbb{R}_b^X$ and $G \subseteq \mathrm{Aut}_\Phi(X)$.

3.3.1 Permutants

In this subsection we introduce a new method for the construction of GENEOs, exploiting the concept of permutant. Consider the conjugation action of G on $\mathrm{Aut}_\Phi(X)$:

$$\begin{aligned} \alpha : G \times \mathrm{Aut}_\Phi(X) &\to \mathrm{Aut}_\Phi(X) \\ (g, f) &\mapsto gfg^{-1}. \end{aligned}$$

Fixing a conjugating element $g \in G$, we can also consider the inner automorphism α_g:

$$\begin{aligned} \alpha_g : \mathrm{Aut}_\Phi(X) &\to \mathrm{Aut}_\Phi(X) \\ f &\mapsto gfg^{-1}. \end{aligned}$$

Definition 3.3.1. We say that a subset $H \subseteq \mathrm{Aut}_\Phi(X)$ is a *permutant* for G if either $H = \emptyset$ or $gHg^{-1} = H$ for every $g \in G$.

Note that a subset H of $\mathrm{Aut}_\Phi(X)$ is a permutant for G if and only if H is a union of orbits for the conjugation action of G on $\mathrm{Aut}_\Phi(X)$. Let us denote by $\mathrm{Perm}(G)$ the set of all permutants for G.

Example 3.3.2. Let $X = \mathbb{R}$ and $\Phi = \mathbb{R}_b^{\mathbb{R}}$. Obviously, $\mathrm{Aut}_\Phi(\mathbb{R}) = \mathrm{Aut}(\mathbb{R})$, where $\mathrm{Aut}(\mathbb{R})$ is the set of all bijections from $\mathbb{R}$ to itself. Consider the set G of all Euclidean isometries of the real line, i.e. the maps from $\mathbb{R}$ to itself of the form $g(x) = ax + b,\ a, b \in \mathbb{R},\ a = \pm 1$. Fix a real number t. Take the translation $h(x) = x + t$ and its inverse $h^{-1} = x - t$. One could easily check that $H = \{h, h^{-1}\}$ is a permutant for G.

Example 3.3.3. Let us consider the unit circle $S^1 := \{(x, y) \in \mathbb{R}^2 :\ x^2 + y^2 = 1\}$. We also assume that Φ is the set of non-expansive functions from S^1 to $\mathbb{R}$, where S^1 and $\mathbb{R}$ are endowed with the Euclidean distance. Let G be the group generated by the reflection with respect to the line $x = 0$. We consider $H = \{\mathrm{id}_{S^1}, \rho, \rho^2, \rho^3\}$, where ρ is the rotation of $\pi/2$ around the origin $(0, 0)$. It would be easily to check that H is a permutant for G.

In the sequel, we will use the cycle notation to represent permutations.

Example 3.3.4. Consider $X = \{1, 2, 3, 4, 5, 6, 7\}$ and the space $[0, 1]^X$ of all functions from X to the unit interval $[0, 1]$. Note that $\mathrm{Aut}_{[0,1]^X}(X)$ coincides with the symmetric group S_7 over X, endowed with the discrete topology. Moreover, $G_{1,2,3}$ denotes the subset of S_7 that contains the non-trivial permutations that fix the last four elements. Namely,

$$G_{1,2,3} = \{(1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}.$$

Similarly, $G_{4,5,6}$ denotes the subset of S_7 that contains the non-trivial permutations that fix the last four elements. Namely,

$$G_{4,5,6} = \{(4\ 5), (4\ 6), (5\ 6), (4\ 5\ 6), (4\ 6\ 5)\}.$$

Finally, let $G = G_{1,2,3} \cup G_{4,5,6} \cup \{\mathrm{id_X}\}$. One could easily check that $G_{1,2,3}$ and $G_{4,5,6}$ are permutant for G.

Example 3.3.5. Consider the sphere $S^2 := \{(x, y, z) \in \mathbb{R}^3 : \ x^2 + y^2 + z^2 = 1\}$ and the space $[0,1]^{S^2}$ of all functions from S^2 to the unit interval $[0, 1]$. Note that $\mathrm{Aut}_{[0,1]^{S^2}}(S^2) = \mathrm{Aut}(S^2)$. Take the group G of all Euclidean isometries of S^2 as invariance group. Assume that θ is a real number in $[0, 2\pi]$. Since any two rotations through the same angle around two oriented axes are conjugate with respect to the action of G, the set $H(\theta)$ of all rotations through the angle θ is a permutant.

If $H = \{h_1, \dots, h_n\}$ is a finite permutant for G, we can define a GENEOs associated with H in the following way. Let $\bar{a}$ be a real number with $n|\bar{a}| \leqslant 1$, we can consider the operator $F_{\bar{a},H} : \mathbb{R}^X_b \to \mathbb{R}^X_b$ defined by setting

$$F_{\bar{a},H}(\varphi) \coloneqq \bar{a} \sum_{i=1}^{n} \varphi h_i.$$

The following statement holds.

Proposition 3.3.6 *Assume that H is a finite permutant. If $F_{\bar{a},H}(\Phi) \subseteq \Phi$ then $F_{\bar{a},H}$ is a GENEO from (Φ, G) to itself with respect to the identity $\mathrm{id}_G \colon G \to G$.*

Proof. First of all, we prove that $F_{\bar{a},H}$ is id_G-equivariant. Let $\tilde{\alpha}_g : \{1, \dots, n\} \to \{1, \dots, n\}$ be an index permutation such that $\tilde{\alpha}_g(i)$ is the index of the image of h_i through the conjugacy action of g, i.e.

$$\alpha_g(h_i) = g h_i g^{-1} = h_{\tilde{\alpha}_g(i)}, \quad \forall\ i \in \{1, \dots, n\}.$$

We obtain that

$$g h_i = h_{\tilde{\alpha}_g(i)} g.$$

Exploiting this relation we obtain that

$$\begin{aligned} F_{\bar{a},H}(\varphi g) =& \bar{a}(\varphi g h_1 + \dots + \varphi g h_n) \\ =& \bar{a}(\varphi h_{\tilde{\alpha}_g(1)} g + \dots + \varphi h_{\tilde{\alpha}_g(n)} g). \end{aligned}$$

Since $\left\{h_{\tilde{\alpha}_g(1)}, \dots, h_{\tilde{\alpha}_g(n)}\right\} = \{h_1, \dots, h_n\}$, we get

$$F_{\bar{a},H}(\varphi g) = F_{\bar{a},H}(\varphi) g, \quad \forall\ \varphi \in \Phi, \quad \forall\ g \in G.$$

It remains to show that $F_{\bar{a},H}$ is non-expansive:

$$\begin{aligned} \|F_{\bar{a},H}(\varphi_1) - F_{\bar{a},H}(\varphi_2)\|_\infty &= \left\| \bar{a} \sum_{i=1}^{n} (\varphi_1 h_i) - \bar{a} \sum_{i=1}^{n} (\varphi_2 h_i) \right\|_\infty \\ &= |\bar{a}| \left\| \sum_{i=1}^{n} (\varphi_1 h_i - \varphi_2 h_i) \right\|_\infty \\ &\leqslant |\bar{a}| \sum_{i=1}^{n} \|\varphi_1 h_i - \varphi_2 h_i\|_\infty \\ &= |\bar{a}| \sum_{i=1}^{n} \|\varphi_1 - \varphi_2\|_\infty \\ &= n|\bar{a}| \|\varphi_1 - \varphi_2\|_\infty \\ &\leqslant \|\varphi_1 - \varphi_2\|_\infty \end{aligned}$$

for any $\varphi_1, \varphi_2 \in \Phi$. □

Remark 3.3.7. Obviously $H = \{\mathrm{id}_X : X \to X\} \subseteq \mathrm{Aut}_\Phi(X)$ is a permutant for every subgroup G of $\mathrm{Aut}_\Phi(X)$, but the use of Proposition 3.3.6 for this trivial permutant leads to the trivial operator given by a multiple of the identity operator on Φ.

Remark 3.3.8. If the group G is Abelian, every finite subset of G is a permutant for G, since the conjugacy action is just the identity. Hence in this setting, for any chosen finite subset $H = \{g_1, \ldots, g_n\}$ of G and any real number $\bar{a}$, such that $n|\bar{a}| \leqslant 1$, $F_{\bar{a},H}(\varphi) = \bar{a}(\varphi g_1 + \cdots + \varphi g_n)$ is a GENEO from (Φ, G) to itself, provided that $F_{\bar{a},H}$ preserves Φ.

Remark 3.3.9. The operator $F_{\bar{a},H} : \Phi \to \Phi$ introduced in Proposition 3.3.6 is linear, provided that Φ is linearly closed. Indeed, assume that a permutant $H = \{h_1, \ldots, h_n\}$ for G and a real number $\bar{a}$ such that $n|\bar{a}| \leqslant 1$ are given. Let us consider the associated operator $F_{\bar{a},H}(\varphi) = \bar{a}\sum_{i=1}^{n}(\varphi h_i)$, and assume that $F_{\bar{a},H}(\Phi) \subseteq \Phi$. If $\lambda_1, \lambda_2 \in \mathbb{R}$ and $\varphi_1, \varphi_2 \in \Phi$, we have

$$
\begin{aligned}
F_{\bar{a},H}(\lambda_1\varphi_1 + \lambda_2\varphi_2) =& \bar{a}\sum_{i=1}^{n}((\lambda_1\varphi_1 + \lambda_2\varphi_2)h_i) \\
=& \bar{a}\sum_{i=1}^{n}(\lambda_1(\varphi_1 h_i) + \lambda_2(\varphi_2 h_i)) \\
=& \bar{a}\sum_{i=1}^{n}\lambda_1(\varphi_1 h_i) + \bar{a}\sum_{i=1}^{n}\lambda_2(\varphi_2 h_i) \\
=& \lambda_1\left[\bar{a}\sum_{i=1}^{n}(\varphi_1 h_i)\right] + \lambda_2\left[\bar{a}\sum_{i=1}^{n}(\varphi_2 h_i)\right] \\
=& \lambda_1 F_{\bar{a},H}(\varphi_1) + \lambda_2 F_{\bar{a},H}(\varphi_2).
\end{aligned}
$$

3.3.2 Some results concerning permutants

When H contains only the identical homeomorphism, the operator $F_{\bar{a},H}$ is trivial, since it is the multiple by the constant $\bar{a}$ of the identical operator. This section highlights that in some cases this situation cannot be avoided, since non-trivial finite permutants for G are not available. In order to illustrate this problem, we need to introduce the concept of *versatile* group.

Definition 3.3.10. Let G be a group that acts on a set X. We say that G is *versatile* if for every triple $(x, y, z) \in X^3$, with $x \neq z$, and for every finite subset S of X, at least one element $g \in G$ exists such that (1) $g(x) = y$ and (2) $g(z) \notin S$.

Proposition 3.3.11 *Assume that $H = \{h_1, \ldots, h_n\}$ is a permutant for a subgroup G of $\mathrm{Aut}_\Phi(X)$. If G is versatile, then $H = \{\mathrm{id}_X\}$.*

Proof. It is sufficient to prove that if H contains an element $h \neq \mathrm{id}_X$, then G is not versatile. We can assume that $h \equiv h_1$. Since h_1 is different from the identity, a point $\bar{x} \in X$ exists such that $h_1(\bar{x}) \neq \bar{x}$. Let us consider the triple $(h_1(\bar{x}), \bar{x}, \bar{x})$ and the set $S = \{h_1^{-1}(\bar{x}), \ldots, h_n^{-1}(\bar{x})\}$. Suppose that $g \in G$ satisfies Property (1) with respect to the previous triple, that is $g(h_1(\bar{x})) = \bar{x}$. Since the conjugacy action of

g on H is a permutation, we can find an element $h_2 \in H$ such that $h_2 = gh_1g^{-1}$, so that $h_2(g(\bar{x})) = g(h_1(\bar{x})) = \bar{x}$ and hence $g(\bar{x}) = h_2^{-1}(\bar{x}) \in S$. Therefore, g does not satisfy Property (2), for $z = \bar{x}$. Hence we can conclude that no $g \in G$ exists verifying both Properties (1) and (2), i.e. G is not versatile. □

Remark 3.3.12. Definition 3.3.10 immediately implies that if G, G' are two subgroups of $\mathrm{Aut}_\Phi(X)$, $G \subseteq G'$ and G is versatile, then also the group G' is versatile. For example, it is easy to prove that the group G of the isometries of the real plane is versatile. It follows that every group G' of self-homeomorphisms of $\mathbb{R}^2$ containing the isometries of the real plane is versatile. As a consequence of Proposition 3.3.11, every permutant for G' is trivial.

The following definition extends the one of versatile group and is of use in studying permutants. Let $\mathrm{Aut}(X)$ be the set of all bijections from X to itself.

Definition 3.3.13. If k is a positive integer, we say that the group $G \subseteq \mathrm{Aut}(X)$ is *k-weakly versatile* if for every pair $(x, z) \in X \times X$ with $x \neq z$ and every subset S of X with $|S| \leq k$, a $g \in G$ exists such that $g(x) = x$ and $g(z) \notin S$.

The previous definition allows us to highlight an interesting property of permutants.

Lemma 3.3.14 *If G is k-weakly versatile, then every permutant $H \neq \emptyset, \{\mathrm{id}_X\}$ has cardinality strictly greater than k.*

Proof. By contradiction, assume that a non-empty permutant $H = \{h_1, \ldots, h_r\} \neq \{\mathrm{id}_X\}$ exists, with $1 \leq r \leq k$. Since $H \neq \{\mathrm{id}_X\}$, we can assume that h_1 is not the identity. Let us take a point $x \in X$ such that $h_1(x) \neq x$ and set $z := h_1(x)$, $S := \{h_1(x), \ldots, h_r(x)\}$. G is k-weakly versatile and hence a $g \in G$ exists, such that $g(x) = x$ and $g(z) \notin S$. Since $gHg^{-1} = H$, an index i exists such that $gh_1 = h_ig$. It follows that $g(z) = g(h_1(x)) = h_i(g(x)) = h_i(x) \in S$, against the assumption that $g(z) \notin S$. □

3.3.3 Permutant measures

In Proposition 3.3.6 only finite permutants are involved. The aim of this subsection is to generalize the concept of permutant in order to be able to consider also infinite permutants in the construction of new GENEOs.

Definition 3.3.15. A finite Borel signed measure μ on $(\mathrm{Aut}_\Phi(X), D_{\mathrm{Aut}})$ is called a *permutant measure* with respect to G if μ is invariant under the conjugation action of G (i.e., $\mu(H) = \mu(gHg^{-1})$ for every $g \in G$ and every Borel set H in $\mathrm{Aut}_\Phi(X)$).

Example 3.3.16. If $\mathrm{Aut}_\Phi(X)$ is a compact group, then the Haar measure on $\mathrm{Aut}_\Phi(X)$ is a permutant measure.

However, there exist permutant measures that are not Haar measures, as you can see in the next example.

Example 3.3.17. Let X be the set of the vertices of a cube in $\mathbb{R}^3$. Let us consider $\Phi = \mathbb{R}^X$, and the group G of the orientation-preserving isometries of $\mathbb{R}^3$ that take X to X. We note that D_X induces the discrete topology on X. Let π_1, π_2, π_3 be the three planes that contain the center of mass of X and are parallel to a face of the cube. Let $h_i : X \to X$ be the orthogonal symmetry with respect to π_i, for $i \in \{1, 2, 3\}$. We have that the set $\{h_1, h_2, h_3\}$ is an orbit under the conjugation action of G. We can now define a permutant measure μ_2 on the group $\mathrm{Aut}_\Phi(X)$ by setting $\mu_2(h_1) = \mu_2(h_2) = \mu_2(h_3) = c$, where c is a positive real number, and $\mu_2(h) = 0$ for any $h \in \mathrm{Aut}_\Phi(X)$ with $h \notin \{h_1, h_2, h_3\}$. We also observe that while the cardinality of G is 24, the cardinality of the support $\mathrm{supp}(\mu_2) := \{h \in \mathrm{Aut}_\Phi(X) : \mu_2(h) \neq 0\}$ of the signed measure μ_2 is 3.

Now, we would like recall that $(\mathbb{R}_b^X, \|\cdot\|_\infty)$ is a Banach space. Consider the function $L_\varphi : \mathrm{Aut}_\Phi(X) \to \Phi$ defined as $L_\varphi(g) := \varphi g$ for each $g \in G$. Note that L_φ is non-expansive. For Bochner integral and related concepts, we refer the interested reader to Appendix B. In particular, we can state that, if Φ is contained in a linear separable space, L_φ is strongly μ-measurable, since it is continuous. This condition on Φ is fulfilled in interesting and imporant examples: when Φ is finite dimensional or when Φ is compact. This is due to the following Lemma. Before proceeding, let us recall that $\mathrm{span}_\mathbb{K}(S)$ is the vector space generated by S with coefficients in the field $\mathbb{K}$.

Lemma 3.3.18 *Let $(X, \|\cdot\|)$ be a normed linear space. X is separable if and only if there exists a compact subset K such that $\overline{\mathrm{span}_\mathbb{R}(K)} = X$.*

Proof. First, assume that X is a separable normed linear space. Note that if $(p_i)_{i\in\mathbb{N}}$ is a sequence converging to p in X, then the set $D = \{p_i : i \in \mathbb{N}\} \cup \{p\} \subseteq X$ is compact. Let $X^* = \{y_i : i \in \mathbb{N}\}$ be a countable dense subset of X. Now, set $x_i = \frac{1}{1+(i+1)\|y_i\|} y_i$ for $i \in \mathbb{N}$. Then we have $\|x_i\| < \frac{1}{i+1}$, and $x_i \to 0$. Consider the compact set $K = \{x_i : i \in \mathbb{N}\} \cup \{0\}$. Since $y_i \in \mathrm{span}_\mathbb{R}(K)$, it follows that $\overline{\mathrm{span}_\mathbb{R}(K)} \supseteq \overline{D} = X$.

Now, in order to complete the proof it will suffice to show that if K is compact, then $\overline{\mathrm{span}_\mathbb{R}(K)}$ is a separable linear space. Since K is compact, we can take a countable dense subset K^* of K. Let us consider the space $\mathrm{span}_\mathbb{Q}(K^*)$, which is the finite linear combinations of elements in K^* with rational coefficients. One could easily check that $\mathrm{span}_\mathbb{Q}(K^*)$ is a countable set. Let us consider an element $x \in \mathrm{span}_\mathbb{R}(K)$. By definition $\mathrm{span}_\mathbb{R}(K)$, we can write $x = \sum_{i=1}^n \alpha_i k_i$, where $\alpha_i \in \mathbb{R}$ and $k_i \in K$. Let us fix a positive real number ε. Since K^* is dense in K and $\mathbb{Q}$ is dense $\mathbb{R}$, for every k_i and α_i, there exist $\tilde{k}_i \in K^*$ and $\tilde{\alpha}_i \in \mathbb{Q}$ such that $\|k_i - \tilde{k}_i\| < \varepsilon$

and $|\alpha_i - \tilde{\alpha}_i| < \varepsilon$. We have that:

$$\begin{aligned}\left\|\sum_{i=1}^{n} \alpha_i k_i - \sum_{i=1}^{n} \tilde{\alpha}_i \tilde{k}_i\right\| &= \left\|\sum_{i=1}^{n} (\alpha_i k_i - \tilde{\alpha}_i k_i + \tilde{\alpha}_i k_i - \tilde{\alpha}_i \tilde{k}_i)\right\| \\ &\leq \sum_{i=1}^{n} |\alpha_i - \tilde{\alpha}_i| \|k_i\| + \sum_{i=1}^{n} |\tilde{\alpha}_i| \|k_i - \tilde{k}_i\| \\ &\leq \varepsilon \sum_{i=1}^{n} \|k_i\| + \varepsilon \sum_{i=1}^{n} |\tilde{\alpha}_i - \alpha_i + \alpha_i| \\ &\leq \varepsilon \sum_{i=1}^{n} \|k_i\| + \varepsilon \sum_{i=1}^{n} |\tilde{\alpha}_i - \alpha_i| + \varepsilon \sum_{i=1}^{n} |\alpha_i| \\ &\leq \varepsilon \sum_{i=1}^{n} \|k_i\| + n\varepsilon^2 + \varepsilon \sum_{i=1}^{n} |\alpha_i|\end{aligned}$$

Since ε can be arbitrarily small, we can say that $\mathrm{span}_{\mathbb{R}}(K) \subseteq \overline{\mathrm{span}_{\mathbb{Q}}(K^*)}$. Hence, $\overline{\mathrm{span}_{\mathbb{R}}(K)} = \overline{\mathrm{span}_{\mathbb{Q}}(K^*)}$. This implies that $\overline{\mathrm{span}_{\mathbb{R}}(K)}$ contains the countable set $\mathrm{span}_{\mathbb{Q}}(K^*)$, and the statement is proved. □

Moreover, when Φ is compact or $\Phi = \mathbb{R}^X$, where X is finite set, Proposition B.0.5 implies L_φ is μ-Bochner integrable, since $L_\varphi(\mathrm{Aut}_\Phi(X))$ is bounded for Theorem 4 and uniform continuity of L_φ. Assume Φ is contained in a linear separable space and L_φ is μ-Bochner integrable. Let H be μ-measurable permutant for G with μ is a permutant measure, and $\mu(H) \neq 0$. We consider the operator $F_H \colon \Phi \to \mathbb{R}_b^X$ defined as

$$F_H(\varphi)(x) := \frac{1}{\mu(H)} \int_H L_\varphi(h)\; d\mu(h).$$

Since L_φ is strongly μ-measurable and μ is finite, F_H is well defined for every $\varphi \in \Phi$.

Proposition 3.3.19 *(F_H, id_G) is a GENEO from (Φ, G) to $(\mathbb{R}_b^X, G)$.*

Proof. It will suffice to prove that F_H is id_G-equivariant and non-expansive. For any $\varphi \in \Phi$ and $g \in G$, we have that:

$$\begin{aligned}F_H(\varphi g) &= \frac{1}{\mu(H)} \int_H L_{\varphi g}(h)\; d\mu(h) \\ &= \frac{1}{\mu(H)} \int_H \varphi g h\; d\mu(h) \\ &= \frac{1}{\mu(H)} \int_H \varphi g g^{-1} h' g\; d\mu(g^{-1} h' g) \\ &= \frac{1}{\mu(H)} \int_H \varphi h' g\; d\mu(h') \\ &= \left(\frac{1}{\mu(H)} \int_H \varphi h'\; d\mu(h')\right) g \\ &= F_H(\varphi) g.\end{aligned}$$

Hence, F_H is id_G-equivariant. Moreover, for any $\varphi_1, \varphi_2 \in \Phi$:

$$\begin{aligned}
\|F_H(\varphi_1) - F_H(\varphi_2)\|_\infty &= \left\| \frac{1}{\mu(H)} \int_H L_{\varphi_1} \, d\mu - \frac{1}{\mu(H)} \int_H L_{\varphi_2} \, d\mu \right\|_\infty \\
&= \left\| \frac{1}{\mu(H)} \int_H L_{\varphi_1} - L_{\varphi_2} \, d\mu \right\|_\infty \\
&\leq \frac{1}{\mu(H)} \int_H \|L_{\varphi_1} - L_{\varphi_2}\|_\infty \, d\mu \\
&\leq \frac{1}{\mu(H)} \int_H \|\varphi_1 h - \varphi_2 h\|_\infty \, d\mu(h) \\
&= \|\varphi_1 - \varphi_2\|_\infty \frac{1}{\mu(H)} \int_H \, d\mu(h) \\
&= \|\varphi_1 - \varphi_2\|_\infty .
\end{aligned}$$

Thus, F_H is non-expansive and the statement is proved. □

3.4 Representation of linear GENEOs via permutant measures

A natural question arises from Proposition 3.3.19: Which linear GENEOs can be represented as GENEOs associated with a permutant measure? In this section we will see that, under suitable assumption, it is possible to represent every linear GENEO (and GEO) by means of permutant measures.

Let $\mathbb{R}^X \cong \mathbb{R}^n$ be the vector space of all functions from a finite set $X = \{x_1, \dots, x_n\}$ to $\mathbb{R}$. We would like to recall that $\mathbb{R}^X$ has the canonical basis $\{\mathbb{1}_{x_j}\}_j$, where $\mathbb{1}_x \colon X \to \mathbb{R}$ is the function taking the value 1 at x and the value 0 at every point y with $y \neq x$. We also consider the group $\mathrm{Aut}(X)$ of all permutations on X and a subgroup G of $\mathrm{Aut}(X)$. Note that $\mathrm{Aut}(X) = \mathrm{Aut}_{\mathbb{R}^X}(X)$. We recall that $\mathbb{R}^X$ is endowed with the L^∞-norm.

Remark 3.4.1. In this case, D_X induces the discrete topology on X. If we endow X with the discrete topology, $\mathbb{R}^X$ coincides with $C^0(X, \mathbb{R})$.

For the sake of simplicity, in the following we give another definition of permutant measure that, under our assumptions, is equivalent to the Definition 3.3.15:

Definition 3.4.2. A finite signed measure μ on $\mathrm{Aut}(X)$ is called a *permutant measure* with respect to G if each subset H of $\mathrm{Aut}(X)$ is measurable and μ is invariant under the conjugation action of G (i.e., $\mu(H) = \mu(gHg^{-1})$ for every $g \in G$). Equivalently, we can say that a signed measure μ on $\mathrm{Aut}(X)$ is a permutant measure with respect to G if each singleton $\{h\} \subseteq \mathrm{Aut}(X)$ is measurable and $\mu(\{h\}) = \mu(\{ghg^{-1}\})$ for every $g \in G$.

With a slight abuse of notation, we will denote by $\mu(h)$ the signed measure of the singleton $\{h\}$ for each $h \in \mathrm{Aut}(X)$.

Example 3.4.3. Let us consider a positive integer number n and the finite set

$$X := \left\{ \left(\cos \frac{2\pi k}{n}, \sin \frac{2\pi k}{n} \right) \in \mathbb{R}^2 : k \in \mathbb{N}, 0 \leq k \leq n-1 \right\}.$$

Let G be the group of all rotations of X of an angle α around the point $(0,0)$, with α multiple of $\frac{2\pi}{n}$. After fixing an integer number m, consider the map $\bar{h} \in \mathrm{Aut}(X)$ that takes each point $\left(\cos\frac{2\pi k}{n}, \sin\frac{2\pi k}{n}\right)$ to the point $\left(\cos\frac{2\pi(k+m)}{n}, \sin\frac{2\pi(k+m)}{n}\right)$. Moreover, we define the function $\mu_1 \colon \mathcal{P}(\mathrm{Aut}(X)) \to \mathbb{R}$ that takes each subset C of $\mathrm{Aut}(X)$ to 1 if $\bar{h} \in C$ and to 0 if $\bar{h} \notin C$, where $\mathcal{P}(\mathrm{Aut}(X))$ is the power set of $\mathrm{Aut}(X)$. Since the orbit of $\bar{h}$ under the conjugation action of G is the singleton $\{\bar{h}\}$, the function μ_1 is a permutant measure. We also observe that while the cardinality of G is n, the cardinality of the support $\mathrm{supp}(\mu_1) := \{h \in \mathrm{Aut}(X) : \mu_1(h) \neq 0\}$ of the signed measure μ_1 is 1.

Permutant measures give a simple method to build GEOs. Moreover, the following result can be seen as a particular case of Proposition 3.3.19.

Proposition 3.4.4 *If μ is a permutant measure with respect to G, then the map $F_\mu : \mathbb{R}^X \to \mathbb{R}^X$ defined by setting $F_\mu(\varphi) := \sum_{h\in\mathrm{Aut}(X)} \varphi h^{-1}\ \mu(h)$ is a linear GEO.*

Proof. Since $\mathrm{Aut}(X)$ linearly acts on $\mathbb{R}^X$ by composition on the right, F_μ is linear. Moreover, for every $\varphi \in \mathbb{R}^X$ and every $g \in G$

$$\begin{aligned} F_\mu(\varphi g) &= \sum_{h\in\mathrm{Aut}(X)} \varphi g h^{-1}\ \mu(h) \\ &= \sum_{h\in\mathrm{Aut}(X)} \varphi g h^{-1} g^{-1} g\ \mu(g h g^{-1}) \\ &= \sum_{f\in\mathrm{Aut}(X)} \varphi f^{-1} g\ \mu(f) \\ &= F_\mu(\varphi) g, \end{aligned} \tag{3.10}$$

since $\mu(h) = \mu(ghg^{-1})$ and the map $h \mapsto f := ghg^{-1}$ is a bijection from $\mathrm{Aut}(X)$ to $\mathrm{Aut}(X)$. □

Obviously, we have that $F_\mu(\varphi) = \sum_{h\in\mathrm{Aut}(X)} \varphi h^{-1}\ \mu(h) = \sum_{h\in\mathrm{supp}(\mu)} \varphi h^{-1}\ \mu(h)$, where $\mathrm{supp}(\mu) := \{h \in \mathrm{Aut}(X) : \mu(h) \neq 0\}$. In Examples 3.4.3 and 3.3.17 $|\mathrm{supp}(\mu_i)| \ll |G|$ for $i = 1, 2$, and hence in those cases summations on $\mathrm{supp}(\mu_i)$ are simpler than summations on the group G. The condition $|\mathrm{supp}(\mu)| \ll |G|$ is not rare in examples and is the main reason to build GEOs by means of permutant measures, instead of using the representation of GEOs as G-convolutions.

Example 3.4.5. The GEOs associated with the permutant measures defined in Examples 3.4.3 and 3.3.17 are respectively $F_{\mu_1}(\varphi) = \varphi\bar{h}^{-1}$ and $F_{\mu_2}(\varphi) = c\varphi h_1^{-1} + c\varphi h_2^{-1} + c\varphi h_3^{-1}$.

It is interesting to observe that the set $\mathrm{PM}(G)$ of permutant measures with respect to G is a lattice. Indeed, if $\mu_1, \mu_2 \in \mathrm{PM}(G)$, then the measures μ', μ'' on $\mathrm{Aut}(X)$, respectively defined by setting $\mu'(h) := \min\{\mu_1(h), \mu_2(h)\}$ and $\mu''(h) := \max\{\mu_1(h), \mu_2(h)\}$, still belong to $\mathrm{PM}(G)$. Moreover, if $\mu \in \mathrm{PM}(G)$ then $|\mu| \in \mathrm{PM}(G)$. Furthermore, $\mathrm{PM}(G)$ is closed under linear combination. Therefore, $\mathrm{PM}(G)$ has a natural structure of real vector space. We can compute the dimension of $\mathrm{PM}(G)$ by considering the conjugation action of G on $\mathrm{Aut}(X)$.

Proposition 3.4.6 $\dim \mathrm{PM}(G) = |{}^{\mathrm{Aut}(X)}/_G|$.

Proof. Consider a permutant measure μ on $\mathrm{Aut}(X)$. Now, we define the function $f_\mu \colon {}^{\mathrm{Aut}(X)}/_G \to \mathbb{R}$ by setting $f_\mu(\mathcal{O}) = \mu(h)$, where $h \in \mathcal{O}$. Since μ is invariant under the conjugation action of G, f_μ is well defined. One could easily check that the map $\mu \mapsto f_\mu$ is an isomorphism between $\mathrm{PM}(G)$ and the space $\mathbb{R}^{\mathrm{Aut}(X)/G}$ of all real-valued functions on ${}^{\mathrm{Aut}(X)}/_G$. Since $\{\mathbb{1}_\mathcal{O}\}_{\mathcal{O}\in \mathrm{Aut}(X)/G}$ is a basis for $\mathbb{R}^{\mathrm{Aut}(X)/G}$, $\dim \mathbb{R}^{\mathrm{Aut}(X)/G} = |{}^{\mathrm{Aut}(X)}/_G|$. Hence, the statement is proved. □

Proposition 3.4.6 and the well-known Burnside's Lemma imply that $\dim \mathrm{PM}(G) = \frac{1}{|G|}\sum_{g\in G} |\mathrm{Aut}(X)^g|$. We recall that $\mathrm{Aut}(X)^g$ denotes the set of elements fixed by the action of g, i.e., $\mathrm{Aut}(X)^g := \{h \in \mathrm{Aut}(X) | ghg^{-1} = h\}$.

From Proposition 3.4.6 the next corollary immediately follows.

Corollary 3.4.7 $\dim \mathrm{PM}(G) = \log_2 |\mathrm{Perm}(G)|$.

We stress that when the group G becomes larger and larger the lattice $\mathrm{PM}(G)$ becomes smaller and smaller. This duality implies that the method described by Proposition 3.4.4 is particularly interesting when G is large. In some sense, this duality is analogous to the one described in [32, Subsection 3.1].

We can prove the following result.

Theorem 13 *If G transitively acts on X, then for every linear group equivariant operator F for $(\mathbb{R}^X, G)$ a permutant measure μ exists such that*

$$F(\varphi) = F_\mu(\varphi) := \sum_{h\in \mathrm{Aut}(X)} \varphi h^{-1}\, \mu(h)$$

for every $\varphi \in \mathbb{R}^X$, and $\sum_{h\in\mathrm{Aut}(X)} |\mu(h)| = \max_{\varphi\in\mathbb{R}^X\setminus\{\mathbf{0}\}} \frac{\|F(\varphi)\|_\infty}{\|\varphi\|_\infty}$.

In order to prove this statement, let us consider the matrix $B = (b_{ij})$ associated with F with respect to the basis $\{\mathbb{1}_{x_1}, \dots, \mathbb{1}_{x_n}\}$.

Remark 3.4.8. We observe that $\mathbb{1}_x h^{-1} = \mathbb{1}_{h(x)}$ for every $h \in \mathrm{Aut}(X)$ and every $x \in X$.

In the following, for every $g \in G$ we will denote by $\sigma_g : \{1, \dots, n\} \to \{1, \dots, n\}$ the function defined by setting $\sigma_g(j) = i$ if and only if $g(x_j) = x_i$. We observe that $\sigma_{g^{-1}} = \sigma_g^{-1}$.

We need the following lemmas.

Lemma 3.4.9 *An n-tuple of real numbers $\alpha = (\alpha_1, \dots, \alpha_n)$ exists such that each row and each column of B can be obtained by permuting α.*

Proof. Let us choose a function $\mathbb{1}_{x_j}$ and a permutation $g \in G$. By equivariance we have that

$$F(\mathbb{1}_{x_j} g) = F(\mathbb{1}_{x_j})g.$$

The left-hand side of the equation can be rewritten as:

$$F(\mathbb{1}_{x_j} g) = F(\mathbb{1}_{g^{-1}(x_j)}) = \sum_{i=1}^{n} b_{i\sigma_g^{-1}(j)} \mathbb{1}_{x_i}.$$

On the right-hand side we get

$$F(\mathbb{1}_{x_j})g = \left(\sum_{i=1}^{n} b_{ij}\mathbb{1}_{x_i}\right) g = \sum_{i=1}^{n} b_{ij}(\mathbb{1}_{x_i}g) = \sum_{i=1}^{n} b_{ij}(\mathbb{1}_{g^{-1}(x_i)}) = \sum_{s=1}^{n} b_{\sigma_g(s)j}\mathbb{1}_{x_s}$$

by setting $x_s = g^{-1}(x_i)$. Therefore, we obtain the following equation:

$$\sum_{i=1}^{n} b_{i\sigma_g^{-1}(j)}\mathbb{1}_{x_i} = \sum_{s=1}^{n} b_{\sigma_g(s)j}\mathbb{1}_{x_s}.$$

This immediately implies that $b_{i\sigma_g^{-1}(j)} = b_{\sigma_g(i)j}$, for any $i \in \{1,\ldots,n\}$. Since this equality holds for any $j \in \{1,\ldots,n\}$ and any $g \in G$, we have that $b_{ij} = b_{\sigma_g(i)\sigma_g(j)}$ for every $i,j \in \{1,\ldots,n\}$ and every $g \in G$.

Now we are ready to show that all the rows of B are permutations of the first row, and all the columns are permutations of the first column. Since G is transitive, for every $p,q \in \{1,\ldots,n\}$ there exists $g_{pq} \in G$ such that $g_{pq}(x_p) = x_q$. Consider the $\bar{\imath}$-th row of B. We know that $b_{\bar{\imath}j} = b_{\sigma_{g_{\bar{\imath}1}}(\bar{\imath})\sigma_{g_{\bar{\imath}1}}(j)} = b_{1\sigma_{g_{\bar{\imath}1}}(j)}$, for any $j \in \{1,\ldots,n\}$. Since $\sigma_{g_{\bar{\imath}1}}$ is a permutation, the $\bar{\imath}$-th row is a permutation of the first row. By the same arguments, we can assert that every column of B is a permutation of the first column of B.

Let us now consider a real number y, and denote by $r(y)$ (respectively $s(y)$) the number of times y occurs in each row (respectively column) of B. Both $nr(y)$ and $ns(y)$ represent the number of times y appears in B. Since $nr(y) = ns(y)$, each row and column contains the same elements (counted with multiplicity). Hence, the statement of our lemma is proved. □

The following result is well known [26].

Lemma 3.4.10 Birkhoff–von Neumann decomposition *Let M be a $n\times n$ real matrix with non-negative entries, such that both the sum of the elements of each row and the sum of the elements of each column is equal to $\bar{c}$. Then for every $h \in \mathrm{Aut}(X)$ a non-negative real number $c(h)$ exists such that $\sum_{h\in\mathrm{Aut}(X)} c(h) = \bar{c}$ and $M = \sum_{h\in\mathrm{Aut}(X)} c(h)P(h)$, where $P(h)$ is the permutation matrix associated with h.*

We recall that the permutation matrix associated with the permutation $h : X \to X$ is the $n \times n$ real matrix $(p_{ij}(h))$ defined by setting $p_{ij}(h) = 1$ if $h(x_j) = x_i$ and $p_{ij}(h) = 0$ if $h(x_j) \neq x_i$. Equivalently, we can define the permutation matrix associated with the permutation $h : X \to X$ as the $n \times n$ real matrix $P(h)$ such that $P(h)e_j = e_{\sigma_h(j)}$ for every column vector $e_j := {}^t(0,\ldots,1,\ldots,0) \in \mathbb{R}^n$ (where 1 is in the j-th position).

We observe that $P\left(h^{-1}\right) = P(h)^{-1}$ and $P(h_1h_2) = P(h_1)P(h_2)$ for every $h, h_1, h_2 \in \mathrm{Aut}(X)$.

Remark 3.4.11. In general, the representation $M = \sum_{h\in\mathrm{Aut}(X)} c(h)P(h)$, stated in Lemma 3.4.10, is not unique. As an example, consider the set $X = \{1,2,3\}$ and the group $G = \mathrm{Aut}(X)$. Let $F\colon \mathbb{R}^X \to \mathbb{R}^X$ be the linear application that maps $\mathbb{1}_j$ to $\sum_{i\in X} \mathbb{1}_i$, for any $j \in X$. One could easily check that F is a linear GEO for $(\mathbb{R}^X, G)$. Indeed, we have that $F(\mathbb{1}_j h) = F(\mathbb{1}_j) = F(\mathbb{1}_j)h$ for any $j \in X$ and any

$h \in \mathrm{Aut}(X)$. The matrix B associated with F with respect to the basis $\{\mathbb{1}_j\}_j$ is:

$$B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

One could represent B at least in two different ways:

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

and

$$B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

We proceed in our proof of Theorem 13 by taking the linear maps $F^{\oplus}, F^{\ominus} : \mathbb{R}^X \to \mathbb{R}^X$ defined by setting $F^{\oplus}(\mathbb{1}_{x_j}) := \sum_{i=1}^{n} \max\{b_{ij}, 0\}\mathbb{1}_{x_i}$ and $F^{\ominus}(\mathbb{1}_{x_j}) := \sum_{i=1}^{n} \max\{-b_{ij}, 0\}\mathbb{1}_{x_i}$ for every index $j \in \{1, \ldots, n\}$. We can easily check that

1. $F^{\oplus}, F^{\ominus}$ are linear GEOs;
2. The matrices associated with $F^{\oplus}$ and $F^{\ominus}$ with respect to the basis $\{\mathbb{1}_{x_1}, \ldots, \mathbb{1}_{x_n}\}$ of $\mathbb{R}^X$ are $B^{\oplus} = \left(b_{ij}^{\oplus}\right) = (\max\{b_{ij}, 0\})$ and $B^{\ominus} = \left(b_{ij}^{\ominus}\right) = (\max\{-b_{ij}, 0\})$, respectively (in particular, $B^{\oplus}, B^{\ominus}$ are non-negative matrices);
3. $F = F^{\oplus} - F^{\ominus}$ and $B = B^{\oplus} - B^{\ominus}$;
4. Lemma 3.4.9 and the definitions of $B^{\oplus}$, $B^{\ominus}$ imply that two n-tuples of real numbers $\alpha^{\oplus} = (\alpha_1^{\oplus}, \ldots, \alpha_n^{\oplus})$, $\alpha^{\ominus} = (\alpha_1^{\ominus}, \ldots, \alpha_n^{\ominus})$ exist such that each row and each column of $B^{\oplus}$ can be obtained by permuting $\alpha^{\oplus}$, and each row and each column of $B^{\ominus}$ can be obtained by permuting $\alpha^{\ominus}$.

From Property (4) and Lemma 3.4.10 this result follows:

Corollary 3.4.12 *For every $h \in \mathrm{Aut}(X)$ two non-negative real numbers $c^{\oplus}(h), c^{\ominus}(h)$ exist, such that $F^{\oplus}(\varphi) = \sum_{h\in\mathrm{Aut}(X)} c^{\oplus}(h)\varphi h^{-1}$ and $F^{\ominus}(\varphi) = \sum_{h\in\mathrm{Aut}(X)} c^{\ominus}(h)\varphi h^{-1}$ for every $\varphi \in \mathbb{R}^X$.*

Proof. Let us start by considering the statement concerning $c^{\oplus}(h)$ and $F^{\oplus}(h)$. Without loss of generality, since $F^{\oplus}$ is linear, it will suffice to prove the existence of a suitable non-negative function $c^{\oplus}(h)$, such that $F^{\oplus}(\mathbb{1}_{x_j}) = \sum_{h\in\mathrm{Aut}(X)} c^{\oplus}(h)\mathbb{1}_{x_j}h^{-1}$, for any $j \in \{1, \ldots, n\}$. The column coordinate vector of the function $F^{\oplus}(\mathbb{1}_{x_j})$ relative to the basis $\{\mathbb{1}_{x_1}, \ldots, \mathbb{1}_{x_n}\}$ is $B^{\oplus}e_j$. Property (4) and Lemma 3.4.10 imply that for every $h \in \mathrm{Aut}(X)$ a non-negative real number $c^{\oplus}(h)$ exists, such that

$$B^{\oplus}e_j = \sum_{h\in\mathrm{Aut}(X)} c^{\oplus}(h)P(h)e_j = \sum_{h\in\mathrm{Aut}(X)} c^{\oplus}(h)e_{\sigma_h(j)}.$$

Since the column vector $e_{\sigma_h(j)}$ represents the column coordinate vector of the function $\mathbb{1}_{h(x_j)}$ relative to the basis $\{\mathbb{1}_{x_1}, \ldots, \mathbb{1}_{x_n}\}$, we can conclude that

$$F^{\oplus}(\mathbb{1}_{x_j}) = \sum_{h \in \mathrm{Aut}(X)} c^{\oplus}(h) \mathbb{1}_{h(x_j)} = \sum_{h \in \mathrm{Aut}(X)} c^{\oplus}(h) \mathbb{1}_{x_j} h^{-1}.$$

The proof of the statement concerning $c^{\ominus}$ and $F^{\ominus}$ is analogous. □

Remark 3.4.13. In general, the function $c\colon \mathrm{Aut}(X) \to \mathbb{R}$ associated with the Birkhoff–von Neumann decomposition does not induce a permutant measure, i.e., the function μ_c that takes each subset H of $\mathrm{Aut}(X)$ to the value $\mu_c(H) := \sum_{h \in H} c(h)$ is not a permutant measure. For example, let us consider the set $X = \{1, 2, 3, 4\}$ and the group S_4 of all permutations of X. Let us define a linear GEO $F : \mathbb{R}^X \to \mathbb{R}^X$ for $(\mathbb{R}^X, S_4)$ by setting $F(\mathbb{1}_j) = \sum_{i \in X} \mathbb{1}_i$, for every index j. After fixing the basis $\{\mathbb{1}_j\}_j$, the matrix B associated with F has the following form:

$$B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

As guaranteed by Lemma 3.4.10, B can be decomposed as follows:

$$B = P(\mathrm{id}_X) + P(\sigma) + P(\sigma^2) + P(\sigma^3),$$

where $\sigma = (1\ 2\ 3\ 4) \in S_4$, in cycle notation. Let $\langle \sigma \rangle$ be the cyclic group generated by σ. The function $c : \mathrm{Aut}(X) \to \mathbb{R}$ associated with the previous decomposition of B is defined as follows: $c(h) = 1$ if $h \in \langle \sigma \rangle$, otherwise $c(h) = 0$. Let us now consider the permutation $g = (1\ 2) \in S_4$, in cycle notation. Since $\sigma^2 = (1\ 3)(2\ 4)$, we have that:

$$c(g\sigma^2 g^{-1}) = c((1\ 2)(1\ 3)(2\ 4)(1\ 2)) = c((1\ 4)(2\ 3)) = 0.$$

Since $c(\sigma^2) = 1$, c is not invariant under the conjugation action of S_4, and hence μ_c is not a permutant measure.

Let us now go back to the proof of Theorem 13 and consider the functions $c^{\oplus}, c^{\ominus}$: $\mathrm{Aut}(X) \to \mathbb{R}$ introduced in Corollary 3.4.12. In order to define the permutant measure μ on $\mathrm{Aut}(X)$ we will need the next lemma.

Lemma 3.4.14 *If $g \in G$ then $B^{\oplus} P(g) = P(g) B^{\oplus}$.*

Proof. Let us consider a permutation $g \in G$. The function $R_{g^{-1}} \colon \mathbb{R}^X \to \mathbb{R}^X$, which maps φ to φg^{-1}, is a linear application. Furthermore, $R_{g^{-1}}(\mathbb{1}_{x_j}) = \mathbb{1}_{x_j} g^{-1} = \mathbb{1}_{g(x_j)}$ for every index j. Hence, the matrix N associated to $R_{g^{-1}}$ with respect to the basis $\{\mathbb{1}_{x_1}, \ldots, \mathbb{1}_{x_n}\}$ verifies the equality $N e_j = e_{\sigma_g(j)}$, so that $N = P(g)$ (we set $e_j := {}^t(0, \ldots, 1, \ldots, 0) \in \mathbb{R}^n$, where 1 is in the j-th position). Since $F^{\oplus}$ is a GEO, the equality $F^{\oplus} R_{g^{-1}} = R_{g^{-1}} F^{\oplus}$ holds. This immediately implies that $B^{\oplus} P(g) = P(g) B^{\oplus}$. □

An analogous lemma holds for the matrix $B^{\ominus}$.

Lemma 3.4.14 guarantees that $P(g)B^{\oplus}P(g)^{-1} = B^{\oplus}$ for every $g \in G$. From this equality and Lemma 3.4.10 it follows that

$$\begin{aligned} B^{\oplus} &= \overbrace{\frac{1}{|G|}B^{\oplus} + \ldots + \frac{1}{|G|}B^{\oplus}}^{|G| \text{ summands}} \qquad (3.11) \\ &= \frac{1}{|G|}\sum_{g\in G} P(g)B^{\oplus}P(g)^{-1} \\ &= \frac{1}{|G|}\sum_{g\in G} P(g)\left(\sum_{h\in \mathrm{Aut}(X)} c^{\oplus}(h)P(h)\right)P(g)^{-1} \\ &= \sum_{h\in \mathrm{Aut}(X)}\sum_{g\in G} \frac{c^{\oplus}(h)}{|G|}P(g)P(h)P(g)^{-1} \\ &= \sum_{h\in \mathrm{Aut}(X)} \frac{c^{\oplus}(h)}{|G|}\sum_{g\in G} P(ghg^{-1}). \end{aligned}$$

Therefore, for every index j we have that

$$\begin{aligned} B^{\oplus}e_j &= \sum_{h\in \mathrm{Aut}(X)} \frac{c^{\oplus}(h)}{|G|}\sum_{g\in G} P(ghg^{-1})e_j \qquad (3.12) \\ &= \sum_{h\in \mathrm{Aut}(X)} \frac{c^{\oplus}(h)}{|G|}\sum_{g\in G} e_{\sigma_{ghg^{-1}}(j)}. \end{aligned}$$

This means that

$$\begin{aligned} F^{\oplus}(\mathbb{1}_{x_j}) &= \sum_{h\in \mathrm{Aut}(X)} \frac{c^{\oplus}(h)}{|G|}\sum_{g\in G} \mathbb{1}_{ghg^{-1}(x_j)} \qquad (3.13) \\ &= \sum_{h\in \mathrm{Aut}(X)} \frac{c^{\oplus}(h)}{|G|}\sum_{g\in G} \mathbb{1}_{x_j}gh^{-1}g^{-1}. \end{aligned}$$

Since $F^{\oplus}$ is linear, it follows that

$$F^{\oplus}(\varphi) = \sum_{h\in \mathrm{Aut}(X)} \frac{c^{\oplus}(h)}{|G|}\sum_{g\in G} \varphi gh^{-1}g^{-1} \qquad (3.14)$$

for every $\varphi \in \mathbb{R}^X$.

We observe that the permutations $gh^{-1}g^{-1}$ in the previous summation are not guaranteed to be different from each other, for g varying in G and h varying in $\mathrm{Aut}(X)$.

For each $h \in \mathrm{Aut}(X)$, let us consider the orbit $\mathcal{O}(h)$ of h under the conjugation action of G on $\mathrm{Aut}(X)$, and set

$$\begin{aligned} \mu^{\oplus}(h) &:= \sum_{f\in \mathcal{O}(h)} \frac{c^{\oplus}(f)}{|\mathcal{O}(f)|} = \sum_{f\in \mathcal{O}(h)} \frac{c^{\oplus}(f)}{|\mathcal{O}(h)|} \\ \mu^{\ominus}(h) &:= \sum_{f\in \mathcal{O}(h)} \frac{c^{\ominus}(f)}{|\mathcal{O}(f)|} = \sum_{f\in \mathcal{O}(h)} \frac{c^{\ominus}(f)}{|\mathcal{O}(h)|}. \end{aligned}$$

In other words, we define the measures $\mu^{\oplus}(h), \mu^{\ominus}(h)$ of each permutation h as the averages of the functions $c^{\oplus}, c^{\ominus}$ along the orbit of h under the conjugation action of G. Let G_h be the stabilizer subgroup of G with respect to h, i.e., the subgroup of G containing the elements that fix h by conjugation. We recall that by conjugating h with respect to every element of G we obtain each element of the orbit $\mathcal{O}(h)$ exactly $|G_h|$ times, and the well-known relation $|G_h||\mathcal{O}(h)| = |G|$ (cf. [3]). Let us now set $\delta(f, h) = 1$ if f and h belong to the same orbit under the conjugation action of G, and $\delta(f, h) = 0$ otherwise.

We observe that the following properties hold for $f, h \in \mathrm{Aut}(X)$:

1. $G_{h^{-1}} = G_h$;
2. if $f \in \mathcal{O}(h)$ then G_f is isomorphic to G_h;
3. $f^{-1} \in \mathcal{O}(h^{-1}) \iff f \in \mathcal{O}(h) \iff h \in \mathcal{O}(f)$.

Therefore, equality (3.14) implies

$$\begin{aligned}
F^{\oplus}(\varphi) &= \sum_{h\in\mathrm{Aut}(X)} \frac{c^{\oplus}(h)}{|G|}|G_{h^{-1}}| \sum_{f^{-1}\in\mathcal{O}(h^{-1})} \varphi f^{-1} \qquad (3.15)\\
&= \sum_{h\in\mathrm{Aut}(X)} \frac{c^{\oplus}(h)}{|G|}|G_h| \sum_{f\in\mathcal{O}(h)} \varphi f^{-1}\\
&= \sum_{h\in\mathrm{Aut}(X)} \frac{c^{\oplus}(h)}{|G|}|G_h| \sum_{f\in\mathrm{Aut}(X)} \delta(f,h)\varphi f^{-1}\\
&= \sum_{f\in\mathrm{Aut}(X)} \left(\sum_{h\in\mathrm{Aut}(X)} \frac{c^{\oplus}(h)}{|G|}|G_h|\delta(f,h)\right) \varphi f^{-1}\\
&= \sum_{f\in\mathrm{Aut}(X)} \left(\sum_{h\in\mathrm{Aut}(X)} \frac{c^{\oplus}(h)}{|\mathcal{O}(h)|}\delta(f,h)\right) \varphi f^{-1}\\
&= \sum_{f\in\mathrm{Aut}(X)} \left(\sum_{h\in\mathcal{O}(f)} \frac{c^{\oplus}(h)}{|\mathcal{O}(h)|}\right) \varphi f^{-1}\\
&= \sum_{f\in\mathrm{Aut}(X)} \varphi f^{-1}\, \mu^{\oplus}(f).
\end{aligned}$$

The definition of $\mu^{\oplus}$ immediately implies that $\mu^{\oplus}(H) = \mu^{\oplus}(gHg^{-1})$ for every $g \in G$ and every subset H of $\mathrm{Aut}(X)$. In other words, $\mu^{\oplus}$ is a non-negative permutant measure with respect to G. Quite analogously, we can prove the equality $F^{\ominus}(\varphi) = \sum_{f\in\mathrm{Aut}(X)} \varphi f^{-1}\, \mu^{\ominus}(f)$, and that $\mu^{\ominus}$ is a non-negative permutant measure with respect to G. As a result, the function $\mu := \mu^{\oplus} - \mu^{\ominus}$ is a permutant measure and the equality $F(\varphi) = \sum_{f\in\mathrm{Aut}(X)} \varphi f^{-1}\, \mu(f)$ holds, since $F = F^{\oplus} - F^{\ominus}$.

It remains to prove that $\sum_{h\in\mathrm{Aut}(X)} |\mu(h)| = \max_{\varphi\in\mathbb{R}^X\setminus\{\mathbf{0}\}} \frac{\|F(\varphi)\|_\infty}{\|\varphi\|_\infty}$.

This statement is trivial if $F \equiv \mathbf{0}$, since in this case μ is the null measure. Hence we can assume that F is not the null map and B is not the null matrix. In order to proceed, we need the next statement.

Proposition 3.4.15 *If $f_1, f_2 \in \mathrm{Aut}(X)$ and an index $s \in \{1, \ldots, n\}$ exists, such that $f_1(x_s) = f_2(x_s)$ (i.e., $\sigma_{f_1}(s) = \sigma_{f_2}(s)$), then either $c^{\oplus}(f_1) = 0$, or $c^{\ominus}(f_2) = 0$, or both.*

Proof. By applying the equality $F^{\oplus}(\varphi) = \sum_{h\in\mathrm{Aut}(X)} c^{\oplus}(h)\varphi h^{-1}$ for $\varphi = \mathbb{1}_{x_1}$, we obtain that

$$\begin{aligned} b^{\oplus}_{\sigma_{f_1}(s)s} &= \left(\sum_{i=1}^{n} b^{\oplus}_{is}\mathbb{1}_{x_i}\right)(x_{\sigma_{f_1}(s)}) \\ &= F^{\oplus}(\mathbb{1}_{x_s})(x_{\sigma_{f_1}(s)}) \\ &= F^{\oplus}(\mathbb{1}_{x_s})(f_1(x_s)) \\ &= \sum_{h\in\mathrm{Aut}(X)} c^{\oplus}(h)\mathbb{1}_{x_s}h^{-1}(f_1(x_s)) \\ &\geq c^{\oplus}(f_1)\mathbb{1}_{x_s}f_1{}^{-1}(f_1(x_s)) \\ &= c^{\oplus}(f_1)\mathbb{1}_{x_s}(x_s) \\ &= c^{\oplus}(f_1). \end{aligned} \tag{3.16}$$

Analogously, the inequality $b^{\ominus}_{\sigma_{f_2}(s)s} \geq c^{\ominus}(f_2)$ holds. Therefore,

$c^{\oplus}(f_1) > 0 \implies b^{\oplus}_{\sigma_{f_1}(s)s} > 0 \implies b^{\ominus}_{\sigma_{f_1}(s)s} = 0 \implies b^{\ominus}_{\sigma_{f_2}(s)s} = 0 \implies c^{\ominus}(f_2) = 0$.

It follows that either $c^{\oplus}(f_1) = 0$, or $c^{\ominus}(f_2) = 0$, or both. □

Corollary 3.4.16 *For every $f \in \mathrm{Aut}(X)$ either $c^{\oplus}(f) = 0$, or $c^{\ominus}(f) = 0$, or both.*

Proof. Set $f_1 = f_2$ in Proposition 3.4.15. □

Let us now set $c := c^{\oplus} - c^{\ominus}$. Corollary 3.4.16 implies that $|c(h)| = c^{\oplus}(h) + c^{\ominus}(h)$ for every $h \in \mathrm{Aut}(X)$. The definitions of $\mu^{\oplus}$ and $\mu^{\ominus}$ immediately imply that $\sum_{f\in\mathcal{O}(h)} \mu^{\oplus}(f) = \sum_{f\in\mathcal{O}(h)} c^{\oplus}(f)$ and $\sum_{f\in\mathcal{O}(h)} \mu^{\ominus}(f) = \sum_{f\in\mathcal{O}(h)} c^{\ominus}(f)$ for each $h \in \mathrm{Aut}(X)$. It follows that $\sum_{f\in\mathcal{O}(h)} |\mu(f)| \leq \sum_{f\in\mathcal{O}(h)} \mu^{\oplus}(f) + \sum_{f\in\mathcal{O}(h)} \mu^{\ominus}(f) = \sum_{f\in\mathcal{O}(h)} c^{\oplus}(f) + \sum_{f\in\mathcal{O}(h)} c^{\ominus}(f) = \sum_{f\in\mathcal{O}(h)} |c(f)|$ for each $h \in \mathrm{Aut}(X)$, and hence

$$\sum_{h\in\mathrm{Aut}(X)} |\mu(h)| \leq \sum_{h\in\mathrm{Aut}(X)} |c(h)|.$$

By setting $\mathbb{1}_X := \sum_{j=1}^{n} \mathbb{1}_{x_j}$ and recalling Corollary 3.4.12, we obtain $F^{\oplus}(\mathbb{1}_X) = \left(\sum_{h\in\mathrm{Aut}(X)} c^{\oplus}(h)\right)\mathbb{1}_X$ and $F^{\ominus}(\mathbb{1}_X) = \left(\sum_{h\in\mathrm{Aut}(X)} c^{\ominus}(h)\right)\mathbb{1}_X$. Since any line in B is a permutation of the first row of B, we get $F^{\oplus}(\mathbb{1}_X) = \left(\sum_{j=1}^{n} b^{\oplus}_{1j}\right)\mathbb{1}_X$ and $F^{\ominus}(\mathbb{1}_X) = \left(\sum_{j=1}^{n} b^{\ominus}_{1j}\right)\mathbb{1}_X$. As a consequence, the equalities $\sum_{h\in\mathrm{Aut}(X)} c^{\oplus}(h) = \sum_{j=1}^{n} b^{\oplus}_{1j}$ and $\sum_{h\in\mathrm{Aut}(X)} c^{\ominus}(h) = \sum_{j=1}^{n} b^{\ominus}_{1j}$ hold, and therefore $\sum_{h\in\mathrm{Aut}(X)} |c(h)| = \sum_{h\in\mathrm{Aut}(X)} c^{\oplus}(h) + \sum_{h\in\mathrm{Aut}(X)} c^{\ominus}(h) = \sum_{j=1}^{n} b^{\oplus}_{1j} + \sum_{j=1}^{n} b^{\ominus}_{1j} = \sum_{j=1}^{n} |b_{1j}|$.

Let us now consider the function $\bar{\varphi} := \sum_{j=1}^{n} \mathrm{sgn}(b_{1j})\mathbb{1}_{x_j} \in \mathbb{R}^X \setminus \{\mathbf{0}\}$. By recalling that any line in B is a permutation of the first row of B, we have that $\sum_{j=1}^{n} |b_{1j}| =$

$|\sum_{j=1}^{n} b_{1j}\text{sgn}\,(b_{1j})\,| \geq |\sum_{j=1}^{n} b_{ij}\text{sgn}\,(b_{1j})\,|$ for every index i. It follows that

$$\begin{aligned}\|F(\bar{\varphi})\|_\infty &= \left\| \sum_{i=1}^{n} \left(\sum_{j=1}^{n} b_{ij}\text{sgn}\,(b_{1j}) \right) \mathbb{1}_{x_i} \right\|_\infty \\ &= \sum_{j=1}^{n} |b_{1j}| \\ &= \sum_{h\in\text{Aut}(X)} |c(h)| \\ &\geq \sum_{h\in\text{Aut}(X)} |\mu(h)|.\end{aligned}$$

Since $\|\bar{\varphi}\|_\infty = 1$, $\frac{\|F(\bar{\varphi})\|_\infty}{\|\bar{\varphi}\|_\infty} \geq \sum_{h\in\text{Aut}(X)} |\mu(h)|$.

For every function $\varphi \in \mathbb{R}^X$ we have that $F(\varphi) = F_\mu(\varphi) := \sum_{h\in\text{Aut}(X)} \varphi h^{-1}\, \mu(h)$. Hence, $\|F(\varphi)\|_\infty \leq \sum_{h\in\text{Aut}(X)} \|\varphi h^{-1}\|_\infty\ |\mu(h)| = \|\varphi\|_\infty \sum_{h\in\text{Aut}(X)} |\mu(h)|$. Therefore, $\frac{\|F(\varphi)\|_\infty}{\|\varphi\|_\infty} \leq \sum_{h\in\text{Aut}(X)} |\mu(h)|$ for every $\varphi \in \mathbb{R}^X \setminus \{\mathbf{0}\}$.

In conclusion, $\sum_{h\in\text{Aut}(X)} |\mu(h)| = \max_{\varphi\in\mathbb{R}^X\setminus\{\mathbf{0}\}} \frac{\|F(\varphi)\|_\infty}{\|\varphi\|_\infty}$. □

Example 3.4.17. The simplest non-trivial example concerning the statement of Theorem 13 can be described as follows. Let $X = \{1,2\}$ and $G = \text{Aut}(X) = \{\text{id}_X, (1\ 2)\}$. Let us consider the linear GEO $F : \mathbb{R}^X \to \mathbb{R}^X$ defined by setting $F(\mathbb{1}_1) := \mathbb{1}_1 - \mathbb{1}_2$ and $F(\mathbb{1}_2) := \mathbb{1}_2 - \mathbb{1}_1$. By defining $\mu(\text{id}_X) := 1$ and $\mu((1\ 2)) := -1$, we get that μ is a permutant measure with respect to G and $F(\varphi) = \sum_{h\in\text{Aut}(X)} \varphi h^{-1}\, \mu(h)$ for every $\varphi \in \mathbb{R}^X$. Furthermore, $\sum_{h\in\text{Aut}(X)} |\mu(h)| = 2 = \frac{\|F(\mathbb{1}_1-\mathbb{1}_2)\|_\infty}{\|\mathbb{1}_1-\mathbb{1}_2\|_\infty} = \max_{\varphi\in\mathbb{R}^X\setminus\{\mathbf{0}\}} \frac{\|F(\varphi)\|_\infty}{\|\varphi\|_\infty}$.

We now observe that the assumption that G transitively acts on X cannot be removed from Theorem 13.

Example 3.4.18. Let us consider the set $X = \{1,2\}$ and the group $G = \{\text{id}_X\} \subseteq \text{Aut}(X) = \{\text{id}_X, (1\ 2)\}$. Take the operator $F \colon \mathbb{R}^X \to \mathbb{R}^X$ defined by setting $F(\mathbb{1}_i) = \mathbb{1}_1$ for any $i \in X$. Although F is a linear GEO, there does not exist a permutant measure μ on $\text{Aut}(X)$, such that $F(\varphi) = F_\mu(\varphi) := \sum_{h\in\text{Aut}(X)} \varphi h^{-1}\, \mu(h)$ for every $\varphi \in \mathbb{R}^X$.

By contradiction, let us assume that such a permutant measure μ exists. Then,

$$\mathbb{1}_1 = F(\mathbb{1}_1) = \mathbb{1}_1\text{id}_X\mu(\text{id}_X) + \mathbb{1}_1(1\ 2)\mu((1\ 2)) = \mathbb{1}_1\mu(\text{id}_X) + \mathbb{1}_2\mu((1\ 2)).$$

Since $\{\mathbb{1}_1, \mathbb{1}_2\}$ is a basis for $\mathbb{R}^X$, the equalities $\mu(\text{id}_X) = 1$ and $\mu((1\ 2)) = 0$ must hold.

It follows that

$$F(\mathbb{1}_2) = \mathbb{1}_2\text{id}_X\mu(\text{id}_X) + \mathbb{1}_1(1\ 2)\mu((1\ 2)) = \mathbb{1}_2.$$

This contradicts the assumption that $F(\mathbb{1}_2) = \mathbb{1}_1$.

Example 3.4.19. Let us set $X = \{1, 2, 3, 4\}$ and $G = \mathrm{Aut}(X)$. Let F be a linear GEO with respect to G. Let $B = (b_{ij})$ be the matrix associated with F with respect to the basis $\{\mathbb{1}_{x_1}, \ldots, \mathbb{1}_{x_n}\}$. In the proof of Lemma 3.4.9 we have seen that $b_{ij} = b_{\sigma_g(i)\sigma_g(j)}$ for any $g \in G$. It follows that two values $\alpha, \beta \in \mathbb{R}$ exist, such that $b_{ij} = \alpha$ if $i = j$ and $b_{ij} = \beta$ if $i \neq j$. By using the cycle notation, let us set $\sigma = (1\ 2\ 3\ 4) \in \mathrm{Aut}(X)$ and $\langle\sigma\rangle = \{\mathrm{id}_X, \sigma, \sigma^2 = (1\ 3)(2\ 4), \sigma^3 = (1\ 4\ 3\ 2)\}$, i.e., the cyclic group generated by σ. We have that $B = \alpha P(\mathrm{id}_X) + \beta P(\sigma) + \beta P(\sigma^2) + \beta P(\sigma^3)$. Therefore, by setting $c(\mathrm{id}_X) := \alpha$, $c(\sigma) := c(\sigma^2) := c(\sigma^3) := \beta$, and $c(h) := 0$ for every $h \notin \langle\sigma\rangle$, we get $F(\varphi) = \sum_{h \in \mathrm{Aut}(X)} \varphi h^{-1}\, c(h)$.

However, the signed measure c is not a permutant measure, since the orbits under the conjugation action of G are the sets

$$\begin{array}{lcl}
\mathcal{O}(\mathrm{id}_X) & = & \{\mathrm{id}_X\} \\
\mathcal{O}(\sigma) & = & \{\sigma = (1\ 2\ 3\ 4), (1\ 2\ 4\ 3), (1\ 3\ 2\ 4), (1\ 3\ 4\ 2), (1\ 4\ 2\ 3), \sigma^3 = (1\ 4\ 3\ 2)\} \\
\mathcal{O}(\sigma^2) & = & \{(1\ 2)(3\ 4), \sigma^2 = (1\ 3)(2\ 4), (1\ 4)(2\ 3)\} \\
\mathcal{O}((1\ 2)) & = & \{(1\ 2), (1\ 3), (1\ 4), (2\ 3), (2\ 4), (3\ 4)\} \\
\mathcal{O}((1\ 2\ 3)) & = & \{(1\ 2\ 3), (1\ 2\ 4), (1\ 3\ 2), (1\ 3\ 4), (1\ 4\ 2), (1\ 4\ 3), (2\ 3\ 4), (2\ 4\ 3)\}
\end{array}$$

and, according to our definition, c is not constant on the orbits $\mathcal{O}(\sigma)$ and $\mathcal{O}(\sigma^2)$.

Following the proof of Theorem 13, we can get a permutant measure μ by computing an average on the orbits. In other words, we can set

$$\mu(h) := \begin{cases} c(\mathrm{id}_X) = \alpha, & \text{if } h = \mathrm{id}_X \\ \sum_{h \in \mathcal{O}(\sigma)} \frac{c(h)}{|O(\sigma)|} = \sum_{h \in \mathcal{O}(\sigma^2)} \frac{c(h)}{|O(\sigma^2)|} = \frac{\beta}{3}, & \text{if } h \in \mathcal{O}(\sigma) \cup \mathcal{O}(\sigma^2) \\ 0, & \text{otherwise.} \end{cases}$$

By making this choice, the equality $F(\varphi) = \sum_{g \in \mathrm{Aut}(X)} \varphi h^{-1}\, \mu(h)$ holds for every $\varphi \in \mathbb{R}^X$, i.e., F is the linear GEO associated with the permutant measure μ.

Proposition 3.4.4 and Theorem 13 immediately imply the following statement.

Theorem 14 *Assume that $G \subseteq \mathrm{Aut}(X)$ transitively acts on the finite set X and F is a map from $\mathbb{R}^X$ to $\mathbb{R}^X$. The map F is a linear group equivariant operator for $(\mathbb{R}^X, G)$ if and only if a permutant measure μ exists such that $F(\varphi) = \sum_{h \in \mathrm{Aut}(X)} \varphi h^{-1}\, \mu(h)$ for every $\varphi \in \mathbb{R}^X$.*

Our main result about the representation of linear GEOs can be adapted to GENEOs.

Theorem 15 *Assume that $G \subseteq \mathrm{Aut}(X)$ transitively acts on the finite set X and F is a map from $\mathbb{R}^X$ to $\mathbb{R}^X$. The map F is a linear group equivariant non-expansive operator for $(\mathbb{R}^X, G)$ if and only if a permutant measure μ exists such that $F(\varphi) = \sum_{h \in \mathrm{Aut}(X)} \varphi h^{-1}\, \mu(h)$ for every $\varphi \in \mathbb{R}^X$, and $\sum_{h \in \mathrm{Aut}(X)} |\mu(h)| \leq 1$.*

Proof. If F is a linear GENEO for $(\mathbb{R}^X, G)$, then Theorem 13 guarantees that in $\mathrm{PM}(G)$ a permutant measure μ exists, such that $F(\varphi) = \sum_{h \in \mathrm{Aut}(X)} \varphi h^{-1}\, \mu(h)$ for every $\varphi \in \mathbb{R}^X$, and $\sum_{h \in \mathrm{Aut}(X)} |\mu(h)| = \max_{\varphi \in \mathbb{R}^X \setminus \{\mathbf{0}\}} \frac{\|F(\varphi)\|_\infty}{\|\varphi\|_\infty}$. Since F is non-expansive, the inequality $\sum_{h \in \mathrm{Aut}(X)} |\mu(h)| \leq 1$ follows. This proves the first implication in our statement.

Let us now assume that a permutant measure μ exists such that

$$F(\varphi) = \sum_{h \in \mathrm{Aut}(X)} \varphi h^{-1} \ \mu(h)$$

for every $\varphi \in \mathbb{R}^X$, with $\sum_{h \in \mathrm{Aut}(X)} |\mu(h)| \leq 1$. Then Proposition 3.4.4 states that F is a linear group equivariant operator for $(\mathbb{R}^X, G)$. Moreover,

$$\begin{aligned} \|F(\varphi)\|_\infty &= \left\| \sum_{h \in \mathrm{Aut}(X)} \varphi h^{-1} \ \mu(h) \right\|_\infty \\ &\leq \sum_{h \in \mathrm{Aut}(X)} \left\| \varphi h^{-1} \right\|_\infty |\mu(h)| \\ &= \sum_{h \in \mathrm{Aut}(X)} \|\varphi\|_\infty |\mu(h)| \\ &= \|\varphi\|_\infty \left(\sum_{h \in \mathrm{Aut}(X)} |\mu(h)| \right) \\ &\leq \|\varphi\|_\infty. \end{aligned}$$

This proves that F is non-expansive, and concludes the proof of the second implication in our statement. $\square$

Chapter 4

Hilbert spaces and Riemannian manifolds of GENEOs

The focus on spaces of GENEOs stresses the need for a study of the topological and geometric structure of these spaces, in order to simplify their exploration and use. In particular, in this capter we show how we can endow a space of GENEOs with the structure of a Riemannian manifold, so making available the use of gradient descent methods for the minimization of cost functions. The Riemannian structure we propose is based on the comparison of the action of GENEOs on data, according to the model described in [31]. This comparison is made by taking into account the probability distribution on the set of admissible signals we are considering. As an application of this approach, we also describe a procedure to select a finite set of representative GENEOs in the considered manifold.

4.1 New mathematical setting

Let X be a set and $(V, \langle\cdot,\cdot\rangle_V)$ be a finite dimensional inner product space whose elements belong to the space $\mathbb{R}_b^X$ of bounded functions from X to $\mathbb{R}$. Then, V and $\mathbb{R}^n$ are isomorphic as vector spaces, where n is the dimension of V. Denote the norm arising from $\langle\cdot,\cdot\rangle_V$ by $\|\cdot\|_V$. For any $\varphi \in V$, let us consider its L^∞-norm $\|\varphi\|_\infty = \sup_{x\in X} |\varphi(x)|$. Note that $\|\cdot\|_V$ and $\|\cdot\|_\infty$ are equivalent, since V is finite dimensional. Hence, there exist two real numbers $\alpha, \beta > 0$ such that $\alpha\|\cdot\|_\infty \leq \|\cdot\|_V \leq \beta\|\cdot\|_\infty$. This implies that $\|\cdot\|_V$ and $\|\cdot\|_\infty$ induce the same topology τ_V on V.

Let us now choose a Borel measure λ on V and an integrable function $f : V \to \mathbb{R}$ such that $\int_V f d\lambda = 1$. Let us define a probability Borel measure μ on V by setting $\mu(A) = \int_A f d\lambda$ for any Borel set A in V. Finally, let us assume that the essential support of μ (i.e. the smallest closed subset C of V such that $f = 0$ μ-almost everywhere outside C) is a compact subspace Φ of V. We will call Φ the space of *admissible signals* in probability. Let us choose a subgroup G of $\mathrm{Aut}_\Phi(X)$. We assume that the following conditions are satisfied:

1. $\langle\cdot,\cdot\rangle_V$ is invariant under the action of G (i.e., $\langle\varphi_1 g, \varphi_2 g\rangle_V = \langle\varphi_1, \varphi_2\rangle_V$ for any $\varphi_1, \varphi_2 \in V$ and $g \in G$).

2. f, λ (and hence μ) are invariant under the action of G (i.e., if $g \in G$ then $f(\varphi g) = f(\varphi)$ for any $\varphi \in V$, and $\lambda(A) = \lambda(g(A))$ for every Borel set A in V).

Condition (1) immediately implies that $\|\cdot\|_V$ is invariant under the action of G.

Under the previous assumptions, we will say that V is a (finite dimensional) *probability inner product space with equivariance group* G.

4.2 The pseudo-distance on X induced by admissible signals

In this subsection, we show that the distance on Φ, induced by $\|\cdot\|_V$, can be pulled back to a pseudo-distance on X.

The following lemma prepares us to define a new pseudo-distance on X.

Lemma 4.2.1 *Let $x_1, x_2 \in X$. Then, the map $\xi_{x_1,x_2} : \Phi \to \mathbb{R}$ defined by $\xi_{x_1,x_2}(\varphi) = |\varphi(x_1) - \varphi(x_2)|$ for any $\varphi \in \Phi$, is continuous with respect to the topology induced by the norm $\|\cdot\|_V$ on Φ, and hence it is an integrable random variable with respect to λ.*

Proof. Let $\varphi_1, \varphi_2 \in \Phi$. Then,

$$\begin{aligned}
|\xi_{x_1,x_2}(\varphi_1) - \xi_{x_1,x_2}(\varphi_2)| &= \big|\,|\varphi_1(x_1) - \varphi_1(x_2)| - |\varphi_2(x_1) - \varphi_2(x_2)|\,\big| \\
&\le |(\varphi_1(x_1) - \varphi_1(x_2)) - (\varphi_2(x_1) - \varphi_2(x_2))| \\
&= |(\varphi_1(x_1) - \varphi_2(x_1)) - (\varphi_1(x_2) - \varphi_2(x_2))| \\
&\le |\varphi_1(x_1) - \varphi_2(x_1)| + |\varphi_1(x_2) - \varphi_2(x_2)| \\
&\le 2\|\varphi_1 - \varphi_2\|_\infty \le \frac{2}{\alpha}\|\varphi_1 - \varphi_2\|_V.
\end{aligned}$$

Therefore, ξ_{x_1,x_2} is continuous with respect to $\|\cdot\|_V$. □

Now, we define a pseudo-distance Δ_X on X as follows:

$$\Delta_X(x_1, x_2) = \int_\Phi |\varphi(x_1) - \varphi(x_2)| f(\varphi)\, d\lambda, \quad \forall\, x_1, x_2 \in X.$$

In plain words, the distance between two points x_1, x_2 is set to be the expected value of the function ξ_{x_1,x_2}. We observe that $\Delta_X \le D_X$, so that the topology τ_{D_X} induced by D_X is finer than the topology τ_{Δ_X} induced by Δ_X.

In the sequel, whenever not differently specified, we assume that X is equipped with the topology arising from Δ_X.

Proposition 4.2.2 *Each function $\varphi_0 \in \Phi$ is continuous with respect to Δ_X.*

Proof. After choosing an $\varepsilon > 0$, let us consider the ball $B_\varepsilon = B_\Phi\left(\varphi_0, \frac{\varepsilon\alpha}{4}\right) = \left\{\varphi \in \Phi : \|\varphi_0 - \varphi\|_V \le \frac{\varepsilon\alpha}{4}\right\}$. Since φ_0 is in the essential support of f, $\mu(B_\varepsilon)$ is positive. For any $\varphi \in B_\varepsilon$, we have that $\|\varphi_0 - \varphi\|_\infty \le \frac{1}{\alpha}\|\varphi_0 - \varphi\|_V \le \frac{\varepsilon}{4}$, and hence for every $x \in X$

$$\begin{aligned}|\varphi(x)-\varphi(x_0)| &= |(\varphi_0(x)-\varphi_0(x_0))+(\varphi(x)-\varphi_0(x))+(\varphi_0(x_0)-\varphi(x_0))|\\ &\geq |\varphi_0(x)-\varphi_0(x_0)|-|\varphi(x)-\varphi_0(x)|-|\varphi_0(x_0)-\varphi(x_0)|\\ &\geq |\varphi_0(x)-\varphi_0(x_0)|-2\|\varphi_0-\varphi\|_\infty\\ &\geq |\varphi_0(x)-\varphi_0(x_0)|-\frac{2}{\alpha}\|\varphi_0-\varphi\|_V\\ &\geq |\varphi_0(x)-\varphi_0(x_0)|-\varepsilon/2.\end{aligned}$$

This implies that

$$\begin{aligned}\Delta_X(x,x_0) &= \int_\Phi |\varphi(x)-\varphi(x_0)|f(\varphi)\,d\lambda\\ &\geq \int_{B_\varepsilon} |\varphi(x)-\varphi(x_0)|f(\varphi)\,d\lambda\\ &\geq \mu(B_\varepsilon)(|\varphi_0(x)-\varphi_0(x_0)|-\varepsilon/2).\end{aligned}$$

It follows that $|\varphi_0(x)-\varphi_0(x)| \leq \frac{\Delta_X(x,x_0)}{\mu(B_\varepsilon)}+\varepsilon/2$. Therefore, if $\Delta_X(x,x_0) \leq (\varepsilon/2)\,\mu(B_\varepsilon)$, then $|\varphi_0(x)-\varphi_0(x_0)| \leq \varepsilon$. □

We now recall that the initial topology τ_{in} on X with respect to Φ is the coarsest topology on X such that each function φ in Φ is continuous. From Proposition 4.2.2 and the definition of τ_{in}, it follows that $\tau_{\text{in}} \subseteq \tau_{\Delta_X}$. The definition of the pseudo-metric Δ_X immediately implies that $\Delta_X \leq D_X$, and hence $\tau_{\Delta_X} \subseteq \tau_{D_X}$. Since Φ is compact, we know that τ_{in} and τ_{D_X} are the same (see Theorem 2). Hence, $\tau_{D_X} = \tau_{\Delta_X} = \tau_{\text{in}}$.

The definition of the pseudo-metric Δ_X on X relies on Φ. Thus, properties on Φ naturally induce properties on X, as shown in the proof of the following statement.

Proposition 4.2.3 *Since Φ is totally bounded, (X,Δ_X) is totally bounded.*

Proof. Theorem 1 proves that X is totally bounded with respect to D_X. Since $\Delta_X \leq D_X$, X is also totally bounded with respect to Δ_X. □

Corollary 4.2.4 *If (X,Δ_X) is complete, then (X,Δ_X) is compact.*

Proof. It follows from Proposition 4.2.3, by recalling that in pseudo-metric spaces a set is compact if and only if it is complete and totally bounded [34]. □

4.3 The group G of Φ-preserving isometries

We know that the elements of $\mathrm{Aut}_\Phi(X)$ are isometries with respect to D_X. In the following, we show that each element of G is also an isometry with respect to Δ_X.

Proposition 4.3.1 *Each $g \in G$ is an isometry with respect to Δ_X.*

Proof. Let λ_g be the Borel measure on V defined by setting $\lambda_g(A) = \lambda(g(A))$ for any Borel set A in Φ. From the invariance of λ under the action of G, $\lambda_g = \lambda$. By applying a change of variable, the invariance of f under the action of each $g \in G$ implies that

$$\begin{aligned}\Delta_X(g(x_1), g(x_2)) &= \int_\Phi |\varphi g(x_1) - \varphi g(x_2)| f(\varphi)\, d\lambda \\ &= \int_\Phi |\varphi g(x_1) - \varphi g(x_2)| f(\varphi g)\, d\lambda_g \\ &= \int_\Phi |\varphi(x_1) - \varphi(x_2)| f(\varphi)\, d\lambda = \Delta_X(x_1, x_2)\end{aligned}$$

for any $x_1, x_2 \in X$. □

Now, we turn G into a pseudo-distance space by using the measure and the norm on Φ. To do this, we need the following lemma.

Lemma 4.3.2 *Let $g_1, g_2 \in G$. Then, the map $\xi_{g_1,g_2} : \Phi \to \mathbb{R}$ sending each φ to $\|\varphi g_1 - \varphi g_2\|_V$ is a continuous map with respect to $\|\cdot\|_V$, and hence an integrable random variable with respect to λ.*

Proof. Let $\varphi_1, \varphi_2 \in \Phi$. Since $\|\cdot\|_V$ is invariant with respect to G, we have that

$$\begin{aligned}|\xi_{g_1,g_2}(\varphi_1) - \xi_{g_1,g_2}(\varphi_2)| &= |\|\varphi_1 g_1 - \varphi_1 g_2\|_V - \|\varphi_2 g_1 - \varphi_2 g_2\|_V| \\ &\le \|(\varphi_1 g_1 - \varphi_1 g_2) - (\varphi_2 g_1 - \varphi_2 g_2)\|_V \\ &= \|(\varphi_1 g_1 - \varphi_2 g_1) - (\varphi_1 g_2 - \varphi_2 g_2)\|_V \\ &\le \|\varphi_1 g_1 - \varphi_2 g_1\|_V + \|\varphi_1 g_2 - \varphi_2 g_2)\|_V = 2\|\varphi_1 - \varphi_2\|_V.\end{aligned}$$

Therefore, ξ_{g_1,g_2} is continuous with respect to $\|\cdot\|_V$. □

Now, we can define the following pseudo-distance on G:

$$\Delta_G(g_1, g_2) = \int_\Phi \|\varphi g_1 - \varphi g_2\|_V f(\varphi)\, d\lambda, \tag{4.1}$$

for any $g_1, g_2 \in G$. We observe that

$$\begin{aligned}\Delta_G(g_1, g_2) &= \int_\Phi \|\varphi g_1 - \varphi g_2\|_V f(\varphi)\, d\lambda \\ &\le \int_\Phi \beta\|\varphi g_1 - \varphi g_2\|_\infty f(\varphi)\, d\lambda \\ &\le \int_\Phi \beta D_G(g_1, g_2) f(\varphi)\, d\lambda \\ &= \beta D_G(g_1, g_2) \int_\Phi f(\varphi)\, d\lambda \\ &= \beta D_G(g_1, g_2)\end{aligned}$$

for every $g_1, g_2 \in G$. The inequality $\Delta_G \le \beta D_G$, implies that the topology τ_{Δ_G} induced by Δ_G cannot be strictly finer than the topology τ_{D_G} induced by D_G.

In the sequel, whenever not differently specified, we consider G as a pseudo-distance space (and hence a topological space) with respect to Δ_G.

Lemma 4.3.3 *Let $f, g, h \in G$. We have that*

$$\Delta_G(f, g) = \Delta_G(fh, gh) = \Delta_G(hf, hg). \tag{4.2}$$

Proof. The first equality follows directly from the invariance of the norm $\|\cdot\|_V$ with respect to G. Now we show that $\Delta_G(f, g) = \Delta_G(hf, hg)$. By applying a change of variable, the invariance of λ and f under the action of G implies that

$$\begin{aligned}
\Delta_G(hf, hg) &= \int_\Phi \|\varphi hf - \varphi hg\|_V f(\varphi)\, d\lambda \\
&= \int_\Phi \|\varphi hf - \varphi hg\|_V f(\varphi h)\, d\lambda_h \\
&= \int_\Phi \|\varphi f - \varphi g\|_V f(\varphi)\, d\lambda \\
&= \Delta_G(f, g).
\end{aligned}$$

where λ_h is the Borel measure on V defined by setting $\lambda_h(A) = \lambda(h(A))$ for any Borel set A in Φ. □

Proposition 4.3.4 *The group G is a topological group. Further, the action of G on Φ by composition on the right is continuous.*

Proof. First, we show that G is a topological group. Let $\sigma : G \times G \to G$ and $\iota : G \to G$ be the composition and the inverse maps, respectively. We consider the product topology on $G \times G$. We must show that σ and ι are continuous. To show that σ is continuous, let $(g_1, g_2), (g_1', g_2') \in G \times G$. Using Lemma 1.2.15, we have that

$$\begin{aligned}
\Delta_G(g_1 g_2, g_1' g_2') &= \Delta_G(g_1, g_1' g_2' g_2^{-1}) \\
&\leq \Delta_G(g_1, g_1') + \Delta_G(g_1', g_1' g_2' g_2^{-1}) \\
&= \Delta_G(g_1, g_1') + \Delta_G(id_X, g_2' g_2^{-1}) \\
&= \Delta_G(g_1, g_1') + \Delta_G(g_2, g_2').
\end{aligned}$$

It follows that the composition map σ is continuous. Now, we show that ι is continuous. Consider $h_1, h_2 \in G$. We have that

$$\begin{aligned}
\Delta_G(h_1^{-1}, h_2^{-1}) &= \Delta_G(h_1^{-1} h_2, h_2^{-1} h_2) \\
&= \Delta_G(h_1^{-1} h_2, id_X) \\
&= \Delta_G(h_1^{-1} h_2, h_1^{-1} h_1) \\
&= \Delta_G(h_2, h_1) \\
&= \Delta_G(h_1, h_2).
\end{aligned}$$

This proves that ι is an isometry, and hence it is continuous.

Therefore, G is a topological group.

Let us now assume that $\rho : \Phi \times G \to \Phi$ is the natural action of G on Φ (i.e. $\rho(\varphi, g) = \varphi g$ for any $\varphi \in \Phi$ and $g \in G$). We have to prove that ρ is continuous, when $\Phi \times G$ is endowed with the product topology. Let $(\varphi_1, g_1), (\varphi_2, g_2) \in \Phi \times G$ and

$\varepsilon > 0$. Let us define B_ε as the ball $B_\Phi(\varphi_2, \frac{\varepsilon}{4})$ in Φ with respect to $\|\cdot\|_V$. Since φ_2 is in the essential support of f, $\mu(B_\varepsilon)$ is positive. We will show that if $\|\varphi_1 - \varphi_2\|_V \leq \frac{\varepsilon}{4}$ and $\Delta_G(g_1, g_2) \leq \frac{\varepsilon}{4}\mu(B_\varepsilon)$, then $\|\rho(\varphi_1, g_1) - \rho(\varphi_2, g_2)\| \leq \varepsilon$. Since $\|\varphi - \varphi_2\|_V \leq \frac{\varepsilon}{4}$ for every $\varphi \in B_\varepsilon$, we have that

$$\begin{aligned}
\Delta_G(g_1, g_2) &= \int_\Phi \|\varphi g_1 - \varphi g_2\|_V f(\varphi)\, d\lambda \\
&\geq \int_{B_\varepsilon} \|\varphi g_1 - \varphi g_2\|_V f(\varphi)\, d\lambda \\
&= \int_{B_\varepsilon} \|(\varphi g_1 - \varphi_2 g_1) + (\varphi_2 g_1 - \varphi_2 g_2) + (\varphi_2 g_2 - \varphi g_2)\|_V f(\varphi)\, d\lambda \\
&\geq \int_{B_\varepsilon} (\|\varphi_2 g_1 - \varphi_2 g_2\|_V - \|\varphi g_1 - \varphi_2 g_1\|_V - \|\varphi_2 g_2 - \varphi g_2\|_V)\, f(\varphi)\, d\lambda \\
&= \int_{B_\varepsilon} (\|\varphi_2 g_1 - \varphi_2 g_2\|_V - 2\|\varphi - \varphi_2\|_V)\, f(\varphi)\, d\lambda \\
&\geq \mu(B_\varepsilon) \left(\|\varphi_2 g_1 - \varphi_2 g_2\|_V - \frac{\varepsilon}{2} \right).
\end{aligned}$$

It follows that $\|\varphi_2 g_1 - \varphi_2 g_2\|_V \leq \frac{\Delta_G(g_1,g_2)}{\mu(B_\varepsilon)} + \frac{\varepsilon}{2}$. Therefore,

$$\begin{aligned}
\|\rho(\varphi_1, g_1) - \rho(\varphi_2, g_2)\|_V &= \|\varphi_1 g_1 - \varphi_2 g_2\|_V \\
&\leq \|\varphi_1 g_1 - \varphi_2 g_1\|_V + \|\varphi_2 g_1 - \varphi_2 g_2\|_V \\
&\leq \|\varphi_1 - \varphi_2\|_V + \frac{\Delta_G(g_1, g_2)}{\mu(B_\varepsilon)} + \frac{\varepsilon}{2} \\
&\leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Consequently, ρ is continuous. □

In order to study the compactness of G, we need the following result.

Proposition 4.3.5 *Since Φ is totally bounded, (G, Δ_G) is totally bounded.*

Proof. By Theorem 4, G is totally bounded with respect to D_G. Since $\Delta_G \leq \beta D_G$, G is also totally bounded with respect to Δ_G. □

Therefore, the following statement holds, by recalling that in pseudo-metric spaces a set is compact if and only if it is complete and totally bounded [34].

Corollary 4.3.6 *If (G, Δ_G) is complete, then (G, Δ_G) is compact.*

4.4 Hilbert space of GEOs

Let $V_1 \subseteq \mathbb{R}_b^{X_1}$ and $V_2 \subseteq \mathbb{R}_b^{X_2}$ be two finite dimensional probability inner product spaces with equivariance groups G_1 and G_2, respectively. Let $\langle \cdot, \cdot \rangle_{V_i}$, $\| \cdot \|_{V_i}$, λ_{V_i}, f_{V_i} be the inner product, the norm, the Borel measure and the probability density function considered on V_i, for $i = 1, 2$ (all of them are G_i-invariant). Moreover, we define a probability Borel measure μ_i on V_i by setting $\mu_i(A) = \int_A f_{V_i} d\lambda_{V_i}$ for any Borel set A in V_i, for $i = 1, 2$. We will assume that the essential supports Φ_1 and

Φ_2 of f_1 and f_2 are respectively compact. Moreover, we recall that since $\|\cdot\|_{V_i}$ and $\|\cdot\|_\infty$ are equivalent in V_i, there exist two real numbers $\beta_i, \alpha_i > 0$ such that

$$\alpha_i\|\cdot\|_\infty \leq \|\cdot\|_{V_i} \leq \beta_i\|\cdot\|_\infty \tag{4.3}$$

for $i = 1, 2$.

Let us consider the Lebesgue-Bochner space $L^2(\Phi_1, V_2)$ of all square integrable maps from Φ_1 to V_2. Explicitly:

$$L^2(\Phi_1, V_2) := \left\{ F \colon \Phi_1 \to V_2 : \left(\int_{\Phi_1} \|F(\varphi)\|^2_{V_2} f_{V_1}(\varphi)\, d\lambda_{V_1} \right)^{\frac{1}{2}} < \infty \right\}$$

Now, we define an inner product on $L^2(\Phi_1, V_2)$ as follows:

$$\langle F_1, F_2 \rangle = \int_{\Phi_1} \langle F_1(\varphi), F_2(\varphi) \rangle_{V_2} f_{V_1}(\varphi)\, d\lambda_{V_1} \quad \forall\, F_1, F_2 \in L^2(\Phi_1, V_2).$$

It is well known that the space $L^2(\Phi_1, V_2)$ is a Hilbert space.

Let us select a homomorphism $T \colon G_1 \to G_2$. For the rest of the chapter, we will consider only GEOs (or GENEOs) from Φ_1 to Φ_2 with respect to T. With an abuse of notation, in this chapter the term GEO will refer to a continuous function $F \colon (\Phi_1, \|\cdot\|_{V_1}) \to (\Phi_2, \|\cdot\|_{V_2})$ such that F is T-equivariant. The following remark ensures us that the two definitions of GEO we have given are equivalent:

Remark 4.4.1. Since $\|\cdot\|_{V_1}$ and $\|\cdot\|_\infty$ are equivalent in V_1, and $\|\cdot\|_{V_2}$ and $\|\cdot\|_\infty$ are equivalent in V_2, a GEO $F \colon (\Phi_1, \|\cdot\|_{V_1}) \to (\Phi_2, \|\cdot\|_{V_2})$ is continuous (and hence Borel measurable) also with respect to the L^∞-norm defined on Φ_1 and Φ_2.

We define the following norms on the space of GEOs from Φ_1 to Φ_2:

1. $|||F|||_\infty := \sup_{\varphi\in\Phi_1} \|F(\varphi)\|_\infty$
2. $|||F|||_{V_2} := \sup_{\varphi\in\Phi_1} \|F(\varphi)\|_{V_2}$
3. $|||F|||_{L_2} := \left(\int_{\Phi_1} \|F(\varphi)\|^2_{V_2} f_{V_1}(\varphi)\, d\lambda_{V_1} \right)^{\frac{1}{2}}$.

Lemma 4.4.2 *$|||\cdot|||_\infty$ and $|||\cdot|||_{V_2}$ induce the same topology on the space of GEOs from Φ_1 to Φ_2.*

Proof. We already know that for any φ in Φ_1:

$$\alpha_2\|F(\varphi)\|_\infty \leq \|F(\varphi)\|_{V_2} \leq \beta_2\|F(\varphi)\|_\infty.$$

Hence, we have that:

$$\alpha_2 \sup_{\varphi\in\Phi_1} \|F(\varphi)\|_\infty \leq \sup_{\varphi\in\Phi_1} \|F(\varphi)\|_{V_2} \leq \beta_2 \sup_{\varphi\in\Phi_1} \|F(\varphi)\|_\infty.$$

Therefore, $|||\cdot|||_\infty$ and $|||\cdot|||_{V_2}$ are equivalent norms, and hence they induce the same topology on the space of GEOs from Φ_1 to Φ_2. □

Lemma 4.4.3 *The topology induced by $|||\cdot|||_{V_2}$ on the space of GEOs from Φ_1 to Φ_2 is finer that the one induced by $|||\cdot|||_{L_2}$.*

Proof. Let us consider a GEO F from Φ_1 to Φ_2. We have that

$$\begin{aligned}
|||F|||_{L_2} &= \left(\int_{\Phi_1} \|F(\varphi)\|^2_{V_2} f_{V_1}(\varphi)\ d\lambda_{V_1}\right)^{\frac{1}{2}} \\
&\leq \left(\int_{\Phi_1} \left(\sup_{\varphi\in\Phi_1} \|F(\varphi)\|^2_{V_2}\right) f_{V_1}(\varphi)\ d\lambda_{V_1}\right)^{\frac{1}{2}} \\
&= |||F|||_{V_2} \left(\int_{\Phi_1} f_{V_1}(\varphi)\ d\lambda_{V_1}\right)^{\frac{1}{2}} \\
&= |||F|||_{V_2}.
\end{aligned}$$

The above inequality implies directly the statement of the lemma. □

In general, the space of GEOs from Φ_1 to Φ_2 is a subspace of $L^2(\Phi_1, V_2)$, but it is not a Hilbert space. An interesting question immediately arises: Which is a natural Hilbert space where the space of GEOs lives in? Before proceeding, we have to give some definitions and results. First we recall some notions on convergence.

Definition 4.4.4. Consider $F \in L^2(\Phi_1, V_2)$ and a sequence $(F_i)_{i\in\mathbb{N}}$ in $L^2(\Phi_1, V_2)$.

1. We say that $F_i \to F$ in $L^2(\Phi_1, V_2)$ if and only if $|||F_i - F|||_{L_2} \to 0$ as $i \to \infty$.
2. We say that $F_i \to F$ in measure if and only if for every $\varepsilon > 0$, $\mu_1(\{\varphi \in \Phi_1 \colon \|F_i(\varphi) - F(\varphi)\|_{V_2} \geq \varepsilon\}) \to 0$ as $i \to \infty$.
3. We say that F_i is Cauchy in measure if and only if for every $\varepsilon > 0$, $\mu_1(\{\varphi \in \Phi_1 \colon \|F_r(\varphi) - F_s(\varphi)\|_{V_2} \geq \varepsilon\}) \to 0$ as $r, s \to \infty$.

Proposition 4.4.5 *If $F_i \to F$ in $L^2(\Phi_1, V_2)$, then $F_i \to F$ in measure.*

Proof. Let us consider the set $E_{i,\varepsilon} = \{\varphi \in \Phi_1 \colon \|F_i(\varphi) - F(\varphi)\|_{V_2} \geq \varepsilon\}$. Then we have that:

$$\begin{aligned}
\left(\int_{\Phi_1} \|F_i(\varphi) - F(\varphi)\|^2_{V_2} f_{V_1}(\varphi)\ d\lambda_{V_1}\right)^{\frac{1}{2}} &\geq \left(\int_{E_{i,\varepsilon}} \|F_i(\varphi) - F(\varphi)\|^2_{V_2} f_{V_1}(\varphi)\ d\lambda_{V_1}\right)^{\frac{1}{2}} \\
&\geq \varepsilon \mu_1(E_{i,\varepsilon})^{\frac{1}{2}}.
\end{aligned}$$

Hence,

$$\mu_1(E_{i,\varepsilon})^{\frac{1}{2}} \leq \varepsilon^{-1} |||F_i - F|||_{L_2} \to 0, \text{ for } i \to \infty.$$

□

Theorem 16 *Suppose that the sequence $(F_i)_{i\in\mathbb{N}}$ in $L^2(\Phi_1, V_2)$ is Cauchy in measure. Then there exists a measurable function $F \colon \Phi_1 \to V_2$ such that $F_i \to F$ in measure, and there exists a subsequence $(F_{i_j})_{j\in\mathbb{N}}$ that converges to F almost everywhere. Moreover, if there is a measurable function $\bar{F} \colon \Phi_1 \to V_2$ such that $F_i \to \bar{F}$ in measure, then $\bar{F} = F$ almost everywhere.*

Proof. Let us choose a subsequence $(F_{i_j})_{j\in\mathbb{N}}$ of $(F_i)_{i\in\mathbb{N}}$ in $L^2(\Phi_1, V_2)$, such that if $E_j = \{\varphi \in \Phi_1 \colon \|F_{i_j}(\varphi) - F_{i_{j+1}}(\varphi)\|_{V_2} \geq 2^{-j}\}$, then $\mu_1(E_j) \leq 2^{-j}$.

Now, let us consider $\mathcal{E}_k = \bigcup_{j=k}^{\infty} E_j$. Then $\mu_1(\mathcal{E}_k) \leq \sum_{j=k}^{\infty} 2^{-j} = 2^{1-k}$. Moreover, if $\varphi \notin \mathcal{E}_k$, for $r \geq s \geq k$ we have

$$\|F_{i_r}(\varphi) - F_{i_s}(\varphi)\|_{V_2} \leq \sum_{t=s}^{r-1} \|F_{i_{t+1}}(\varphi) - F_{i_t}(\varphi)\|_{V_2} \leq \sum_{t=s}^{r-1} 2^{-t} = 2^{1-s}. \tag{4.4}$$

Take $\mathcal{E} = \bigcap_{k=1}^{\infty} \mathcal{E}_k = \limsup E_j$. By definition of $\mathcal{E}$, $\mu_1(\mathcal{E}) = 0$. We note that the sequence $(F_{i_j})_{j\in\mathbb{N}}$ is pointwise Cauchy on $\mathcal{E}^c$ (which is the complement of $\mathcal{E}$). Let us define a function $F \colon \Phi_1 \to V_2$ as follows: $F(\varphi)$ is the zero function on X_1 for any $\varphi \in \mathcal{E}$, and $F(\varphi) := \lim_{j\to\infty} F_{i_j}(\varphi)$ on $\mathcal{E}^c$ (such a limit exists since V_2 is complete). Then $F_{i_j} \to F$ almost everywhere and F is measurable. Also, (4.4) implies that $\|F_{i_j}(\varphi) - F(\varphi)\|_{V_2} \leq 2^{1-j}$ for $\varphi \notin \mathcal{E}_k$ and $j \geq k$. Since $\mu_1(\mathcal{E}_k) \to 0$ as $k \to \infty$, $F_{i_j} \to F$ in measure. Now, for every index i we can write

$$\{\varphi \in \Phi_1 \colon \|F_i(\varphi) - F(\varphi)\|_{V_2} \geq \varepsilon\} \subseteq \\ \left\{\varphi \in \Phi_1 \colon \|F_i(\varphi) - F_{i_j}(\varphi)\|_{V_2} \geq \frac{\varepsilon}{2}\right\} \cup \left\{\varphi \in \Phi_1 \colon \|F_{i_j}(\varphi) - F(\varphi)\|_{V_2} \geq \frac{\varepsilon}{2}\right\}.$$

Note that the sets on the right hand side both have small measure when i and j are large. Hence, $F_i \to F$ in measure. Likewise, if $F_i \to \bar{F}$ in measure, we have that

$$\left\{\varphi \in \Phi_1 \colon \|F(\varphi) - \bar{F}(\varphi)\|_{V_2} \geq \varepsilon\right\} \subseteq \\ \left\{\varphi \in \Phi_1 \colon \|F_i(\varphi) - F(\varphi)\|_{V_2} \geq \frac{\varepsilon}{2}\right\} \cup \left\{\varphi \in \Phi_1 \colon \|F_i(\varphi) - \bar{F}(\varphi)\|_{V_2} \geq \frac{\varepsilon}{2}\right\}$$

for all i, hence $\mu_1(\{\varphi \in \Phi_1 \colon \|F(\varphi) - \bar{F}(\varphi)\|_{V_2} \geq \varepsilon\}) = 0$ for all ε. Letting ε tend to zero, we conclude that $F = \bar{F}$ almost everywhere.

□

Corollary 4.4.6 *If $F_i \to F$ in $L^2(\Phi_1, V_2)$, there is a subsequence $(F_{i_j})_{j\in\mathbb{N}}$ such that $F_{i_j} \to F$ almost everywhere.*

Proof. Combining Proposition 4.4.5 and Theorem 16, we can immediately state the result. □

Furthermore, in order to consider a Hilbert space of operators we need a a weaker definition of GEO:

Definition 4.4.7. An operator F in $L^2(\Phi_1, V_2)$ is called **μ-GEO** if it is T-equivariant almost everywhere, i.e. for any $g \in G$ and almost any $\varphi \in \Phi$, $F(\varphi g) = F(\varphi)T(g)$.

Proposition 4.4.8 *The space of μ-GEOs from Φ_1 to Φ_2 is a Hilbert space.*

Proof. It will suffice to show that the space of μ-GEOs is a closed subspace of $L^2(\Phi_1, V_2)$. Assume that the sequence $(F_i)_{i\in\mathbb{N}}$ of μ-GEOs converges to F in $L^2(\Phi_1, V_2)$. It will suffice to prove that F is a μ-GEO. Since $(F_i)_{i\in\mathbb{N}}$ converges in $L^2(\Phi_1, V_2)$, by

Corollary 4.4.6 there is a subsequence $(F_{i_j})_{j\in\mathbb{N}}$ such that $(F_{i_j})_{j\in\mathbb{N}}$ converges to F almost everywhere. Since T-equivariance is a pointwise property, we have that

$$F(\varphi g) = \lim_{j\to\infty} F_{i_j}(\varphi g) = \lim_{j\to\infty} F_{i_j}(\varphi)T(g) = F(\varphi)T(g)$$

for any $g \in G$ and almost any $\varphi \in \Phi$. □

In this context, we will now consider a slightly different definition of GENEO. A GENEO from $(\Phi_1, \|\cdot\|_{V_1})$ to $(\Phi_2, \|\cdot\|_{V_2})$ is a GEO F from Φ_1 to Φ_2 such that

$$\|F(\varphi) - F(\varphi')\|_{V_2} \leq \|\varphi - \varphi'\|_{V_1}$$

for every $\varphi, \varphi' \in \Phi_1$.

Remark 4.4.9. Assume (F_i) is a sequence of GENEOs from $(\Phi_1, \|\cdot\|_{V_1})$ to $(\Phi_2, \|\cdot\|_{V_2})$ with respect to $T : G_1 \to G_2$. If (F_i) converges to $F\colon \Phi_1 \to \Phi_2$ with respect to the norm $|||\cdot|||_{V_2}$, then $F(\varphi) = \lim_{i\to\infty} F_i(\varphi)$ with respect to the norm $\|\cdot\|_{V_2}$ for every $\varphi \in \Phi$. Indeed, we have that $\|F(\varphi) - F_i(\varphi)\|_{V_2} \leq |||F - F_i|||_{V_2}$ for every index i.

Lemma 4.4.10 *If (F_i) is a sequence of GENEOs from $(\Phi_1, \|\cdot\|_{V_1})$ to $(\Phi_2, \|\cdot\|_{V_2})$ with respect to T converging to a function F from Φ_1 to Φ_2 with respect to $|||\cdot|||_{V_2}$, then F is a GENEO from $(\Phi_1, \|\cdot\|_{V_1})$ to $(\Phi_2, \|\cdot\|_{V_2})$.*

Proof. First we prove that F is non-expansive. For any pair φ_1, φ_2 in Φ_1, we have that:

$$\begin{aligned}
\|F(\varphi_1) - F(\varphi_2)\|_{V_2} &= \left\| \lim_{i\to\infty} F_i(\varphi_1) - \lim_{i\to\infty} F_i(\varphi_2) \right\|_{V_2} \\
&= \lim_{i\to\infty} \|F_i(\varphi_1) - F_i(\varphi_2)\|_{V_2} \\
&\leq \lim_{i\to\infty} \|\varphi_1 - \varphi_2\|_{V_1} \\
&= \|\varphi_1 - \varphi_2\|_{V_1}.
\end{aligned}$$

Now we show that F is equivariant. For any φ in Φ_1 and any g in G_1, we have that:

$$\begin{aligned}
F(\varphi g) &= \lim_{i\to\infty} F_i(\varphi g) \\
&= \lim_{i\to\infty} (F_i(\varphi)T(g)) \\
&= \left(\lim_{i\to\infty} F_i(\varphi) \right) T(g) \\
&= F(\varphi)T(g).
\end{aligned}$$

□

Theorem 17 *The space of GENEOs from $(\Phi_1, \|\cdot\|_{V_1})$ to $(\Phi_2, \|\cdot\|_{V_2})$ with respect to T is compact with respect to the norm $|||\cdot|||_{L_2}$.*

Proof. Let (F_i) be a sequence of GENEOs from $(\Phi_1, \|\cdot\|_{V_1})$ to $(\Phi_2, \|\cdot\|_{V_2})$. Because of Lemma 4.4.2, it will suffice to prove that there exists a subsequence (F_{i_j}) of

(F_i) that converges in the $|||\cdot|||_\infty$-topology. If F is a GENEO from $(\Phi_1, \|\cdot\|_{V_1})$ to $(\Phi_2, \|\cdot\|_{V_2})$, from (4.3) it follows:

$$\alpha_2 \|F(\varphi) - F(\varphi')\|_\infty \leq \|F(\varphi) - F(\varphi')\|_{V_2} \leq \|\varphi - \varphi'\|_{V_1} \leq \beta_1 \|\varphi - \varphi'\|_\infty$$

and hence

$$\|F(\varphi) - F(\varphi')\|_\infty \leq \frac{\beta_1}{\alpha_2} \|\varphi - \varphi'\|_\infty.$$

If we scale F by $\frac{\alpha_2}{\beta_1}$, we obtain a GENEO from $(\Phi_1, \|\cdot\|_\infty)$ to $(\Phi_2, \|\cdot\|_\infty)$. We can consider the associated sequence $(\frac{\alpha_2}{\beta_1} F_i)$ of GENEOs from $(\Phi_1, \|\cdot\|_\infty)$ to $(\Phi_2, \|\cdot\|_\infty)$. Since the space of GENEOs from $(\Phi_1, \|\cdot\|_\infty)$ to $(\Phi_2, \|\cdot\|_\infty)$ is compact with respect to $|||\cdot|||_\infty$, we can extract a subsequence $(\frac{\alpha_2}{\beta_1} F_{i_j})$ that converges with respect to $|||\cdot|||_\infty$ to a GENEO $\bar{F}$ from $(\Phi_1, \|\cdot\|_\infty)$ to $(\Phi_2, \|\cdot\|_\infty)$ (see [5]). This immediately implies that (F_{i_j}) converges to $\frac{\alpha_2}{\beta_1} \bar{F}$ with respect to $|||\cdot|||_{V_2}$. Lemma 4.4.10 ensures us that $\frac{\alpha_2}{\beta_1} \bar{F}$ is a GENEO from $(\Phi_1, \|\cdot\|_{V_1})$ to $(\Phi_2, \|\cdot\|_{V_2})$. Because of Lemma 4.4.3, we have that (F_{i_j}) converges to $\frac{\alpha_2}{\beta_1} \bar{F}$ with respect to $|||\cdot|||_{L_2}$. □

4.5 Submanifolds of GENEOs

In this subsection we discuss how it is possible to define a Riemannian structure on a manifold of GENEOs. We also briefly recall some basic definitions and results about Hilbert manifolds and Riemannian manifolds. For further details and the general theory, see [1, 36].

Since $L^2(\Phi, V_2)$ is a Hilbert space, it has a natural structure of Hilbert manifold with a single global chart given by the identity function on $L^2(\Phi, V_2)$. Each tangent space $T_p L^2(\Phi, V_2)$ at any point $p \in L^2(\Phi, V_2)$ is canonically isomorphic to $L^2(\Phi, V_2)$ itself. We can give a Riemannian structure on $L^2(\Phi, V_2)$ by defining a metric g as $g(v, w)(p) = \langle v, w \rangle$ for every $v, w \in T_p L^2(\Phi, V_2)$, where $\langle \cdot, \cdot \rangle$ is the inner product on $L^2(\Phi, V_2)$. Now, let M be a C^k-submanifold of $L^2(\Phi, V_2)$, such that each element of M is a GENEO. Hence, M naturally inherits a Riemannian structure from $L^2(\Phi, V_2)$.

4.5.1 An example

Let us consider X equal to the torus $S^1 \times S^1$, endowed with the metric induced by the usual embedding in $\mathbb{R}^4$ (implied by the usual embedding of S^1 in $\mathbb{R}^2$). We assume that $\chi : \mathbb{R}^+ \to \mathbb{R}$ is a continuous function whose support is contained in the interval $[0, \pi]$. As usual, parameterize S^1 by $\theta \in [0, 2\pi]$. We set $V_1 = V_2$ equal to the $2n$-dimensional vector space whose elements are the functions $\varphi : X \to \mathbb{R}$ that can be expressed as

$$\varphi(\alpha, \beta) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} a_i^j \chi \left(\sqrt{\left(\alpha - \frac{2\pi i}{n}\right)^2 + \left(\beta - \frac{2\pi j}{n}\right)^2} \right)$$

with $a_i^j \in \mathbb{R}$ for every $i, j \in \{0, \ldots, n-1\}$. Moreover, we take $\Phi_1 = \Phi_2$ equal to the compact subset of V_1 whose elements are the functions $\varphi \in V_1$ with $|a_i^j| \leq 1$ for every $i, j \in \{0, \ldots, n-1\}$.

Set G equal to the finite group whose elements are the isometries that preserve the set $Y = \left\{\left(\frac{2\pi i}{n}, \frac{2\pi j}{n}\right) : i, j \in \{0, \ldots, n-1\}\right\} \subseteq S^1 \times S^1$. It would be easy to check that every $g \in G$ is a Φ_1-operation.

In this example, our manifold of GENEOs is the one containing all GENEOs F defined by taking T equal to the identical homomorphism and setting

$$F(\varphi)(\alpha, \beta) := \frac{1}{\sqrt{m}} \left(\sum_{t=1}^{m} \sum_{r,s \in \{-1,1\}} u_t \varphi \left(\alpha + r\frac{k_t \pi}{n}, \beta + s\frac{k_t \pi}{n} \right) \right)$$

for any $\varphi \in \Phi_1$, with $u_t \in \mathbb{R}$ and $k_t \in \mathbb{N}$ for every $t \in \{1, \ldots, m\}$, such that $\sum_{t=1}^{m} u_t^2 = 1$.

Note that the manifold M is an $(m-1)$-dimensional sphere.

Chapter 5

Conclusions

The first contribution of this dissertation consists of giving a novel, formal and mathematical framework for machine learning, based on the study of metric, topological and geometric properties of operator spaces acting on function spaces. This approach is dual to the classical one: instead of focusing on data, our approach concentrates on suitable operators defined on the functions that represent the data. We focus on the study of the space of group equivariant non-expansive operators (GENEOs). From the mathematical point of view, equivariance means that our observer wants to respect some intrinsic symmetries of the set of admissible signals. In applications, choosing to work on a space of operators equivariant with respect to specific transformations allows us to inject in the system pre-existing knowledge. Indeed, the operators will be blind to the action of the group on the data, hence reducing the dimensionality of the space to be explored during optimisation. The choice of working with non-expansive operators is justified both by the possibility of proving the compactness of the spaces of GENEOs (under the assumption of compactness of the spaces of measurements), and by the fact that in practical applications we are usually interested in operators that compress the information we have as an input. The rationale of our approach is based on the assumption that the main interest in machine learning does not consist of the analysis and the approximation of data, but in the analysis and the approximation of the observers looking at the data. A simple example can make this idea clearer: if we consider images representing skin lesions, we are not mainly interested in the images per se but rather in approximating the judgement given by the physicians about such images.

Presenting our mathematical model, we have shown how the space of GENEOs is suitable for machine learning. By using pseudo-metrics, we defined a topology on the space of GENEOs which is induced by the one we defined on the function space of data. We built the necessary machinery to define maps between GENEOs whose groups of equivariance are different from each other. This definition is fundamental, because it allows one to compose operators hierarchically, in the same fashion as computational units are linked in an artificial neural network. Thereafter, by taking advantage of known and novel results in persistent homology, we proved compactness and convexity of the space of GENEOs under suitable hypotheses. These results guarantee that any operator can always be approximated by a finite number of operators belonging to the same space. It is important to stress the use of persistent homology in our model: the metric comparison of GENEOs is a key point in our approach and persistent homology allows for a fast comparison of functions, so allowing for a fast comparison of GENEOs. In order to enable us to populate and investigate better and better the space of GENEOs, we introduced various methods to build new classes of operators. Finally, in the application is often crucial to navigate in the space of operators following the gradient. For this reason we defined a Riemannian structure on a manifold of GENEOs, starting from studying a natural Hilbert space where GENEOs live in.

We would like to conclude by mentioning some interesting problems and new lines of research that naturally arise in our mathematical model. Given a dataset and an equivariance group, building equivariant operators could be, in general, a difficult task. We have already introduced some methods for constructing GENEOs in Chapter 3. Nevertheless, new methods should be developed, in order to get good approximations of the spaces of GENEOs. For example, we are able to define a class of GENEOs by means of permutant measures (see Section 3.3). Basically, this technique is grounded in using a symmetric weighted average to build new GENEOs. As a next step in this direction, we could generalize our approach through the use of other suitable symmetric functions. Furthermore, in many applications group actions may not be enough. As an example, we can take a dataset that contains digit images and the group of planar rotations as the invariance group. If we take an image of a "6" and rotate it through an angle of π, we obtain something that could be seen as an image of a "9" by an agent. But, for its own purposes, that agent might not admit that this could happen. This represents a significant limitation of our theory. In some sense, we need less algebraic structure in order to have a suitable mathematical model. Hence, a fundamental advancement in our research could be to generalize and adapt our framework to equivariance with respect to sets of transformations, instead of groups. A first step in this direction is made in [16]. We plan to devote further research to these issues.

Appendix A

Persistent Homology

In persistent homology, data are modelled in metric spaces. The usual first step is to filter the data so to obtain a family of nested topological spaces that captures the topological information at multiple scales. A common way to obtain a filtration is by considering the sublevel sets of a continuous function, hence the name *sublevel set persistence*. Let φ be a real-valued continuous function on a topological space X. Persistent homology represents the changes of the homology groups of the sublevel set $X_t = \varphi^{-1}((-\infty, t])$ varying t in $\mathbb{R}$. We can see the parameter t as an increasing time, whose changes produce the birth and the death of k-dimensional holes in the sublevel set X_t. We observe that the number of independent 0-dimensional holes of X_t equals the number of connected components of X_t minus one, 1-dimensional holes refer to tunnels and 2-dimensional holes to voids.

If $u, v \in \mathbb{R}$ and $u < v$, we can consider the inclusion i of X_u into X_v. If $\check{H}$ denotes the singular or Čech homology functor, such an inclusion induces a homomorphism $i_k : \check{H}_k(X_u) \to \check{H}_k(X_v)$ between the homology groups of X_u and X_v in degree k.

Definition A.0.1. The group $PH_k^{\varphi}(u, v) := i_k(\check{H}_k(X_u))$ is called the kth **persistent homology group** with respect to the function $\varphi : X \to \mathbb{R}$, computed at the point (u, v). The rank $r_k(\varphi)(u, v)$ of $PH_k^{\varphi}(u, v)$ is called the kth **persistent Betti numbers function** (PBN) with respect to the function $\varphi : X \to \mathbb{R}$, computed at the point (u, v).

Remark A.0.2. Let X and Y be two homeomorphic spaces and let $h : Y \to X$ be a homeomorphism. Then the persistent homology group with respect to the function $\varphi : X \to \mathbb{R}$ and the persistent homology group with respect to the function $\varphi h : Y \to \mathbb{R}$ are isomorphic at each point (u, v) in the domain. More precisely, the isomorphism is the one that maps each singular homology class $[c = \sum_{i=1}^{r} a_i \cdot \sigma_i] \in PH_k^{\varphi}(u, v)$ to the homology class $[c' = \sum_{i=1}^{r} a_i \cdot (h^{-1}\sigma_i)] \in PH_k^{\varphi h}(u, v)$, where each σ_i is a singular simplex involved in the representation of the cycle c. Therefore we can say that the persistent homology groups and the persistent Betti numbers functions are invariant under the action of the group $\mathrm{Homeo}(X)$ of all homeomorphisms from X to itself.

Persistent Betti numbers functions can be completely described by multisets called **persistence diagrams**. Another equivalent description is given by barcodes (cf. [13]).

Definition A.0.3. The kth persistence diagram is the multiset of all the pairs $p_j = (b_j, d_j)$, where b_j and d_j are the times of birth and death of the jth k-dimensional hole. When a hole never dies, we set its time of death equal to ∞.

The multiplicity $m(p_j)$ says how many holes share both the time of birth b_j and the time of death d_j. For technical reasons, the points (t, t) on the diagonal are added to each persistence diagram, each one with infinite multiplicity.

Each persistence diagram $\mathbb{D}$ can contain an infinite number of points. For every $q \in \Delta^* := \{(x, y) \in \mathbb{R}^2 \ : \ x < y\} \cup \{(x, \infty) : x \in \mathbb{R}\}$, the equality $m(q) = 0$ means that q does not belong to the persistence diagram $\mathbb{D}$. We define on $\bar{\Delta}^* := \{(x, y) \in \mathbb{R}^2 \ : \ x \leq y\} \cup \{(x, \infty) : x \in \mathbb{R}\}$ a pseudo-metric as follows

$$d^*((x, y), (x', y')) := \min \left\{ \max\{|x - x'|, |y - y'|\}, \max \left\{ \frac{y - x}{2}, \frac{y' - x'}{2} \right\} \right\} \tag{A.1}$$

by agreeing that $\infty - y = \infty$, $y - \infty = -\infty$ for $y \neq \infty$, $\infty - \infty = 0$, $\frac{\infty}{2} = \infty$, $|\pm\infty| = \infty$, $\min\{\infty, c\} = c$, $\max\{\infty, c\} = \infty$.

The pseudo-metric d^* between two points p and p' takes the smaller value between the cost of moving p to p' and the cost of moving p' and p onto $\Delta := \{(x, y) \in \mathbb{R}^2 \ : \ x = y\}$. Obviously, $d^*(p, p') = 0$ for every $p, p' \in \Delta$. If $p \in \Delta^+ := \{(x, y) \in \mathbb{R}^2 \ : \ x < y\}$ and $p' \in \Delta$, then $d^*(p, p')$ equals the distance, induced by the max-norm, between p and Δ. Points at infinity have a finite distance only to the other points at infinity, and their distance equals the Euclidean distance between abscissas.

We can compare persistence diagrams by means of the **bottleneck distance** (also called **matching distance**) δ_{match}.

Definition A.0.4. Let $\mathbb{D}$, $\mathbb{D}'$ be two persistence diagrams. We define the *bottleneck distance* δ_{match} between $\mathbb{D}$ and $\mathbb{D}'$ by setting

$$\delta_{\text{match}} \left(\mathbb{D}, \mathbb{D}'\right) := \inf_{\sigma} \sup_{p \in \mathbb{D}} d^* \left(p, \sigma\left(p\right)\right), \tag{A.2}$$

where σ varies in the set of all bijections from the multiset $\mathbb{D}$ to the multiset $\mathbb{D}'$.

For further information about persistence diagrams and the bottleneck distance, we refer the reader to [27, 22]. Each persistent Betti numbers function is associated with exactly one persistence diagram, and (if we use Čech homology) every persistence diagram is associated with exactly one persistent Betti numbers function. Then the metric δ_{match} induces a pseudo-metric d_{match} on the sets of the persistent Betti numbers functions [15].

Appendix B

Bochner integral

Now we have to recall some concept and basic results from the Bochner integral Theory (for further details, see [25, 43]).

Definition B.0.1. Let $(E, \|\cdot\|)$ be a Banach space and $(A, \mathcal{A}, \mu)$ a σ-finite (signed) measure space. A function $s\colon A \to E$ is called μ-simple function if it is of the form

$$s = \sum_{n=1}^{N} \mathbb{1}_{A_n} x_n,$$

where for any $1 \leq n \leq N$, $A_n \in \mathcal{A}$ and $x_n \in E$. The symbol $\mathbb{1}_{A_n}$ denotes the indicator function of the set A_n, that is $\mathbb{1}_{A_n}(a) = 1$ if $a \in A_n$ and $\mathbb{1}_{A_n}(a) = 0$ if $a \notin A_n$.

Definition B.0.2. A function $f\colon A \to E$ is strongly μ-measurable if there exists a sequence $(f_n)_{n\in\mathbb{N}}$ of μ-simple functions pointwise converging to f μ-almost everywhere.

Remark B.0.3. Let $(A, \mathcal{A}, \mu)$ be a σ-finite (signed) Borel measure space. By Pettis's measurability Theorem (see [25]), if $(E, \|\cdot\|)$ is separable, $f\colon A \to E$ is Borel measurable with respect to the norm topology of E implies that f is strongly μ-measurable. Hence, every continuous function $f\colon A \to E$ is strongly μ-measurable.

Definition B.0.4. A function $f\colon A \to E$ is μ-Bochner integrable if f is strongly μ-measurable and there exists a sequence of μ-simple functions $s_n\colon A \to E$ such that $\lim_{n\to\infty} \int_A \|s_n - f\| \, d\mu = 0$.

It follows that every μ-simple function is μ-Bochner integrable. Now, take $f = \sum_{n=1}^{N} \mathbb{1}_{A_n} x_n$. We set

$$\int_A f \, d\mu := \sum_{n=1}^{N} \mu(A_n) x_n.$$

One could easily check that this definition is independent of the representation of f. If f is μ-Bochner integrable, the limit

$$\int_A f \, d\mu := \lim_{n\to\infty} \int_A s_n \, d\mu$$

exists in E and is called the Bocher integral of f with respect to μ. It is routine to check that this definition is independent of the approximating sequence $(s_n)_{n\in\mathbb{N}}$.

Proposition B.0.5 *A strongly μ-measurable function $f\colon A \to E$ is μ-Bochner integrable if and only if $\int_A \|f\|\ d\mu < \infty$. Moreover we have that*

$$\left\|\int_A f\ d\mu\right\| \leq \int_A \|f\|\ d\mu.$$

Bibliography

[1] R. Abraham, J. E. Marsden, and T. Ratiu, *Manifolds, tensor analysis, and applications*, vol. 75, Springer Science & Business Media, 2012.

[2] F. Anselmi, L. Rosasco, and T. Poggio, *On invariance and selectivity in representation learning*, Information and Inference: A Journal of the IMA **5** (2016), no. 2, 134–158.

[3] M. Aschbacher, *Finite group theory*, 2 ed., Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2000.

[4] Y. Bengio, A. C. Courville, and P. Vincent, *Representation learning: A review and new perspectives*, IEEE Transactions on Pattern Analysis and Machine Intelligence **35** (2013), 1798–1828.

[5] M. G. Bergomi, P. Frosini, D. Giorgi, and N. Quercioli, *Towards a topological–geometrical theory of group equivariant non-expansive operators for data analysis and machine learning*, Nature Machine Intelligence **1** (2019), 423–433.

[6] S. Botteghi, M. Brasini, P. Frosini, and N. Quercioli, *On the finite representation of group equivariant operators via permutant measures*, arXiv preprint arXiv:2008.06340 (2020).

[7] R. Brüel Gabrielsson and G. Carlsson, *Exposition and interpretation of the topology of neural networks*, 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), 2019, pp. 1069–1076.

[8] P. Bubenik and J. A. Scott, *Categorification of persistent homology*, Discrete & Computational Geometry **51** (2014), no. 3, 600–627.

[9] F. Cagliari, B. Di Fabio, and C. Landi, *The natural pseudo-distance as a quotient pseudo-metric, and applications*, Forum Mathematicum **27** (01 May. 2015), no. 3, 1729 – 1742.

[10] F. Camporesi, P. Frosini, and N. Quercioli, *On a new method to build group equivariant operators by means of permutants*, 2nd International Cross-Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE) (Hamburg, Germany) (Andreas Holzinger, Peter Kieseberg, A Min Tjoa, and Edgar Weippl, eds.), Machine Learning and Knowledge Extraction, vol. LNCS-11015, Springer International Publishing, August 2018, Part 4: MAKE-Topology, pp. 265–272.

[11] G. Carlsson, *Topology and data*, Bull. Amer. Math. Soc. (N.S.) **46** (2009), no. 2, 255–308.

[12] G. Carlsson and R. Brüel Gabrielsson, *Topological approaches to deep learning*, Topological Data Analysis (Cham) (Nils A. Baas, Gunnar E. Carlsson, Gereon Quick, Markus Szymik, and Marius Thaule, eds.), Springer International Publishing, 2020, pp. 119–146.

[13] G. Carlsson and A. Zomorodian, *The theory of multidimensional persistence*, Discrete Comput. Geom. **42** (2009), no. 1, 71–93.

[14] P. Cascarano, P. Frosini, N. Quercioli, and A. Saki, *On the geometric and Riemannian structure of the spaces of group equivariant non-expansive operators*, arXiv preprint arXiv:2103.02543 (2021).

[15] A. Cerri, B. Di Fabio, M. Ferri, P. Frosini, and C. Landi, *Betti numbers in multidimensional persistent homology are stable functions*, Math. Methods Appl. Sci. **36** (2013), no. 12, 1543–1557.

[16] W. Chachólski, A. De Gregorio, N. Quercioli, and F. Tombari, *Landscapes of data sets and functoriality of persistent homology*, arXiv preprint arXiv:2002.05972 (2020).

[17] F. Chazal, D. Cohen-Steiner, M. Glisse, L. J. Guibas, and S. Y. Oudot, *Proximity of persistence modules and their diagrams*, Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry (New York, NY, USA), SCG '09, Association for Computing Machinery, 2009, p. 237–246.

[18] F. Chazal, D. Cohen-Steiner, L. J. Guibas, F. Mémoli, and S. Y. Oudot, *Gromov-hausdorff stable signatures for shapes using persistence*, Computer Graphics Forum, vol. 28, Wiley Online Library, 2009, pp. 1393–1403.

[19] S. Chowdhury, N. Clause, F. Mémoli, J. A. Sánchez, and Z. Wellner, *New families of stable simplicial filtration functors*, Topology and its Applications **279** (2020), 107254.

[20] T. S. Cohen and M. Welling, *Group equivariant convolutional networks*, Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48, ICML'16, JMLR.org, 2016, p. 2990–2999.

[21] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, *Stability of persistence diagrams*, vol. 37, 01 2007, pp. 263–271.

[22] ———, *Stability of persistence diagrams*, Discrete Comput. Geom. **37** (2007), no. 1, 103–120.

[23] T. K. Dey, F. Fan, and Y. Wang, *Computing topological persistence for simplicial maps*, Proceedings of the Thirtieth Annual Symposium on Computational Geometry (New York, NY, USA), SOCG'14, Association for Computing Machinery, 2014, p. 345–354.

[24] T. K. Dey, F. Mémoli, and Y. Wang, *Multiscale mapper: Topological summarization via codomain covers*, Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms (USA), SODA '16, Society for Industrial and Applied Mathematics, 2016, p. 997–1013.

[25] J. Diestel and J. J. Uhl, Jr., *Vector measures*, American Mathematical Society, Providence, R.I., 1977, With a foreword by B. J. Pettis, Mathematical Surveys, No. 15.

[26] F. Dufossé, K. Kaya, I. Panagiotas, and B. Uçar, *Further notes on Birkhoff–von Neumann decomposition of doubly stochastic matrices*, Linear Algebra and its Applications **554** (2018), 68 – 78.

[27] H. Edelsbrunner and J. Harer, *Persistent homology—a survey*, Surveys on discrete and computational geometry, Contemp. Math., vol. 453, Amer. Math. Soc., Providence, RI, 2008, pp. 257–282.

[28] H. Edelsbrunner and D. Morozov, *Persistent homology: theory and practice*, European Congress of Mathematics, Eur. Math. Soc., Zürich, 2013, pp. 31–50.

[29] R. Engelking, *General topology*, Sigma series in pure mathematics, Heldermann, 1989.

[30] P. Frosini, *Does intelligence imply contradiction?*, Cognitive Systems Research **10** (2009), no. 4, 297–315.

[31] ______, *Towards an observer-oriented theory of shape comparison*, Eurographics Workshop on 3D Object Retrieval (A. Ferreira, A. Giachetti, and D. Giorgi, eds.), The Eurographics Association, 2016, pp. 5–8.

[32] P. Frosini and G. Jabłoński, *Combining persistent homology and invariance groups for shape comparison*, Discrete Comput. Geom. **55** (2016), no. 2, 373–409.

[33] P. Frosini and N. Quercioli, *Some remarks on the algebraic properties of group invariant operators in persistent homology*, 1st International Cross-Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE) (Reggio, Italy) (Andreas Holzinger, Peter Kieseberg, A Min Tjoa, and Edgar Weippl, eds.), Machine Learning and Knowledge Extraction, vol. LNCS-10410, Springer International Publishing, August 2017, Part 1: MAKE Topology, pp. 14–24.

[34] S. A. Gaal, *Point set topology*, Pure and Applied Mathematics, Vol. XVI, Academic Press, New York-London, 1964.

[35] R. Ghrist, *Barcodes: the persistent topology of data*, Bull. Amer. Math. Soc. (N.S.) **45** (2008), no. 1, 61–75.

[36] W. P.A. Klingenberg, *Riemannian geometry*, De Gruyter, Berlin, Boston, 2011.

[37] Y. LeCun, Y. Bengio, et al., *Convolutional networks for images, speech, and time series*, The handbook of brain theory and neural networks **3361** (1995), no. 10.

[38] M. Lesnick, *The theory of the interleaving distance on multidimensional persistence modules*, Found. Comput. Math. **15** (2015), no. 3, 613–650.

[39] R. Li, S. Zhong, and C. Swartz, *An improvement of the Arzela–Ascoli theorem*, Topology and its Applications **159** (2012), 2058–2061.

[40] D. Marcos, M. Volpi, N. Komodakis, and D. Tuia, *Rotation equivariant vector field networks*, 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 5058–5067.

[41] J. Masci, D. Boscaini, M. Bronstein, and P. Vandergheynst, *Geodesic convolutional neural networks on riemannian manifolds*, Proceedings of the IEEE International Conference on Computer Vision Workshops, 2015, pp. 37–45.

[42] R. Gonzalez-Díaz N. Atienza and M. Soriano-Trigueros, *On the stability of persistent entropy and new summary functions for topological data analysis*, Pattern Recognition **107** (2020), 107509.

[43] C. Prévôt and M. Röckner, *A concise course on stochastic partial differential equations*, Lecture Notes in Mathematics, Springer Berlin Heidelberg, 2007.

[44] N. Quercioli, *Some new methods to build group equivariant non-expansive operators in TDA*, arXiv preprint arXiv:2010.08823 (2020).

[45] V. Robins, *Towards computing homology from finite approximations*, Topology proceedings, vol. 24, 1999, pp. 503–532.

[46] M. Scolamiero, W. Chachólski, A. Lundman, R. Ramanujam, and S. Öberg, *Multidimensional persistence and noise*, Found. Comput. Math. **17** (2017), no. 6, 1367–1406.