Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE
Corso di Laurea Magistrale in Matematica

FROM GROUP EQUIVARIANT TO PARTIAL GROUP EQUIVARIANT NON-EXPANSIVE OPERATORS

Tesi di Laurea in Topological Data Analysis

Relatore:
Chiar.mo Prof.
Patrizio Frosini

Presentata da:
Lucia Ferrari

Correlatori:
Dott. Nicola Quercioli
Dott.ssa Francesca Tombari

VI Sessione
Anno Accademico 2021/2022

Abstract

Incorporating symmetries as an inductive bias into neural network architectures, Geometric Deep Learning (GDL) and Topological Data Analysis (TDA) have given an improvement in the development of deep learning models. In recent years, a line of research has emerged forming a bridge between GDL and TDA: a topological-geometrical theory of Group Equivariant Non-Expansive Operators (GENEOs). In the theory of GENEOs the collection of all symmetries is represented by a group, but in some applications, the group axioms are not maintained since real-world data rarely follows strict mathematical symmetries due to noisy or incomplete data or to symmetry breaking features. The main aim of this thesis is to give a generalization of the results obtained for GENEOs to a new mathematical framework where the property of equivariance is maintained only for some transformations, encoding a partial equivariance with respect to the action of the group of all transformations. To this end, we introduce the concept of Partial Group Equivariant Non-Expansive Operator (P-GENEO), extending the results obtained for GENEOs to a more general set-up, where the sets of transformations are represented by subsets with a weaker structure than the one of group.

Contents

Introduzione

Nell'ultimo decennio, lo sviluppo sempre maggiore del deep learning e la complessità delle sue strutture hanno portato a un crescente interesse nella ricerca di strumenti per esplorare meglio questa frontiera. In particolare, diverse tecniche geometriche sono state incorporate nel deep learning, dando origine al nuovo campo del Geometric Deep Learning (GDL).
Questo approccio geometrico al deep learning viene sfruttato con un duplice scopo. Da un lato, la geometria fornisce un quadro matematico comune per studiare le architetture delle reti neurali. Dall'altro, è possibile incorporare un bias geometrico nel modello di deep learning, basato sulla conoscenza pregressa dell'insieme di dati. In questo secondo caso, l'obiettivo del GDL è lo studio delle *simmetrie* dell'insieme di dati, cioè delle caratteristiche che sono *invarianti* sotto determinate trasformazioni. Per codificare queste proprietà, lo schema generale della maggior parte delle architetture di deep learning è modellato tenendo conto di una certa *equivarianza di gruppo*. Se consideriamo un insieme di dati e un gruppo che codifica le sue simmetrie, cioè le trasformazioni che portano dati legittimi in dati legittimi (per esempio la rotazione o la traslazione di un'immagine), l'equivarianza di gruppo è la proprietà che garantisce che tali simmetrie siano mantenute dopo l'azione di un operatore sull'insieme di dati. In altre parole, un operatore è detto equivariante se l'azione del gruppo sull'insieme di dati influisce nello stesso modo sull'output.
Esiste una stretta connessione tra il Geometric Deep Learning e la Topological Data Analysis (TDA). Negli ultimi anni si è assistito a un fiorente sviluppo di entrambe le discipline e diverse comunità di ricercatori, come matematici, ingegneri e informatici, hanno iniziato a lavorare su problemi in queste aree. Partendo da un terreno comune geometrico, è stato possibile sviluppare un modello matematico per l'equivarianza di gruppo che collega GDL e TDA (si veda [3], [25], [8], [4] e [15]). L'obiettivo principale di questa teoria è lo

studio dei Group Equivariant Non-Expansive Operators (GENEOs) da un punto di vista topologico.
Nella teoria dei GENEO, l'insieme di tutte le simmetrie è rappresentato da un gruppo, ma in alcune applicazioni gli assiomi di gruppo non sono necessariamente validi, poiché i dati del mondo reale raramente seguono simmetrie matematiche rigorose a causa di dati rumorosi o incompleti o di caratteristiche che rompono le simmetrie. A titolo di esempio, possiamo prendere un insieme di dati che contiene immagini di cifre e il gruppo delle rotazioni come gruppo che agisce sui dati. Se prendiamo un'immagine della cifra '6'e la ruotiamo di un angolo di π, otteniamo un'immagine che l'utente interpreterebbe molto probabilmente come '9'. Allo stesso tempo, vogliamo essere in grado di ruotare la cifra '6' di piccoli angoli mantenendo il suo significato.

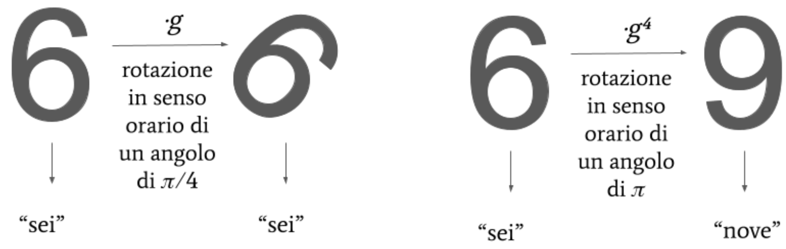

Figure 1: Esempio di rottura di simmetria. Applicando una rotazione g di $\pi/4$, la cifra '6' conserva il suo significato (a sinistra), mentre g^4 non è una trasformazione ammissibile, poiché si ottiene la cifra '9' (a destra).

È quindi auspicabile estendere la teoria dei GENEO, considerando gli insiemi di trasformazioni con una struttura più debole di quella dei gruppi. L'obiettivo principale di questa tesi è fornire una generalizzazione dei risultati ottenuti per i GENEO ad un nuovo quadro matematico in cui la proprietà di equivarianza è mantenuta solo per alcune trasformazioni, codificando un'equivarianza parziale rispetto all'azione del gruppo di tutte le trasformazioni. A tal fine, introdurremo il concetto di *Partial Group Equivariant Non-Expansive Operator (P-GENEO)*. Nella nostra impostazione matematica, consideriamo due insiemi di dati e un altro insieme contenente le trasformazioni ammissibili. Come insiemi di dati, prendiamo un insieme contenente i dati originali e un altro insieme che racchiude le loro variazioni ammissibili; quindi definiamo l'insieme delle trasformazioni ammissibili come un sottoinsieme di tutte le biiezioni che portano ogni elemento del primo insieme di dati nel secondo insieme. Dopo aver definito su ciascuno di questi insiemi una struttura topologica, definiremo i P-GENEO: triple di operatori, ciascuno dei quali agisce su uno degli spazi considerati e legati dalla proprietà di equivarianza. Illustreremo poi alcuni risultati sullo spazio dei P-GENEO, che estendono alcuni risultati già noti contenuti nella teoria dei GENEO. In sintesi, il concetto di P-GENEO viene introdotto per simulare il ruolo dell'osservatore mantenendo le equivarianze solo dove necessario. Ciò consente una maggiore libertà nella scelta delle simmetrie nei dati, a seconda dell'applicazione considerata.

Introduction

Over the past decade, the ever-increasing development of deep learning and the complexity of its structures have led to a growing interest in researching tools to better explore this frontier. In particular, several geometric techniques have been incorporated into deep learning, giving rise to the new field of Geometric Deep Learning (GDL).

This geometric approach to deep learning is exploited with a dual purpose. On one hand, geometry provides a common mathematical framework to study neural network architectures. On the other hand, a geometric bias in the deep learning model, based on prior knowledge of the data set, can be incorporated. In this second case, the focus of GDL is the study of *symmetries* of the data set, i.e., transformations that leave certain features unchanged or *invariant*. In order to encode these properties, the general blueprint of many deep learning architectures is modeled taking into account a certain *group equivariance*. If we consider a data set and a group encoding its symmetries, i.e., transformations that take admissible data to admissible data (for example rotation or translation of an image), the group equivariance is the property guaranteeing that such symmetries are maintained after the action of an operator on the data set. In other words, an operator is called equivariant if the action of the group on the data set affects the output in the same way.

There is a strict connection between Geometric Deep Learning and Topological Data Analysis (TDA). The past few years have witnessed a flourishing development of both these disciplines and different communities of researchers, as mathematicians, engineers, and computer scientists, have started working on problems in these areas. Building on a geometric common ground, it has been possible to develop a mathematical model for group equivariance connecting GDL and TDA (see [3], [25], [8], [4], and [15]). The main goal of this theory is the study of Group Equivariant Non-Expansive Operators

(GENEOs) through a topological point of view, in order to simulate the role of the observer maintaining the invariances.
In the theory of GENEOs, the collection of all symmetries is represented by a group, but in some applications, the group axioms do not necessarily hold since real-world data rarely follows strict mathematical symmetries due to noisy or incomplete data or to symmetry-breaking features. As an example, we can take a data set that contains images of digits and the group of rotations as the group acting on our data. If we take an image of the digit '6' and rotate it by an angle of π, we obtain an image that the user would most likely interpret as '9'. At the same time, we want to be able to rotate the digit '6' by small angles while preserving its meaning.

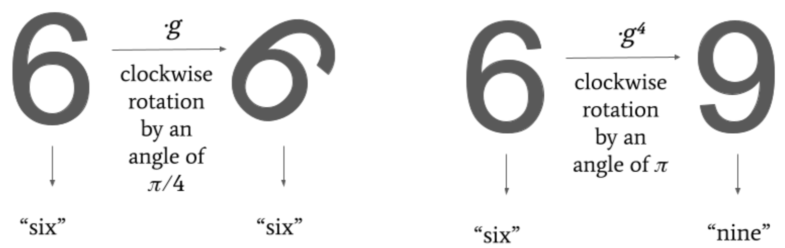

FIGURE 2: Example of a breaking symmetry feature. Applying a rotation g of $\pi/4$, the digit '6' preserves its meaning (left), while g^4 is not an admissible transformation, since we obtain the digit '9' (right).

It is then desirable to extend the theory of GENEOs, by considering the sets of transformations with a weaker structure than the one of group. The main aim of this thesis is to give a generalization of the results obtained for GENEOs to a new mathematical framework where the property of equivariance is maintained only for some transformations, encoding a partial equivariance with respect to the action of the group of all transformations. To this end, we will introduce the concept of *Partial Group Equivariant Non-Expansive Operator (P-GENEO)*.
In our mathematical setting, we consider two data sets paired with sets containing the admissible transformations. As data sets, we take a set containing the original data and another set that encloses their admissible variations;

then we define the set of admissible transformations as a subset of all the bijections that change every elementof the first data set into an element of the second set, by right composition. After defining on each of these sets a topological structure, we will define P-GENEOs: triples of operators, each acting on one of the considered spaces and linked by the equivariance property. We will then illustrate some results on the space of P-GENEOs, which extend some already known findings contained in the theory of GENEOs.
In summary, the concept of P-GENEO is introduced to simulate the role of the observer by maintaining equivariances only where necessary. This allows more freedom in the choice of symmetries in the data, depending on the application considered.

An overview: Deep Learning, Geometric Deep Learning and Topological Data Analysis

Deep Learning The essence of deep learning is to discover intricate structures in large data sets, by identifying different layers of representation. The algorithms used are implemented as a *backpropagation* algorithm, meaning that the machine learns how to change its internal parameters in each layer by propagating the parameter values taken at higher-level layers. In particular, deep-learning methods allow to amplify the relevant aspects of the input, deleting irrelevant variations. For example, in an image there are many features that can be represented in different layers: the edges typically are described in lower layers, while faces or letters are typically described in higher layers. *Artificial neural networks* are a realization of deep multi-layers hierarchies inspired by the processes of signals in the human brain.
In the real world, deep-learning algorithms have many applications, such as computer vision (see [20]), speech processing (see [21]) and medical diagnosis (e.g., to detect carcinogens or for Alzheimer's disease prediction, see [2]), among many others. For further and more detailed information on deep-learning algorithms, we refer to [22].

Geometric Deep Learning Over the past decade, one of the main focuses of deep learning research has been the study of signals such as images, video and speech in a Euclidean domain. However, in recent years there has been a need to deal with more complicated data set structures that reside in high-dimensional spaces or in non-Euclidean domains. Geometric Deep Learning (GDL) is a branch of deep learning that includes a geometric approach in the study of these structures. In particular, there are two dual relevant pathways in the field of GDL:

1. the study of the most successful neural networks architectures, such as CNNs, RNNs, GNNs, and Transformers in a common mathematical framework;

2. incorporate a geometric bias, based on prior knowledge of the data set, into deep learning models.

The theory of GENEOs extends the geometric approach expressed in 2) from data to observers.
There are many possible data set structures that can be used in different applications: *graphs* are commonly used to represent data that are related to each other and interact, such as social networks, molecular structures, and road networks; *point clouds* are used to represent data in 3D space; *manifolds* are a general way to describe non-Euclidean spaces such as spheres or hyperbolic spaces, used for example in computer vision.
The non-Euclidean nature of data implies that there are no familiar properties such as global parametrization, common system of coordinates or vector space structure. Consequently, basic operations such as linear combination or convolution, which are frequently used in deep-learning algorithms on Euclidean data, are not well-defined in non-Euclidean domains. Moreover, learning functions in high dimensions is a hard estimation problem. In fact, the number of parameters in a deep neural network grows exponentially with the dimensionality of the input space, making it difficult to train the network efficiently. In these structures, a geometric approach can allow us to take advantage of the *symmetries*, i.e., transformations that leave certain features of the data set unchanged or *invariant.*, in order to remedy the problem of working in these environments.
In most of the settings considered in GDL, the machine learning system operates on data sets composed of *signals* (functions) on some domain X. The set of symmetries is often represented by a *group*, since in many real-world applications the group axioms are maintained. In the case of non-Euclidean data, GDL methods use *group convolutions* or *equivariant neural networks* to take advantage of the symmetries of the data and improve the performance. Group convolutions apply a convolutional operation to a group of

transformations that preserve the structure of the data, such as rotations or translations. Equivariant neural networks, on the other hand, are designed to be invariant to certain types of transformations while maintaining equivariance to others. In other words, equivariance refers to the property of an operator to maintain the relationship between inputs and outputs even when the input undergoes some transformation or perturbation. An operator is called *invariant* with respect to a group if the action of the group on the input does not change the output, and it is called *equivariant* with respect to a group if the group action on the input affects the output in the same way. More formally, we can consider a data set of $\mathbb{R}^m$-valued signals on a domain X, $\Phi := \{\varphi\colon X \to \mathbb{R}^m\}$ and a group G that encloses its symmetries.
An operator F defined on Φ is G-*invariant* if

$$F(\varphi g) = F(\varphi),$$

i.e., if its output is unaffected by the group action on the input (e.g., Figure 3).

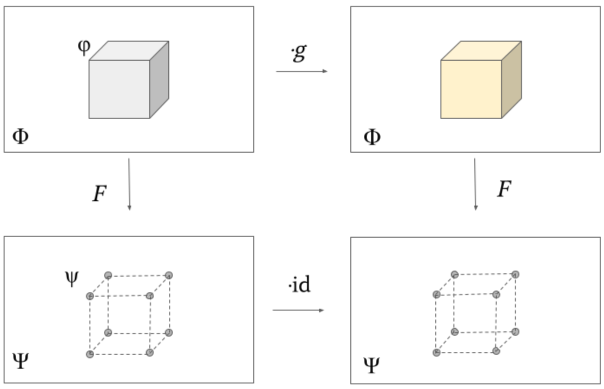

FIGURE 3: Example of *invariance*. Given an image of a cube (top left), color variation does not affect the result obtained by an edge detection operator (F).

We can consider another set of $\mathbb{R}^k$-valued functions on a domain Y, $\Psi := \{\psi\colon Y \to \mathbb{R}^k\}$ and a group H representing its symmetries. Considering a homomorphism $T\colon G \to H$, an operator $F\colon \Phi \to \Psi$ is called T-*equivariant* if

$$F(\varphi g) = F(\varphi)T(g),$$

i.e., the group action on the input affects the output in the same way (e.g., Figure 4). We can observe that T-equivariance becomes invariance when T

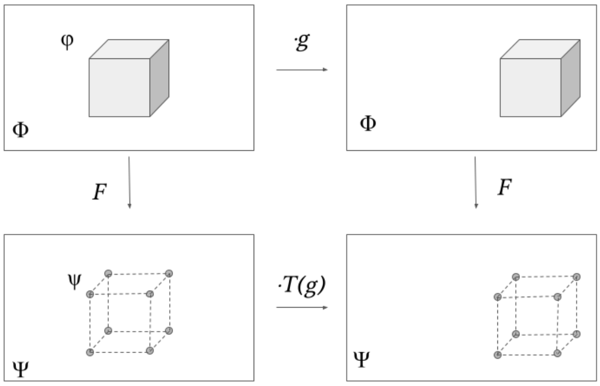

FIGURE 4: Example of *translation equivariance*. Given an image of a cube (top left), translating it by the element g (top right) and then computing the feature map F, we obtain an image (bottom right) that is equivalent to the image obtained by first computing the feature map (bottom left) and then translating the image (bottom right).

is the trivial homomorphism, whose image is the trivial group. For example, a 3D object recognition network may be equivariant to rotations and translations, but invariant to changes in scale or illumination.
By exploiting the symmetries of the data in this way, GDL algorithms can achieve state-of-the-art performance on a range of tasks, from graph classification to 3D point cloud segmentation. For more detailed information see [7], [6], [24], [31], [14].

Topological Data Analysis An area of research strictly connected with GDL is the study of data sets using topological techniques, called Topological Data Analysis (TDA). TDA provides mathematical methods to encode the complex geometry of data sets into a compact and easy-to-handle representation. There are some key points to consider when applying topology to data analysis.
A common goal of data analysis is to infer qualitative information from data, and to understand its organization both from a global and a local perspec-

tive. In this line topology offers the possibility of analyzing the connectivity of a space and, more generally, its structural features, such as clusters, holes, and voids, using methods such as *homology* and *persistent homology.*
Persistent homology (PH) is a mathematical tool that captures topological information at multiple scales. The first step is to filter the data set, commonly through the sublevels of a continuous function. In this way, one obtains a family of nested topological spaces that captures the topological information at different levels. In other words, we consider pairs composed by a topological space and a continuous function defined on it, that induces the sublevel set filtration. In general, PH allows us to study the birth and the death of k-dimensional holes when we move in the sublevel sets of the filtration. The topological persistence can informally be thought of as the evaluation of the importance of topological features with respect to their resistance to noise.
Another useful feature of topology is that it does not depend on the chosen coordinates, but studies the intrinsic geometric properties of objects. In TDA, the word "*summaries*" is referred to the topological features or characteristics of a data set that are obtained through TDA methods such as persistent homology. The final key point in the application of topology in data analysis is that summaries provide a more robust and informative description of the underlying structure of the data set compared to selecting individual parameters. In fact, summaries represent the topological features of a data set in a way that is stable under different choices of parameters. For further and more detailed information on TDA we refer to [10] , [16], [13], and [12].

Relation between GDL and TDA With the increasing development and use of deep learning techniques, the architecture of neural networks is becoming more and more varied and intricate and data sets are often residing on non-Euclidean domains, such as manifold, graphs or grids. GDL and TDA share the common intent of using geometry and topology to study complex data. TDA allows the addition of topological tools in the geometric approach to deep learning.
Specifically, as in GDL, we can apply the topological techniques in two lines of research, with a dual purpose: on one hand, TDA can be applied to deep learning, performing a topological analysis of neural network architectures and opening the way to alternative kinds of abstract representations. For ex-

ample, a topological analysis can be performed to study how the information is processed through the different stages of a convolutional neural network (CNN). Suppose we want to analyze a CNN trained for image classification. In this case, the objective of topological analysis could be to identify the main geometrical or topological features of the different structures within the network, such as connection patterns between neurons, the distribution of link weights, etc (e.g. [17]).
On the other hand, we can apply deep learning to TDA, using prior topological knowledge on the data set structure in order to feed the neural network with a simplified input. An example is the classification of unstructured data, such as text or biomolecules. In this case, topological analysis can be used to generate topological representations of text or biomolecules, which can then be used as input for deep learning models (e.g. [9]).

A bridge between GDL and TDA: GENEOs

In recent years, a line of research has emerged forming a bridge between GDL and TDA: a topological-geometrical theory of Group Equivariant Non-Expansive Operators (GENEOs).
This study operates within a general mathematical framework where any agent capable of acting on a certain set can be formally described as a collection of operators acting on the data. The main goal of this model is to simulate the role of the observer maintaining the wanted invariances/equivariances. In the setting considered, we will assume that our data are not studied directly, but through the action of agents that measure and transform them.
As in many other applications, we represent data by $\mathbb{R}^m$-valued continuous functions defined on some topological space. For the sake of simplicity in this work we will focus on real-valued functions. Hence, we will consider data as points of a space of functions (*signals* or *measurements*) defined on some domain X, endowed with a certain topological structure. For example, take as domain a two-dimensional $n \times n$ grid $X = \mathbb{Z}_n \times \mathbb{Z}_n$, an image in a grey scale φ (i.e. a signal $\varphi\colon X \to \mathbb{R}$) and a transformation F that subsamples the image by coarsening the underlying grid. In our model, we want to preserve topological features and symmetries of the signal after the application of some transformation. The reason for this choice in the managing of data is that we are more interested in the space of the observers transforming the data, rather than the data themselves. This way of looking at data is often

used in GDL, for example in convolution neural networks.
The theory that we are illustrating is the study of group equivariant non-expansive operators (GENEOs) through the use of topological data analysis. We have already described the importance of equivariance in GDL to be able to maintain the symmetries of the data. To be more specific, equivariance means that the observer wants to respect some intrinsic symmetries of the set of admissible signals. In application, working on a space of equivariant operators allows us to inject into the system pre-existing knowledge. This is a remedy against the risk of high computational complexity since it gives us the possibility to reduce the dimension while maintaining the discrimination purposes. On the other hand, the non-expansiveness of an operator is the property of maintaining or decreasing distances and it implies also the continuity of the operators. Formally, a function f from a pseudo-metric space (P, d_p) to a pseudo-metric space (Q, d_q) is called *non-expansive* if

$$d_q(f(x), f(y)) \leq d_p(x, y),$$

for every $x, y \in P$. This choice of working with non-expansive operators guarantees that, if the space of measurements is compact, the space of GENEOs is also compact. From the point of view of applications, non-expansivity models the need of compressing the information given to the operator as input. Moreover, the space of GENEOs is convex if the space of measurements is convex, allowing us to create a GENEO starting from a finite set of preexisting ones. The theory of GENEOs presents also other methods to build classes of GENEOs combining finite or infinite sets of known GENEOs in order to produce new ones.
We said before that the purpose of this theory is to simulate the role of the observer; indeed we focus not on the data itself, but in approximating the way in which the observer looks at the data. To clarify this idea we give an example: if we consider images representing skin lesions, we are not interested in the images per se but rather in approximating the judgment given by the doctors about such images. For further and more detailed information on the theory of GENEOs and its applications we refer to [3], [25], [8], [4], [15], [11], and [1].

An example of application of GENEOs: protein pocket detection

In [5] GENEOnet is introduced. It is a method to identify promising binding sites in proteins, called *pockets*, a key problem in drug design. Over the past few years, many researchers have focused on this problem, involving two main tools: a geometrical analysis for detecting the empty regions within a protein structure and a physicochemical analysis to characterize the interacting and structural properties of the found pockets in order to prioritize them and to identify the correct binding sites. Figure 5 illustrates the result obtained by applying suitable GENEOs to the protein 2QWE.

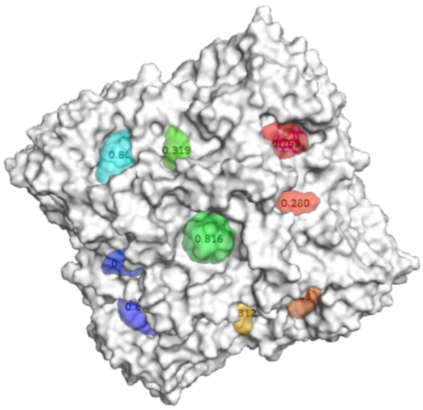

FIGURE 5: Model predictions for protein 2QWE. The global view of the prediction is shown, where different pockets are depicted in different colors and are labelled with their scores. (image courtesy of [5])

Figure 6 shows the comparison between the results obtained by the GENEOnet and other state-of-the-art methods for pocket detection in proteins (the higher the lines, the better). The use of GENEOs allows one to obtain better results, despite using fewer parameters. Moreover, the method assigns a druggability score for each identified pocket. In this way it is possible to rank the pockets on the same molecule by scoring them in decreasing order. In Figure 6 the notation H_j refers to the portion of correct recognitions by the j-th top ranked pocket:

$$H_j = \frac{\#(\text{proteins whose true pocket is hit by the } j-\text{th top ranked})}{\#(\text{proteins})},$$

while T_j refers to the corresponding cumulative quantities

$$T_j = \frac{\#(\text{proteins whose true pocket is hit within the } j-\text{th top ranked})}{\#(\text{proteins})}$$
$$= \sum_{i=1}^{j} H_j.$$

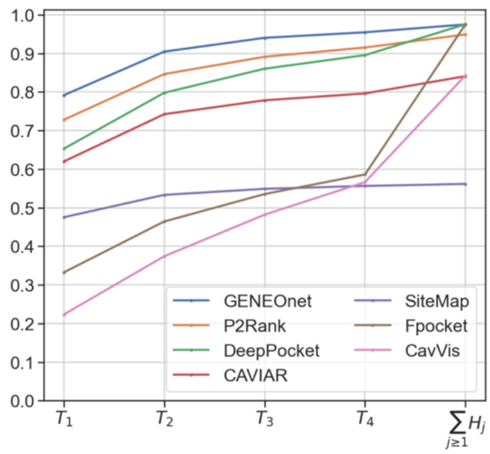

FIGURE 6: Results of comparison on a test set. In this figure, the cumulative frequencies curves are shown. (image courtesy of [5])

From full equivariance to partial equivariance

Incorporating symmetries as an inductive bias into neural network architectures, Geometric Deep Learning and Topological Data Analysis have given an improvement in the development of deep learning models. However, real-world data rarely follows strict mathematical symmetries due to noisy or incomplete data or to symmetry breaking features.

In the theory of GENEOs the collection of all symmetries is represented by a group, but in some applications, the group axioms are not maintained. As

an example, we can take a data set that contains images of digits and the group of rotations as the equivariance group. As shown in Figure 2, if we take an image of the digit '6' and rotate it by an angle of π, we obtain an image that the user would interpret as '9'. At the same time, we want to be able to rotate the digit '6' by small angles while retaining its meaning.
In such model for digit classification, equivariance should be given with respect to a set that is not closed under composition, hence not a group. Thus, in applications it may sometimes be more useful to consider only some transformations and to ignore others. Another example is the simple translation of an image: we can consider a data set that contains images located in a rectangle of the screen and we want to transpose them into another rectangle. In this case, the identity is not relevant, since we are not interested in the original images, but in their transformations. Therefore, there is a necessity to extend the theory of GENEOs, considering weaker structures than groups. The main goal of this thesis is to generalize the known results of GENEOs in this new mathematical framework.

Partial Group Equivariant Non-Expansive Operators (P-GENEOs)

We introduce the concept of Partial Group Equivariant Non-Expansive Operator (P-GENEO) in order to extend the results obtained for GENEOs to a more general set-up. We consider the same mathematical framework described for GENEOs: data are represented as functions defined on topological spaces and they cannot be studied in a direct way, but only through the action of agents that measure and transform them.
In this new model there are some substantial differences with respect to the theory of GENEOs:

1. The user chooses two data sets in input: the set containing the original measurements and another set that encloses the admissible variations of such measurements, defined in the same domain. For example, in the case where the function that represents the digit '6' is being observed, we define an initial space that contains this function and another space that contains certain small rotations of '6', chosen by the observer, but excludes all others.

2. Instead of considering a group of transformations we will consider a

set containing only those that do not change the meaning of our data, i.e., only those bringing each original signal inside the set of admissible variations of measurements. Therefore, by choosing the initial spaces, the user defines also which transformations of the data set by right composition are admissible and which are not.

3. Finally, we define the concept of P-GENEO as a weakening of the concept of GENEO. In this way, depending on the application considered, the observer also has the possibility of restricting the set of admissible transformations.

With these assumptions in mind, in this thesis, we will proceed by extending results obtained for GENEOs, following in particular [3] and [25]. We will study topological and metric properties of P-GENEOs trying to formalize the role of the observer in data analysis, with more freedom in the choice of the environment in which he/she can operate and act. We will define suitable pseudo-metrics for the function spaces, the set of transformations and the set of non-expansive operators. Grounding on certain topological structures, we will proceed by proving compactness and convexity of the space of P-GENEOs, under the assumption that function spaces are compact and convex. These are useful properties from a computational point of view: a compact space can be approximated by a finite set and the convexity guarantees the possibility of creating new P-GENEOs, as convex combination of some preexisting ones. Proceeding along this line, we will realize general methods for building P-GENEOs, starting from a finite or infinite set of known P-GENEOs.

Related works

The main motivation for this dissertation is that real-world data rarely follows mathematical symmetries. This problem is not new in GDL and some steps has been already being made in this direction by researchers in different fields (see [26], [28], [29], [30], and [18]).
Group restriction operation In [30] a *group restriction operation* is proposed, that allows to realize network architectures which are decreasingly equivariant with depth. This is useful, for example, for natural images which show low level features like edges in arbitrary orientations but carry a sense of preferred orientation globally. Each individual layer of a convolutional

network should therefore be adapted to the symmetries present in the length scale of its point of view and this loss of symmetry can be implemented by restricting the equivariance at a certain depth. The group restriction operation yields models that are locally equivariant but are globally invariant only to the level of symmetry present in the data and it is shown experimentally that a restriction at later stages of the model improves the performance.

Approximately Equivariance An operator is called *approximately equivariant* when equivariance holds up to small variations. Formally, let Φ, Ψ be sets of functions endowed with the uniform norm, and $F\colon \Phi \to \Psi$ be an operator between them. Let G, H be groups acting on Φ and Ψ respectively and $T\colon G \to H$ an homomorphism. Consider a real number $\varepsilon > 0$. We say that F is ε-approximately T-equivariant if for any $g \in G$ and for any $\varphi \in \Phi$ we have that $\|F(\varphi g) - F(\varphi)T(g)\|_\infty \leq \varepsilon$. Note that strictly equivariant functions are 0-approximately equivariant.

To address the need for more interpretable priors, in [18] Residual Pathway Priors (RPPs) are illustrated. This is a method for converting hard architectural constraints into soft priors. Practically, RPPs allow to tackle problems in which perfect symmetry has been violated, but approximate symmetry is still present, as it is the case for most real-world physical systems. The core implementation of RPP is to expand each layer in the model into a sum of both a restrictive layer that encodes the hard architectural constraints and a generic more flexible layer but it penalizes the more flexible path via a lower prior probability.

In article [29] a new class of approximately equivariant networks for modeling imperfectly symmetric dynamics is introduced, by relaxing equivariant constraints. This is made in order to perform both strictly equivariant networks and highly flexible models in learning dynamics in the real world.

Partial Equivariance Formally, a set is *partial equivariant* with respect to a group if the equivariance property is maintained for a subset of the group and not necessarily for all transformations in the group.

The article [26] introduces the concept of *Partial G-CNNs* in order to represent a G-CNNs model with a partial equivariance to the group. Convolutional Neural Networks (CNNs) are a widely used deep learning architectures that are translation equivariant. Group equivariant CNNs (G-CNNs) extend equivariance to other symmetry groups. The reasoning is the same as presented before: frequently, invariant transformations can be better represented by a subset of a group than the whole group or a subgroup. Therefore, one

should consider a model where data distribution respects equivariance *partially.* Moreover, as in [30], the optimal level of equivariance may change per layer depending on changes in the likelihood of some transformation for low and high-level features. For example, the orientations of edges in a human face can be represented by full rotation equivariance, while the poses of the face with respect to the camera are better described by a subset of all rotations. Nevertheless, manually tuning layer-wise levels of equivariance is not simple and requires iterations over several possible combinations of equivariance levels. Therefore, *Partial Group equivariant CNNs* are introduced, in order to create a model able to learn optimal levels of equivariance directly from data.
This approach is similar to the one considered in this thesis, but with some differences. Instead of choosing the group elements that can act on the data set, we decided to first define the set containing the admissible variations of data. From this setting, we obtain the set of admissible transformations and then the user can choose a subset, depending on the application considered. In our model is also importart the role of the non-expansiveness. It is the feature that allows to obtain compactness and approximability and that distinguishes our model from much existing literature on equivariant machine learning.

Organization of the thesis

The structure of the thesis is the following.
Chapter 1 is dedicated to describe the mathematical setting. First, we define the data sets Φ and Φ' as real-valued function spaces endowed with the uniform norm, where the functions are defined on some domain X. Φ is the set containing the original measurements and Φ' is the set that encloses the admissible variations of such measurements. The admissible transformations are defined as a set of bijections $\mathrm{Aut}_{\Phi,\Phi'}(X)$ in the domain X that, acting on the right on every signal in Φ, transform the signals into measurements belonging to Φ', without getting out of it. We can observe that we can reduce this setting to the case of GENEOs considering $\Phi = \Phi'$. We will consider the space X and the space $\mathrm{Aut}_{\Phi,\Phi'}(X)$ as pseudo-metric spaces, defining two different pseudo-metrics on both, inherited from Φ and Φ'. By choosing the correct pseudo-metric each time, it is possible to prove some results about the structure of the space of admissible transformations.

Chapter 2 is dedicated to the definition of Partial Group Equivariant Non-Expansive Operator (P-GENEO) and the proof of compactness and convexity of the space of P-GENEOs, under the assumption that function spaces are compact and convex. Since in the setting we have two different data sets in input, a P-GENEO can be represented as a pair of non-expansive operators, each acting on one of the data sets. These operators are different from each other, but are connected by the equivariance property.
In Chapter 3 we present some methods to build P-GENEOs, starting from a finite or an infinite set of known P-GENEOs. We conclude the dissertation with some final observations and outlining some open issues that could be the subject of future research.

Epistemological assumptions

The mathematical framework we have developed is motivated by an epistemological background that is based on the following assumptions:

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

2. We consider two different data sets: the set containing the original measurements and another set that encloses the admissible variation of such measurements, defined on the same domain. With admissible variation, we intend the transformations of measurements that do not change their meaning for the observer.

3. Data are not studied directly, but through the action of agents that measure and transform them. Only the pair (data, agent) matters. In our case the data in input are a pair of data sets and as a consequence also the agent is represented by a pair of operators.

4. Agents are described in terms of how they transform data while maintaining some kind of equivariance.

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

Therefore, instead of just analyzing data, our setting is devoted to the study of the pair (data, agent), where data are a pair of data sets and agents are a pair of equivariant operators. Our purpose is to describe these operators from a geometric and topological point of view, extending the results obtained in the theory of GENEOs.

Chapter 1

Mathematical setting

In our mathematical model, data are represented as function spaces, where the functions are real-valued and defined on a topological space. Our aim is to be able to analyze the case where data can only be transformed up to a certain point. The reason for this setting is that in many applications the data points can change meaning after a transformation. For example, the digit '6' rotated by an angle of π becomes the digit '9', which is obviously different. Therefore, instead of considering a group of transformations we will consider a subset containing only those that do not change the meaning of our data. In order to do this, it will be necessary to define two different function spaces: the space of the initial data and the space in which these data can vary, in an admissible way.

1.1 Data sets and operations

Consider a set X and the normed vector space $(\mathbb{R}_b^X, \|\cdot\|_\infty)$, where $\mathbb{R}_b^X$ is the space of all bounded real-valued functions on X and $\|\cdot\|_\infty$ is the usual uniform norm, i.e., for any $f \in \mathbb{R}_b^X$, $\|f\|_\infty := \sup_{x\in X}|f(x)|$. On the set X we consider transformations given by elements of $\mathrm{Aut}(X)$, i.e., the group of bijections from X to itself. Then, we can consider the right group action $\mathcal{R}$ defined as follows:

$$\mathcal{R}\colon \mathbb{R}_b^X \times \mathrm{Aut}(X) \to \mathbb{R}_b^X, \qquad (\varphi, s) \mapsto \varphi s.$$

Remark 1.1. For every $s \in \mathrm{Aut}(X)$, the map $\mathcal{R}_s\colon \mathbb{R}_b^X \to \mathbb{R}_b^X$, with $\mathcal{R}_s(\varphi) := \varphi s$ preserves the distances. In fact, for any $\varphi_1, \varphi_2 \in \mathbb{R}_b^X$, by bijectivity of s, we have that

$$\begin{aligned}\|\mathcal{R}_s(\varphi_1) - \mathcal{R}_s(\varphi_2)\|_\infty &= \sup_{x\in X} |\varphi_1 s(x) - \varphi_2 s(x)| \\ &= \sup_{y\in X} |\varphi_1(y) - \varphi_2(y)| \\ &= \|\varphi_1 - \varphi_2\|_\infty,\end{aligned}$$

denoting $y = s(x)$.

Our data sets are given by two sets Φ and Φ' of bounded real-valued measurements on X. In our model, X represents the space where the measurements can be made, Φ is the space of permissible measurements, and Φ' is a space in which Φ can vary, without changing the meaning. In other words, we want to be able to apply some admissible transformations on the space X, so that the resulting changes in the measurements in Φ are contained in the space Φ'. Thus, in our model, we consider operations on X in the following admissible way:

Definition 1.2. A (Φ, Φ')**-operation** is an element s of $\mathrm{Aut}(X)$ such that, for any measurement $\varphi \in \Phi$, the composition φs belongs to Φ'.
The set of all (Φ, Φ')-operations is denoted by $\mathrm{Aut}_{\Phi,\Phi'}(X)$.

Remark 1.3. We can observe that the identity function id_X is an element of $\mathrm{Aut}_{\Phi,\Phi'}(X)$ if and only if $\Phi \subseteq \Phi'$.

For any $s \in \mathrm{Aut}_{\Phi,\Phi'}(X)$, the restriction to Φ of the map $\mathcal{R}_s$ takes values in Φ' since $\mathcal{R}_s(\varphi) := \varphi s \in \Phi'$ for any $\varphi \in \Phi$. We can consider the restriction to Φ of the map $\mathcal{R}$:

$$\mathcal{R}\colon \Phi \times \mathrm{Aut}_{\Phi,\Phi'}(X) \to \Phi', \qquad (\varphi, s) \mapsto \varphi s$$

where $\mathcal{R}(\varphi, s) = \mathcal{R}_s(\varphi)$, for every $s \in \mathrm{Aut}_{\Phi,\Phi'}(X)$ and every $\varphi \in \Phi$.

Definition 1.4. Let X be a set. A **perception triple** is a triple (Φ, Φ', S) with $\Phi, \Phi' \subseteq \mathbb{R}_b^X$ and $S \subseteq \mathrm{Aut}_{\Phi,\Phi'}(X)$.

In our model, the subset $S \subseteq \mathrm{Aut}_{\Phi,\Phi'}(X)$ represents the space of all the admissible operations that the observer is allowed to apply to the data set Φ. Also the triple $(\Phi, \Phi', \mathrm{Aut}_{\Phi,\Phi'}(X))$ is a perception triple, called *universal*.

Example 1.5. Given $X = \mathbb{R}^2$, consider two rectangles R and R' in X. Assume $\Phi := \{\varphi\colon X \to [0,1] : \operatorname{supp}(\varphi) \subseteq R\}$ and $\Phi' := \{\varphi'\colon X \to [0,1] : \operatorname{supp}(\varphi') \subseteq R'\}$. We recall that, if we consider a function $f\colon X \to \mathbb{R}$, the *support* of f is the set of points in the domain where the function does not vanish, i.e., $\operatorname{supp}(f) = \{x \in X \mid f(x) \neq 0\}$.
Consider S as the set of translations that bring R into R'. The triple (Φ, Φ', S) is a perception triple. This example could represent a model where rectangular pictures can only be changed by translations belonging to a bounded set.

1.2 Pseudo-metrics on data sets

In our model, data are represented as function spaces, that is, considering a generic set X, a set $\Omega \subseteq \mathbb{R}_b^X$ of bounded real-valued functions $\omega\colon X \to \mathbb{R}$. We can endow the space X with an extended pseudo-metric induced by Ω:

$$D_X^\Omega(x_1, x_2) = \sup_{\omega\in\Omega} |\omega(x_1) - \omega(x_2)|$$

for every $x_1, x_2 \in X$.

Proposition 1.6. *D_X^Ω is an extended pseudo-metric.*

Proof. We have that $D_X^\Omega(x, x) = 0$ for any $x \in X$ and D_X^Ω is symmetric. It remains to show that triangle inequality holds:

$$\begin{aligned} D_X^\Omega(x_1, x_2) &= \sup_{\omega\in\Omega} |\omega(x_1) - \omega(x_2)| \\ &\leq \sup_{\omega\in\Omega} (|\omega(x_1) - \omega(x_3)| + |\omega(x_3) - \omega(x_2)|) \\ &\leq \sup_{\omega\in\Omega} |\omega(x_1) - \omega(x_3)| + \sup_{\omega\in\Omega} |\omega(x_3) - \omega(x_2)| \\ &= D_X^\Omega(x_1, x_3) + D_X^\Omega(x_3, x_2) \end{aligned}$$

for any $x_1, x_2, x_3 \in X$. □

Remark 1.7. The choice of this pseudo-metric over X means that two points can only be distinguished if they assume different values for some measurements. For example, if Φ contains only a constant function and X contains at least two points, the distance between any two points of X is always null.

The pseudo-metric space $X_\Omega := (X, D_X^\Omega)$ can be considered as a topological space with the bases

$$\mathcal{B}_\Omega = \{B_\Omega(x_0, r)\}_{x_0 \in X,\ r \in \mathbb{R}^+} = \big\{\{x \in X :\ D_X^\Omega(x, x_0) < r\}\big\}_{x_0 \in X,\ r \in \mathbb{R}^+},$$

and the induced topology is denoted by τ_Ω. The reason for considering a topological space X, rather than just a set, follows from the need of formalising the assumption that data are stable under small perturbations.

Remark 1.8. In our case, there are two collections of functions Φ and Φ' in $\mathbb{R}_b^X$ representing our data, both of which induce a topology on X. So, in the model, we will consider the pseudo-metric spaces X_Φ and $X_{\Phi'}$ with the same underlying set X.

In general the topologies induced on X by two different subsets of $\mathbb{R}_b^X$ are not comparable, but they are in this special case:

Proposition 1.9. *If $\Phi \subseteq \Phi' \subseteq \mathbb{R}_b^X$, then the topologies $\tau_{\Phi'}$ and $\tau_{\Phi'}$ are comparable and, in particular, $\tau_{\Phi'}$ is finer than τ_Φ.*

Proof. It will suffice to show that every open set $U \in \tau_\Phi$ is also in $\tau_{\Phi'}$. Since $D_X^\Phi(x_1, x_2) \leq D_X^{\Phi'}(x_1, x_2)$ for every $x_1, x_2 \in X$, we have that

$$\begin{aligned} B_{\Phi'}(x_0, r) &= \{x \in X :\ D_X^{\Phi'}(x, x_0) < r\} \\ &= \{x \in X : \sup_{\varphi \in \Phi} |\varphi(x) - \varphi(x_0)| \leq \sup_{\varphi' \in \Phi'} |\varphi'(x) - \varphi'(x_0)| < r\} \\ &\subseteq \{x \in X :\ \sup_{\varphi \in \Phi} |\varphi(x) - \varphi(x_0)| < r\} \\ &= B_\Phi(x_0, r) \end{aligned}$$

for every $x_0 \in X,\ r \in \mathbb{R}^+$. Then, considering an open set $U \in \tau_\Phi$, for every $x_0 \in U$ there exists $r > 0$ such that $x_0 \in B_{\Phi'}(x_0, r) \subseteq B_\Phi(x_0, r) \subseteq U$. Then $U \in \tau_{\Phi'}$ and the statement is true. □

Now, given a set $\Omega \subseteq \mathbb{R}_b^X$, we will prove a result about the compactness of the pseudo-metric space X_Ω. Let us recall the following Lemma (see [19]):

Lemma 1.10. *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.*

Theorem 1. *If Ω is totally bounded, then X_Ω is totally bounded.*

Proof. By Lemma 1.10 it will suffice to prove that every sequence in X admits a Cauchy subsequence with respect to the pseudo-metric D_X^Ω. Consider a sequence $(x_i)_{i\in\mathbb{N}}$ in X_Ω and take a real number $\varepsilon > 0$. Since Ω is totally bounded, we can find a finite subset $\Omega_\varepsilon = \{\omega_1, \dots, \omega_n\}$ such that for every $\omega \in \Omega$ there exists $\omega_r \in \Omega$ for which $\|\omega - \omega_r\|_\infty < \varepsilon$. We can consider now the real sequence $(\omega_1(x_i))_{i\in\mathbb{N}}$, that is bounded since $\Omega \in \mathbb{R}_b^X$. From Bolzano-Weierstrass Theorem it follows that we can extract a convergent subsequence $(\omega_1(x_{i_h}))_{h\in\mathbb{N}}$. Again, we can extract from $(\omega_2(x_{i_h}))_{h\in\mathbb{N}}$ another convergent subsequence $(\omega_2(x_{i_{h_t}}))_{t\in\mathbb{N}}$. Repeating the process, we are able to extract a subsequence of $(x_i)_{i\in\mathbb{N}}$, that for semplicity of notation we can indicate as $(x_{i_j})_{j\in\mathbb{N}}$, such that $(\omega_k(x_{i_j}))_{j\in\mathbb{N}}$ is a convergent subsequence in $\mathbb{R}$ for every $k \in \{1, \dots, n\}$ and hence a Cauchy sequence in $\mathbb{R}$.
By construction, Ω_ε is finite, then we can find an index $\bar{\jmath}$ such that for any $k \in \{1, \dots, n\}$

$$|\omega_k(x_{i_l}) - \omega_k(x_{i_m})| \leq \varepsilon, \quad \text{for every} \quad l, m \geq \bar{\jmath}.$$

Furthermore we have that, for any $\omega \in \Omega$, any $\omega_k \in \Omega_\varepsilon$ and any $l, m \in \mathbb{N}$

$$\begin{aligned} |\omega(x_{i_l}) - \omega(x_{i_m})| &\leq |\omega(x_{i_l}) - \omega_k(x_{i_l})| + |\omega_k(x_{i_l}) - \omega_k(x_{i_m})| \\ &\quad + |\omega_k(x_{i_m}) - \omega(x_{i_m})| \\ &\leq \|\omega - \omega_k\|_\infty + |\omega_k(x_{i_l}) - \omega_k(x_{i_m})| + \|\omega_k - \omega\|_\infty. \end{aligned}$$

We observe that the choice of $\bar{\jmath}$ depends only on ε and Ω_ε, not on k. Then, choosing a $\omega_k \in \Omega_\varepsilon$ such that $\|\omega_k - \omega\|_\infty < \varepsilon$, we get $\|\omega(x_{i_l}) - \omega(x_{i_m})\|_\infty < 3\varepsilon$ for every $\omega \in \Omega$ and every $l, m \geq \bar{\jmath}$. Then,

$$D_X^\Omega(x_{i_l}, x_{i_m}) = \sup_{\omega\in\Omega} |\omega(x_{i_l}) - \omega(x_{i_m})| < 3\varepsilon \quad \text{for every} \quad l, m \geq \bar{\jmath}.$$

Then $(x_{i_j})_{j\in\mathbb{N}}$ is a Cauchy sequence in X_Ω. For Lemma 1.10 the statement holds. □

Corollary 1.11. *If Ω is totally bounded and X_Ω is complete, then X_Ω is compact.*

Proof. From Theorem 1 we have that X_Ω is totally bounded and since by hypothesis it is also complete, it is compact. □

Now, we will prove that the choice of the pseudo-metric D_X^Ω on X makes the functions of Ω non-expansive.

Definition 1.12. Consider two pseudo-metric spaces (P, d_p) and (Q, d_q). A **non-expansive** function from (P, d_p) to (Q, d_q) is a function $f\colon P \to Q$ such that $d_q(f(p_1), f(p_2)) \leq d_p(p_1, p_2)$ for any $p_1, p_2 \in P$.
We call $\mathbf{NE}(P, Q)$ the space of all non-expansive functions from (P, d_p) to (Q, d_q).

Proposition 1.13. $\Omega \subseteq \mathbf{NE}(X_\Omega, \mathbb{R})$.

Proof. For any $x_1, x_2 \in X$ we have that

$$|\omega(x_1) - \omega(x_2)| \leq \sup_{\omega \in \Omega} |\omega(x_1) - \omega(x_2)| = D_X^\Omega(x_1, x_2).$$

□

Then, the topology on X induced by D_X^Ω naturally makes the measurements in Ω continuous. In particular, since the previous results hold for a generic $\Omega \subseteq \mathbb{R}_b^X$, they are also true for Φ and Φ' in our model.

Remark 1.14. The topology on X induced by the pseudo-metric of one of the function spaces does not make the functions of the other continuous: a function $\varphi' \in \Phi'$ may not be continuous from X_Φ to $\mathbb{R}$ and a function $\varphi \in \Phi$ may not be continuous from $X_{\Phi'}$ to $\mathbb{R}$.

Example 1.15. Assume $X = \mathbb{R}$ and for every $a, b \in \mathbb{R}$ consider the functions $\varphi_a\colon X \to \mathbb{R}$ and $\varphi'_b\colon X \to \mathbb{R}$ defined by setting

$$\varphi_a(x) = \begin{cases} 0 & \text{if } x \geq a \\ 1 & \text{otherwise} \end{cases}, \qquad \varphi'_b(x) = \begin{cases} 0 & \text{if } x \leq b \\ 1 & \text{otherwise} \end{cases}.$$

Suppose $\Phi := \{\varphi_a : a \geq 0\}$ and $\Phi' := \{\varphi'_b : b \leq 0\}$, and consider $s \in \mathrm{Aut}_{\Phi,\Phi'}(X)$ as the symmetry with respect to the y-axis, i.e., such that $s(x) = -x$. We can observe that the function $\varphi_1 \in \Phi$ is not continuous from X'_Φ to $\mathbb{R}$; indeed $D_X^{\Phi'}(0, 2) = 0$, but $|\varphi_1(0) - \varphi_1(2)| = 1$.

However, if $\Phi \subseteq \Phi'$, we have that the functions of Φ maintain continuity on $X_{\Phi'}$, indeed:

Proposition 1.16. *If* $\Phi \subseteq \Phi'$*, then* $\Phi \subseteq \boldsymbol{NE}(X_{\Phi'}, \mathbb{R})$.

Proof. By Proposition 1.13 the statement trivially holds since $\Phi \subseteq \Phi' \subseteq \mathbf{NE}(X_{\Phi'}, \mathbb{R})$. □

We recall that the *initial topology* τ_{in} is the coarsest topology that makes all functions in Ω continuous. Under the assumption that Ω is totally bounded, there is another interesting result which gives us a connection between τ_Ω and the initial topology τ_{in} with respect to Ω, when we take the Euclidean topology τ_e on $\mathbb{R}$.

Proposition 1.17. *If Ω is totally bounded, then the topology τ_Ω coincides with τ_{in}.*

Before proceeding with the proof we show the following Lemma:

Lemma 1.18. *If Ω is totally bounded, then for every $\delta > 0$ there exists a finite subset Ω_δ of Ω such that*

$$0 \leq \sup_{\omega \in \Omega} |\omega(x_1) - \omega(x_2)| - \max_{\omega \in \Omega_\delta} |\omega(x_1) - \omega(x_2)| \leq 2\delta$$

for every $x_1, x_2 \in X$.

Proof. Since by hypothesis Ω is totally bounded, there exists a finite subset $\Omega_\delta := \{\omega_1, \dots, \omega_n\}$ of Ω such that for each $\omega \in \Omega$ we can find $\omega_i \in \Omega_\delta$ for which $\|\omega - \omega_i\|_\infty < \delta$. Hence, for every $x \in X$, we have that $|\omega(x) - \omega_i(x)| < \delta$. Considering two points $x_1, x_2 \in X$, for any $\varepsilon > 0$ we can choose a $\omega_\varepsilon \in \Omega$ such that

$$\sup_{\omega \in \Omega} |\omega(x_1) - \omega(x_2)| - |\omega_\varepsilon(x_1) - \omega_\varepsilon(x_2)| \leq \varepsilon.$$

Now, for any $\omega \in \Omega$ we can take an index $i \in \{1, \dots, n\}$ for which $\|\omega - \omega_i\|_\infty < \delta$ and we have that

$$\begin{aligned}
|\omega(x_1) - \omega(x_2)| &= |\omega(x_1) - \omega_i(x_1) + \omega_i(x_1) - \omega_i(x_2) + \omega_i(x_2) - \omega(x_2)| \\
&\leq |\omega(x_1) - \omega_i(x_1)| + |\omega_i(x_1) - \omega_i(x_2)| + |\omega_i(x_2) - \omega(x_2)| \\
&< |\omega_i(x_1) - \omega_i(x_2)| + 2\delta \\
&\leq \max_{\omega_j \in \Omega_\delta} |\omega_j(x_1) - \omega_j(x_2)| + 2\delta.
\end{aligned}$$

We have found the inequality

$$|\omega(x_1) - \omega(x_2)| \leq \max_{\omega_j \in \Omega_\delta} |\omega_j(x_1) - \omega_j(x_2)| + 2\delta,$$

for every $\omega \in \Omega$. Then, we have that

$$\sup_{\omega\in\Omega} |\omega(x_1) - \omega(x_2)| \leq \max_{\omega_j\in\Omega_\delta} |\omega_j(x_1) - \omega_j(x_2)| + 2\delta.$$

On the other hand, since $\Omega_\delta \subseteq \Omega$, we have that

$$\sup_{\omega\in\Omega} |\omega(x_1) - \omega(x_2)| \geq \max_{\omega_j\in\Omega_\delta} |\omega_j(x_1) - \omega_j(x_2)|.$$

Hence, the statement is proved. □

We are now ready to prove Proposition 1.17:

Proof. We already know, because of Proposition 1.13, that every function ω of Ω is continuous with respect to τ_Ω. Hence, τ_Ω is obviously finer than τ_{in}. Consider two bases respectively for the topology τ_Ω and the topology τ_{in}:

$$\mathcal{B}_\Omega = \{B_\Omega(x_0, r), x_0 \in X, r > 0\},$$
$$\mathcal{B}_{\text{in}} = \left\{\bigcap_{i\in I} \omega_i^{-1}(U_i), |I| < \infty, U_i \in \tau_e, \omega_i \in \Omega \text{ for every } i \in I\right\}.$$

In order to prove our statement it will suffice to show that the topology τ_{in} is finer than the topology τ_Ω. Then, considering $x_0 \in B_\Omega(x_0, r) \in \mathcal{B}_\Omega$, we want to prove that there exists a set $V = \bigcap_{i\in I} \omega_i^{-1}(U_i) \in \mathcal{B}_{\text{in}}$ such that $x_0 \in V \subseteq B_\Omega(x_0, r)$.
Since Ω is totally bounded then, because of Lemma 1.18, for every $\delta > 0$ there exists a finite subset $\Omega_\delta := \{\omega_1, \ldots, \omega_n\}$ of Ω such that

$$0 \leq \sup_{\omega\in\Omega} |\omega(x_1) - \omega(x_2)| - \max_{\omega\in\Omega_\delta} |\omega(x_1) - \omega(x_2)| \leq 2\delta \tag{1.1}$$

for every $x_1, x_2 \in X$. Consider now

$$B_\delta(x_0, r) := \left\{x \in X \,\middle|\, \max_{\omega_i\in\Omega_\delta} |\omega_i(x_0) - \omega_i(x)| < r\right\}$$

for every $x_0 \in X$ and every $r > 0$. We can choose $r, \delta > 0$ such that $2\delta < r$. Consider $x_0 \in B_\Omega(x_0, r) \in \mathcal{B}_\Omega$; from inequality (1.1) we have that $B_\delta(x_0, r - 2\delta) \subseteq B_\Omega(x_0, r)$.
We now define in $\mathbb{R}$ the sets $U_i :=]\omega_i(x_0) - r + 2\delta, \omega_i(x_0) + r - 2\delta[$ for $i \in I$ and we consider in X the set $V := \bigcap_{\omega\in\Omega_\delta} \omega_i^{-1}(U_i) \subseteq \mathcal{B}_{\text{in}}$.
We have that $x_0 \in V$ and it remains to show that $V \subseteq B_\Omega(x_0, r)$. If $z \in V$, then $|\omega_i(z) - \omega_i(x_0)| < r - 2\delta$ for every $\omega_i \in \Omega_\delta$. Therefore, $x_0 \in V \subseteq B_\delta(x_0, r - 2\delta) \subseteq B_\Omega(x_0, r)$. Then τ_{in} is finer than τ_Ω. Hence, the statement is true. □

1.3 Pseudo-metrics and structure on the space of operations

Proposition 1.19. *Every element of* $\mathrm{Aut}_{\Phi,\Phi'}(X)$ *is non-expansive from* $X_{\Phi'}$ *to* X_Φ.

Proof. Considering a bijection $s \in \mathrm{Aut}_{\Phi,\Phi'}(X)$ we have that

$$\begin{aligned} D_X^\Phi(s(x_1), s(x_2)) &= \sup_{\varphi\in\Phi} |\varphi s(x_1) - \varphi s(x_2)| \\ &= \sup_{\varphi\in\Phi s} |\varphi(x_1) - \varphi(x_2)| \\ &\le \sup_{\varphi'\in\Phi'} |\varphi'(x_1) - \varphi'(x_2)| = D_X^{\Phi'}(x_1, x_2) \end{aligned}$$

for every $x_1, x_2 \in X$. Then, $s \in \mathbf{NE}(X_{\Phi'}, X_\Phi)$ and the statement is proved. □

Now we are ready to put more structure on $\mathrm{Aut}_{\Phi,\Phi'}(X)$. Considering a set $\Omega \subseteq \mathbb{R}_b^X$ of bounded real-valued functions $\omega\colon X \to \mathbb{R}$, we can endow the set $\mathrm{Aut}(X)$ with a pseudo-metric inherited from Ω:

$$D_{\mathrm{Aut}}^\Omega(s_1, s_2) := \sup_{\omega\in\Omega} \|\omega s_1 - \omega s_2\|_\infty$$

for any s_1, s_2 in $\mathrm{Aut}(X)$.

Remark 1.20. As in Remark 1.8, the sets Φ and Φ' can endow $\mathrm{Aut}(X)$ with two possible pseudo-metrics D_{Aut}^Φ and $D_{\mathrm{Aut}}^{\Phi'}$.
In particular, we can consider $\mathrm{Aut}_{\Phi,\Phi'}(X)$ as a pseudo-metric subspace of $\mathrm{Aut}(X)$ with the induced pseudo-metrics.

Remark 1.21. We observe that, for any s_1, s_2 in $\mathrm{Aut}(X)$,

$$\begin{aligned} D_{\mathrm{Aut}}^\Omega(s_1, s_2) &:= \sup_{\omega\in\Omega} \|\omega s_1 - \omega s_2\|_\infty \\ &= \sup_{x\in X} \sup_{\omega\in\Omega} |\omega(s_1(x)) - \omega(s_2(x))| \\ &= \sup_{x\in X} D_X^\Omega(s_1(x), s_2(x)). \end{aligned} \tag{1.2}$$

From this result, we can see that the metric we defined, which is based on the action of the elements of $\mathrm{Aut}(X)$ on the set Ω, is exactly the uniform metric on X_Ω.

1.3.1 Some observations on the composition of operations

Consider the following sets of transformations:

$$\begin{aligned} \mathrm{Aut}_{\Phi}(X) &:= \{s \in \mathrm{Aut}(X), \varphi s \in \Phi \text{ for every } \varphi \in \Phi\}; \\ \mathrm{Aut}_{\Phi'}(X) &:= \{s \in \mathrm{Aut}(X), \varphi' s \in \Phi' \text{ for every } \varphi' \in \Phi'\}; \\ \mathrm{Aut}_{\Phi,\Phi'}(X) &= \{s \in \mathrm{Aut}(X), \varphi s \in \Phi' \text{ for every } \varphi \in \Phi\} \end{aligned}$$

It can be observed that in $\mathrm{Aut}_{\Phi,\Phi'}(X)$ it seems natural to choose only the pseudo-metric D^{Φ}_{Aut} since we are interested only in transformations of functions in Φ. However, when we look at the composition of certain elements in $\mathrm{Aut}_{\Phi,\Phi'}(X)$, it is necessary to take the pseudo-metric $D^{\Phi'}_{\mathrm{Aut}}$ in order to maintain the continuity of the composition operation, whenever it is admissible. Consider two elements s, t in $\mathrm{Aut}_{\Phi,\Phi'}(X)$ such that st is still an element of $\mathrm{Aut}_{\Phi,\Phi'}(X)$, i.e., for every function $\varphi \in \Phi$ we have that $\varphi st \in \Phi'$. Then, for any $\varphi \in \Phi$ we have that

$$\varphi' := \varphi s \in \Phi s \subseteq \Phi', \quad \varphi' t \in \Phi'.$$

Therefore, t is also an element of $\mathrm{Aut}_{\Phi s,\Phi'}(X)$. By definition Φs is contained in Φ' for every $s \in \mathrm{Aut}_{\Phi,\Phi'}(X)$ and this justifies the choice of considering in $\mathrm{Aut}_{\Phi,\Phi'}(X)$ also the pseudo-metric $D^{\Phi'}_{\mathrm{Aut}}$.
We have shown in particular that if s, t are elements of $\mathrm{Aut}_{\Phi,\Phi'}(X)$ such that st is still an element of $\mathrm{Aut}_{\Phi,\Phi'}(X)$, then t is an element of $\mathrm{Aut}_{\Phi s,\Phi'}(X)$, which is an implication of the following Proposition:

Proposition 1.22. *Let s, t be elements of $\mathrm{Aut}_{\Phi,\Phi'}(X)$. The composition st belongs to $\mathrm{Aut}_{\Phi,\Phi'}(X)$ if and only if t belongs to $\mathrm{Aut}_{\Phi s,\Phi'}(X)$.*

Proof. Consider $s, t \in \mathrm{Aut}_{\Phi,\Phi'}(X)$. If the composition st belongs to $\mathrm{Aut}_{\Phi,\Phi'}(X)$, we have already proved that $t \in \mathrm{Aut}_{\Phi s,\Phi'}(X)$.
On the other hand, if $t \in \mathrm{Aut}_{\Phi s,\Phi'}(X)$ we have that $\bar{\varphi} t \in \Phi'$ for every $\bar{\varphi} \in \Phi s$. Since all the elements of Φs are in the form φs for some $\varphi \in \Phi$, we have that $\varphi st \in \Phi'$ for every $\varphi \in \Phi$. Therefore, $st \in \mathrm{Aut}_{\Phi,\Phi'}(X)$ and the statement is proved. □

Remark 1.23. Let t be an element of $\mathrm{Aut}_{\Phi,\Phi'}(X)$. We can observe that if $s \in \mathrm{Aut}_{\Phi}(X)$, then $\Phi s \subseteq \Phi$ and $st \in \mathrm{Aut}_{\Phi,\Phi'}(X)$.

Corollary 1.24. *Suppose* $\Phi' := \cup_{s\in \mathrm{Aut}_{\Phi,\Phi'}(X)}\Phi s$. *Let* t *be an element of* $\mathrm{Aut}_{\Phi,\Phi'}(X)$. *The composition* st *belongs to* $\mathrm{Aut}_{\Phi,\Phi'}(X)$ *for every* $s \in \mathrm{Aut}_{\Phi,\Phi'}(X)$ *if and only if* t *belongs to* $\mathrm{Aut}_{\Phi'}(X)$.

Proof. Consider $t \in \mathrm{Aut}_{\Phi,\Phi'}(X)$. Because of Proposition 1.22, the composition st belongs to $\mathrm{Aut}_{\Phi,\Phi'}(X)$ for every $s \in \mathrm{Aut}_{\Phi,\Phi'}(X)$ if and only if $t \in \mathrm{Aut}_{\Phi s,\Phi'}(X)$ for every $s \in \mathrm{Aut}_{\Phi,\Phi'}(X)$. However, since $\Phi' = \cup_{s\in \mathrm{Aut}_{\Phi,\Phi'}(X)}\Phi s$, this is true if and only if $t \in \mathrm{Aut}_{\Phi'}(X)$. □

1.3.2 Structure on the set of operations

We denote by $\Pi \subseteq \mathrm{Aut}_{\Phi,\Phi'}(X) \times \mathrm{Aut}_{\Phi,\Phi'}(X)$ the set containing the pairs (s,t) such that st belongs to $\mathrm{Aut}_{\Phi,\Phi'}(X)$.

Lemma 1.25. *Consider* $r, s, t \in \mathrm{Aut}(X)$. *It holds that*

$$D^{\Omega}_{\mathrm{Aut}}(rt, st) = D^{\Omega}_{\mathrm{Aut}}(r, s).$$

Proof. Consider $r, s, t \in \mathrm{Aut}(X)$. Since $\mathcal{R}_t$ preserve distances, we have that:

$$\begin{aligned} D^{\Omega}_{\mathrm{Aut}}(rt, st) &:= \sup_{\omega\in\Omega} \|\omega rt - \omega st\|_\infty \\ &= \sup_{\omega\in\Phi} \|\omega r - \omega s\|_\infty \\ &= D^{\Omega}_{\mathrm{Aut}}(r, s). \end{aligned}$$

□

Lemma 1.26. *Consider* $r, s \in \mathrm{Aut}(X)$ *and* $t \in \mathrm{Aut}_{\Phi,\Phi'}(X)$. *It holds that*

$$D^{\Phi}_{\mathrm{Aut}}(tr, ts) \leq D^{\Phi'}_{\mathrm{Aut}}(r, s).$$

Proof. Consider $r, s \in \mathrm{Aut}(X)$ and $t \in \mathrm{Aut}_{\Phi,\Phi'}(X)$. Since $\Phi t \subseteq \Phi'$, we have that:

$$\begin{aligned} D^{\Phi}_{\mathrm{Aut}}(tr, ts) &= \sup_{\varphi\in\Phi} \|\varphi tr - \varphi ts\|_\infty \\ &= \sup_{\varphi'\in\Phi t} \|\varphi' r - \varphi' s\|_\infty \\ &\leq \sup_{\varphi'\in\Phi'} \|\varphi' r - \varphi' s\|_\infty \\ &= D^{\Phi'}_{\mathrm{Aut}}(r, s). \end{aligned}$$

□

Remark 1.27. We may consider Π to be a pseudo-metric subspace of the product space $(\mathrm{Aut}_{\Phi,\Phi'}(X), D_{\mathrm{Aut}}^{\Phi}) \times (\mathrm{Aut}_{\Phi,\Phi'}(X), D_{\mathrm{Aut}}^{\Phi'})$. In particular, we consider on Π the pseudo-metric

$$D_{\Pi}((s_1, s_2), (t_1, t_2)) = D_{\mathrm{Aut}}^{\Phi}(s_1, s_2) + D_{\mathrm{Aut}}^{\Phi'}(t_1, t_2).$$

Observe that the topology induced by this pseudo-metric is equivalent to the classical product topology.

Proposition 1.28. *Consider* $\Pi \subseteq (\mathrm{Aut}_{\Phi,\Phi'}(X), D_{\mathrm{Aut}}^{\Phi}) \times (\mathrm{Aut}_{\Phi,\Phi'}(X), D_{\mathrm{Aut}}^{\Phi'})$. *Then the function* $\circ\colon \Pi \to (\mathrm{Aut}_{\Phi,\Phi'}(X), D_{\mathrm{Aut}}^{\Phi})$ *that maps* (s,t) *to* st *is continuous.*

Proof. Consider two elements $(s_1, t_1), (s_2, t_2)$ of Π. Because of Lemma 1.25 and Lemma 1.26, we obtain that

$$\begin{aligned} D_{\mathrm{Aut}}^{\Phi}(s_1t_1, s_2t_2) &\leq D_{\mathrm{Aut}}^{\Phi}(s_1t_1, s_2t_1) + D_{\mathrm{Aut}}^{\Phi}(s_2t_1, s_2t_2) \\ &= D_{\mathrm{Aut}}^{\Phi}(s_1, s_2) + D_{\mathrm{Aut}}^{\Phi}(s_2t_1, s_2t_2) \\ &\leq D_{\mathrm{Aut}}^{\Phi}(s_1, s_2) + D_{\mathrm{Aut}}^{\Phi'}(t_1, t_2). \end{aligned}$$

Therefore, the statement is proved. □

Denote as $\Upsilon \subseteq \mathrm{Aut}_{\Phi,\Phi'}(X)$ the set containing the elements s such that s^{-1} belongs to $\mathrm{Aut}_{\Phi,\Phi'}(X)$.

Proposition 1.29. *The function* $(\cdot)^{-1}\colon (\Upsilon, D_{\mathrm{Aut}}^{\Phi'}) \to (\mathrm{Aut}_{\Phi,\Phi'}(X), D_{\mathrm{Aut}}^{\Phi})$, *that maps* s *to* s^{-1}, *is continuous.*

Proof. Consider two bijections $s_1, s_2 \in \Upsilon$. Because of Lemma 1.25 and Lemma 1.26, we obtain that

$$\begin{aligned} D_{\mathrm{Aut}}^{\Phi}(s_1^{-1}, s_2^{-1}) &= D_{\mathrm{Aut}}^{\Phi}(s_1^{-1}s_2, s_2^{-1}s_2) \\ &= D_{\mathrm{Aut}}^{\Phi}(s_1^{-1}s_2, \mathrm{id}_X) \\ &= D_{\mathrm{Aut}}^{\Phi}(s_1^{-1}s_2, s_1^{-1}s_1) \\ &\leq D_{\mathrm{Aut}}^{\Phi'}(s_1, s_2). \end{aligned}$$

Therefore, the statement is true. □

We have previously defined the map

$$\mathcal{R}\colon \Phi \times \mathrm{Aut}_{\Phi,\Phi'}(X) \to \Phi', \qquad (\varphi, s) \mapsto \varphi s$$

where $\mathcal{R}(\Phi, s) = \mathcal{R}_s(\Phi)$, for every $s \in \mathrm{Aut}_{\Phi,\Phi'}(X)$.

Proposition 1.30. *The function $\mathcal{R}$ is continuous, by choosing the pseudo-metric D^{Φ}_{Aut} on $\mathrm{Aut}_{\Phi,\Phi'}(X)$.*

Proof. We have that

$$\begin{aligned}\|\mathcal{R}(\varphi,t)-\mathcal{R}(\overline{\varphi},s)\|_\infty &= \|\varphi t-\overline{\varphi}s\|_\infty\\ &\le \|\varphi t-\varphi s\|_\infty+\|\varphi s-\overline{\varphi}s\|_\infty\\ &= \|\varphi t-\varphi s\|_\infty+\|\varphi-\overline{\varphi}\|_\infty\\ &\le D^{\Phi}_{\mathrm{Aut}}(t,s)+\|\varphi-\overline{\varphi}\|_\infty\end{aligned}$$

for any $\varphi,\overline{\varphi}\in\Phi$ and any $t,s\in\mathrm{Aut}_{\Phi,\Phi'}(X)$. This proves that $\mathcal{R}$ is continuous. □

Now, consider a subset $S\subseteq\mathrm{Aut}_{\Phi,\Phi'}(X)$; we can give a result about the compactness of $(S,D^{\Phi}_{\mathrm{Aut}})$, under suitable assumptions.

Theorem 2. *If Φ and Φ' are totally bounded, then $(S,D^{\Phi}_{\mathrm{Aut}})$ is totally bounded.*

Proof. Consider a sequence $(s_i)_{i\in\mathbb{N}}$ in S and take a real number $\varepsilon>0$. Since 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 $\varphi_r\in\Phi$ for which $\|\varphi-\varphi_r\|_\infty<\varepsilon$. Now, consider the sequence $(\varphi_1 s_i)_{i\in\mathbb{N}}$ in Φ'. Since also Φ' is totally bounded, from Lemma 1.10 it follows that we can extract a Cauchy subsequence $(\varphi_1 s_{i_h})_{h\in\mathbb{N}}$. Again, we can extract another Cauchy subsequence $(\varphi_2 s_{i_{h_t}})_{t\in\mathbb{N}}$. Repeating the process for every $k\in\{1,\ldots,n\}$, we are able to extract a subsequence of $(s_i)_{i\in\mathbb{N}}$, that for simplicity of notation we can indicate as $(s_{i_j})_{j\in\mathbb{N}}$, such that $(\varphi_k s_{i_j})_{j\in\mathbb{N}}$ is a Cauchy sequence for every $k\in\{1,\ldots,n\}$.
By definition Φ_ε is finite, then we can find an index $\bar{\jmath}$ such that for any $k\in\{1,\ldots,n\}$

$$\|\varphi_k s_{i_l}-\varphi_k s_{i_m}\|_\infty\le\varepsilon,\quad\text{for every}\quad l,m\ge\bar{\jmath}. \tag{1.3}$$

Furthermore we have that, for any $\varphi\in\Phi$, any $\varphi_k\in\Phi_\varepsilon$ and any $l,m\in\mathbb{N}$

$$\begin{aligned}\|\varphi s_{i_l}-\varphi s_{i_m}\|_\infty &\le \|\varphi s_{i_l}-\varphi_k s_{i_l}\|_\infty+\|\varphi_k s_{i_l}-\varphi_k s_{i_m}\|_\infty+\|\varphi_k s_{i_m}-\varphi s_{i_m}\|_\infty\\ &= \|\varphi-\varphi_k\|_\infty+\|\varphi_k s_{i_l}-\varphi_k s_{i_m}\|_\infty+\|\varphi_k-\varphi\|_\infty.\end{aligned}$$

We observe that the choice of $\bar{\jmath}$ in (1.3) depends only on ε and Φ_ε, not on k. Then, choosing a $\varphi_k\in\Phi_\varepsilon$ such that $\|\varphi_k-\varphi\|_\infty<\varepsilon$, we get $\|\varphi s_{i_l}-\varphi s_{i_m}\|_\infty<3\varepsilon$ for every $\varphi\in\Phi$ and every $l,m\ge\bar{\jmath}$. Hence,

$$D^{\Phi}_{\mathrm{Aut}}(s_{i_l},s_{i_m})=\sup_{\varphi\in\Phi}\|\varphi s_{i_l}-\varphi s_{i_m}\|_\infty<3\varepsilon$$

Therefore $(s_{i_j})_{j\in\mathbb{N}}$ is a Cauchy sequence. For Lemma 1.10 the statement holds. □

Corollary 1.31. *If* $(S, D^{\Phi}_{\mathrm{Aut}})$ *is complete, then it is also compact.*

Proof. From Theorem 2 we have that S is totally bounded and since by hypothesis it is also complete, the statement holds. □

Chapter 2

The space of P-GENEOs

In this chapter we introduce the concept of Partial Group Equivariant Non-Expansive Operator (P-GENEO). P-GENEOs allow the transformation of data sets, preserving symmetries and distances and maintaining the acceptability conditions of the transformations. We will also describe some topological results about the structure of the space of P-GENEOs. We recall the following definition:

Definition 2.1. Let X be a set. A **perception triple** is a triple (Φ, Φ', S) with $\Phi, \Phi' \subseteq \mathbb{R}_b^X$ and $S \subseteq \mathrm{Aut}_{\Phi,\Phi'}(X)$.
The set X is called the **domain** of the perception triple.

Definition 2.2. Let X, Y be sets and (Φ, Φ', S), (Ψ, Ψ', Q) be perception triples with domains X and Y, respectively. Consider a triple of functions (F, F', T) with the following properties:

- $F\colon \Phi \to \Psi$, $F'\colon \Phi' \to \Psi'$, $T\colon S \to Q$;
- for any $s, t \in S$ such that $st \in S$ it holds that $T(st) = T(s)T(t)$;
- for any $s \in S$ such that $s^{-1} \in S$ it holds that $T(s^{-1}) = T(s)^{-1}$;
- (F, F', T) is *equivariant*, i.e., $F'(\varphi s) = F(\varphi)T(s)$ for every $\varphi \in \Phi$, $s \in S$.

The triple (F, F', T) is called a **perception map** or a **Partial Group Equivariant Operator (P-GEO)** from (Φ, Φ', S) to (Ψ, Ψ', Q).

When the map T is fixed and specified, we will consider simply pairs of operators (F, F') instead of triples (F, F', T) and we say that (F, F') is a P-GEO *associated with* or *with respect to* the map T.
Moreover, in this case we indicate the property of equivariance of the triple (F, F', T) writing that the pair (F, F') is T-*equivariant*.

In Remark 1.3 we observed that $\mathrm{id}_X \in \mathrm{Aut}_{\Phi,\Phi'}(X)$ if and only if $\Phi \subseteq \Phi'$. Then we can consider a perception triple (Φ, Φ', S) with $\Phi \subseteq \Phi'$ and $\mathrm{id}_X \in S \subseteq \mathrm{Aut}_{\Phi,\Phi'}(X)$. Now we will show how a P-GEO from this perception triple behaves.

Lemma 2.3. *Consider two perception triples* (Φ, Φ', S) *and* (Ψ, Ψ', Q) *with domains* X *and* Y*, respectively, and with* $\mathrm{id}_X \in S \subseteq \mathrm{Aut}_{\Phi,\Phi'}(X)$. *Let* (F, F', T) *be a P-GEO from* (Φ, Φ', S) *to* (Ψ, Ψ', Q). *Then* $\Psi \subseteq \Psi'$ *and* $\mathrm{id}_Y \in Q \subseteq \mathrm{Aut}_{\Psi,\Psi'}(Y)$.

Proof. Since (F, F', T) is a P-GEO, by definition, we have that, for any $s, t \in S$ such that $st \in S$, $T(st) = T(s)T(t)$. In particular $\mathrm{id}_X \in S$, then

$$T(\mathrm{id}_X) = T(\mathrm{id}_X \mathrm{id}_X) = T(\mathrm{id}_X)T(\mathrm{id}_X)$$

and hence $T(\mathrm{id}_X) = \mathrm{id}_Y \in Q \subseteq \mathrm{Aut}_{\Psi,\Psi'}(X)$. Moreover, for Remark 1.3, we have that $\Psi \subseteq \Psi'$. □

Proposition 2.4. *Consider two perception triples* (Φ, Φ', S) *and* (Ψ, Ψ', Q) *with domains* X *and* Y*, respectively, and with* $\mathrm{id}_X \in S \subseteq \mathrm{Aut}_{\Phi,\Phi'}(X)$. *Let* (F, F', T) *be a P-GEO from* (Φ, Φ', S) *to* (Ψ, Ψ', Q). *Then* $F'|_\Phi = F$.

Proof. Since (F, F', T) is a P-GEO, it is equivariant and by Lemma 2.3 we have that

$$F'(\varphi) = F'(\varphi \mathrm{id}_X) = F(\varphi)T(\mathrm{id}_X) = F(\varphi)\mathrm{id}_Y = F(\varphi)$$

for every $\varphi \in \Phi$. □

Definition 2.5. Assume that (Φ, Φ', S) and (Ψ, Ψ', Q) are perception triples. If (F, F', T) is a perception map from (Φ, Φ', S) to (Ψ, Ψ', Q) and F, F' are non-expansive , i.e.,

$$\|F(\varphi_1) - F(\varphi_2)\|_\infty \leq \|\varphi_1 - \varphi_2\|_\infty,$$
$$\|F'(\varphi'_1) - F'(\varphi'_2)\|_\infty \leq \|\varphi'_1 - \varphi'_2\|_\infty$$

for every $\varphi_1, \varphi_2 \in \Phi$, $\varphi'_1, \varphi'_2 \in \Phi'$, then (F, F', T) is called a **Partial Group Equivariant Non-Expansive Operator (P-GENEO)**.

In other words, a P-GENEO is a triple (F, F', T) such that F, F' are non-expansive and the following diagram commutes:

$$\begin{array}{ccc} \Phi & \xrightarrow{\mathcal{R}_s} & \Phi' \\ {\scriptstyle F}\big\downarrow & & \big\downarrow{\scriptstyle F'} \\ \Psi & \xrightarrow{\mathcal{R}_{T(s)}} & \Psi' \end{array}$$

for every $s \in S$.

Remark 2.6. We can observe that a GENEO can be represented as a special case of P-GENEO, considering two perception triples (Φ, Φ', S), (Ψ, Ψ', Q) in which:

1. the two function spaces are equal, $\Phi = \Phi'$ and $\Psi = \Psi'$;
2. the subsets containing the invariant transformations S and Q are groups (and then, by definition, the map $T: S \to Q$ is a homomorphism).

In this setting, a P-GENEO (F, F', T) is a triple where the first two operators are equal $F = F'$ (because of Proposition 2.4) and the map T is a homomorphism. Hence, instead of the triple, we can simply write the pair (F, T), that is a GENEO.

Considering two perception triples, we typically want to study the space of all P-GENEOs between them with the map T fixed. Therefore, when the map T is fixed and specified, we will consider simply pairs of operators (F, F') instead of triples (F, F', T), and we say that (F, F') is a P-GENEO *associated with* or *with respect to* the map T.
We denote the set of all P-GENEOs from (Φ, Φ', S) to (Ψ, Ψ', Q) associated with the map $T: S \to Q$ by the symbol

$$\mathcal{F}_T^{all}((\Phi, \Phi', S), (\Psi, \Psi', Q)),$$

or simply $\mathcal{F}_T^{all}$ if the perception triples are clear from the context.

Example 2.7. Let $X = \mathbb{R}^2$. Take a real number $l > 0$. In X consider the square $Q_1 := [0, l] \times [0, l]$, and its translation of a vector $(s_1, s_2) \in \mathbb{R}^2$ $Q_1' := [s_1, l + s_1] \times [s_2, l + s_2]$. Analogously, let us consider a real number $0 < \varepsilon < l$ and two squares inside Q_1 and Q_1', $Q_2 := [\varepsilon, l - \varepsilon] \times [\varepsilon, l - \varepsilon]$ and $Q_2' := [s_1 + \varepsilon, l + s_1 - \varepsilon] \times [s_2 + \varepsilon, l + s_2 - \varepsilon]$, as in Figure 2.1.

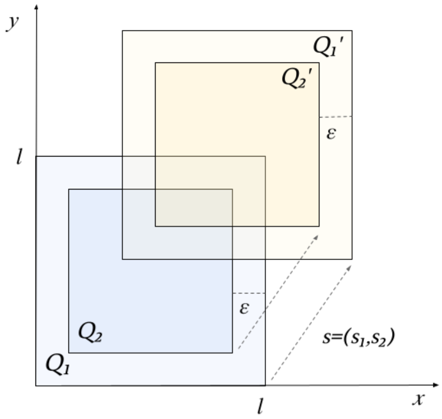

FIGURE 2.1: Square supports of functions.

Consider the following function spaces in $\mathbb{R}_b^X$:

$$\Phi := \{\varphi\colon X \to \mathbb{R} \mid \operatorname{supp}(\varphi) \subseteq Q_1\}$$
$$\Phi' := \{\varphi'\colon X \to \mathbb{R} \mid \operatorname{supp}(\varphi') \subseteq Q_1'\}$$
$$\Psi := \{\psi\colon X \to \mathbb{R} \mid \operatorname{supp}(\psi) \subseteq Q_2\}$$
$$\Psi' := \{\psi'\colon X \to \mathbb{R} \mid \operatorname{supp}(\psi') \subseteq Q_2'\}.$$

Let $S := \{s^{-1}\}$, where s is the translation operator of vector (s_1, s_2). The triples (Φ, Φ', S) and (Ψ, Ψ', S) are perception triples.
This example could model the translation of two nested grey-scale images. We want to build now an operator between these images in order to obtain a transformation that preserves the translation invariance.
We can consider the triple of functions (F, F', T) defined as follows. $F\colon \Phi \to \Psi$ is the operator that maintains the output of functions in Φ at points of Q_2 and set them to zero outside it; analogously $F'\colon \Phi' \to \Psi'$ is the operator that maintains the output of functions in Φ' at points of Q_2' and set them to zero outside it; and $T = \mathrm{id}_S$. Therefore, the triple (F, F', T) is a P-GENEO from (Φ, Φ', S) to (Ψ, Ψ', S). It turns out that the maps are non-expansive and the equivariance holds:

$$F'(\varphi s^{-1}) = F(\varphi)T(s^{-1}) = F(\varphi)s^{-1}$$

for any $\varphi \in \Phi$. From the point of view of application, we are considering two square images and their translations and we apply an operator that 'cuts'

the images, taking into account only the part of the image that interests the observer.
This example justifies the definition of P-GENEO as a triple of operators (F, F', T), without requiring F and F' to be equal in the possibly non-empty intersection of their domains. In fact, if φ is a function contained in $\Phi \cap \Phi'$, its image via F and F' may be different.

2.1 Compactness and convexity of the space of P-GENEOs

Given two perception triples, under specific assumptions about the data sets, it is possible to show two features useful in applications: compactness and convexity. Indeed, from a computational point of view, a compact space can be approximated by a finite set and the convexity guarantees the possibility of creating new P-GENEOs, as convex combination of some preexisting ones.
Before proceeding we need to endow the space of P-GENEOs with a topology.
Let X, Y be sets. Considering two sets $\Omega \subseteq \mathbb{R}_b^X, \Delta \subseteq \mathbb{R}_b^Y$, we can define the distance

$$D_{\mathbf{NE}}^{\Omega}(F_1, F_2) := \sup_{\omega \in \Omega} \|F_1(\omega) - F_2(\omega)\|_\infty$$

for every $F_1, F_2 \in \mathbf{NE}(\Omega, \Delta)$.
Consider the space $\mathcal{F}_T^{all}$ of all the P-GENEOs between two perception triples (Φ, Φ', S) and (Ψ, Ψ', Q) associated with the map T. We define a distance $D_{\text{P-GENEO}}$ on $\mathcal{F}_T^{all}$:

$$D_{\text{P-GENEO}}((F_1, F_1'), (F_2, F_2')) := \max\{D_{\mathbf{NE}}^{\Phi}(F_1, F_2), D_{\mathbf{NE}}^{\Phi'}(F_1', F_2')\}$$

$$= \max\{\sup_{\varphi \in \Phi} \|F_1(\varphi) - F_2(\varphi)\|_\infty, \sup_{\varphi' \in \Phi'} \|F_1'(\varphi') - F_2'(\varphi')\|_\infty\}$$

for every $(F_1, F_1'), (F_2, F_2') \in \mathcal{F}_T^{all}$. Observe that the pseudo-metrics are all well-defined, since by definition $F_1, F_2 \in \mathbf{NE}(\Phi, \Psi)$ and $F_1', F_2' \in \mathbf{NE}(\Phi', \Psi')$.

2.1.1 Compactness

Before proceeding, we recall some definitions:

Definition 2.8. A **relatively compact subset** B of A is a subset whose closure is compact.

Definition 2.9. Let (P, d_P) and (Q, d_Q) be two pseudo-metric spaces. A family $\mathcal{H}$ of functions from P to Q is called **equicontinuous** at a point $p_0 \in P$ if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d_Q(h(p_0), h(p)) < \varepsilon$ for any $h \in \mathcal{H}$ and for any $p \in P$ such that $d_P(p_0, p) < \delta$.
If the family $\mathcal{H}$ is equicontinuous at each point of P, it is called **pointwise equicontinuous** or simply **equicontinuous**.

Definition 2.10. Let (P, d_P) and (Q, d_Q) be two pseudo-metric spaces. A family $\mathcal{H}$ of functions from P to Q is called **uniformly equicontinuous** if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d_Q(h(p_1), h(p_2)) < \varepsilon$ for any $h \in \mathcal{H}$ and for any $p_1, p_2 \in P$ such that $d_P(p_1, p_2) < \delta$.

Note that a uniformly equicontinuous set is equicontinuous at every point. We can now recall a generalization of the Ascoli-Arzelà Theorem for pseudo-metric spaces (see [23]):

Theorem 3 (Generalization of the Ascoli-Arzelà Theorem)**.** *Let P be a compact space and (Q, d) a pseudo-metric space. For $\mathcal{H} \subset C(P, Q)$, the following are equivalent:*

1. *$\mathcal{H}$ is relatively compact in $C(P, Q)$;*
2. *$\mathcal{H}$ is equicontinuous and $K := \{g(p) : g \in \mathcal{H}, p \in P\}$ is relatively compact in (Q, d);*
3. *$\mathcal{H}$ is equicontinuous and $\mathcal{H}(p) := \{g(p) : g \in \mathcal{H}\}$ is relatively compact in (Q, d) for each $p \in P$.*

Corollary 2.11. *If $(P, d_P), (Q, d_Q)$ are compact pseudo-metric spaces, then $\boldsymbol{NE}(P, Q)$ is compact.*

Proof. The non-expansiveness condition directly implies that $\mathbf{NE}(P, Q)$ is uniformly equicontinuous. Indeed, for every $\varepsilon > 0$ we can take $\delta = \varepsilon$ and we have that

$$d_Q(F(p_1), F(p_2)) \leq d_P(p_1, p_2) < \delta = \varepsilon$$

for every $F \in \mathbf{NE}(P, Q)$ and for every $p_1, p_2 \in P$ such that $\|p_1 - p_2\|_\infty < \delta$. Therefore, $\mathbf{NE}(P, Q)$ is also equicontinuous at each point.
Since every subset of a compact have a compact closure and Q is compact, we have that $K := \{F(p) : F \in \mathbf{NE}(P, Q), p \in P\} \subseteq Q$ is relatively compact. Since P is compact, we are in the hypothesis of Theorem 3 considering

$\mathcal{H} = \mathbf{NE}(P, Q) \subset C(P, Q)$ and we have proven that assertion 2 holds. Hence, for assertion 1, $\mathbf{NE}(P, Q)$ is relatively compact. So, in order to prove our statement, it will suffice to show that $\mathbf{NE}(P, Q)$ is closed. Consider a sequence $(F_i)_{i\in\mathbb{N}}$ in $\mathbf{NE}(P, Q)$ such that $\lim_{i\to\infty} F_i = F$. We have that

$$d_Q(F(p_1), F(p_2)) = \lim_{i\to\infty} d_Q(F_i(p_1), F_i(p_2)) \leq d_P(p_1, p_2)$$

for every $p_1, p_2 \in P$. Therefore, $F \in \mathbf{NE}(P, Q)$ and the statement is true. □

Returning to our model, we consider two perception triples (Φ, Φ', S) and (Ψ, Ψ', Q), with domains X and Y, respectively, and the space $\mathcal{F}_T^{all}$ of P-GENEOs between them, associated with the map $T\colon S \to Q$. The following result holds:

Theorem 4. *If Φ, Ψ and Φ', Ψ' are compact, then $\mathcal{F}_T^{all}$ is compact with respect to the pseudo-metric $D_{\text{P}-\text{GENEO}}$.*

Proof. By definition, $\mathcal{F}_T^{all} \subseteq \mathbf{NE}(\Phi, \Psi) \times \mathbf{NE}(\Phi', \Psi')$. Since Φ, Ψ and Φ', Ψ' are compact, for Corollary 2.11 the spaces $\mathbf{NE}(\Phi, \Psi)$ and $\mathbf{NE}(\Phi', \Psi')$ are compact, and then, from Tychonoff's Theorem, the product $\mathbf{NE}(\Phi, \Psi) \times \mathbf{NE}(\Phi', \Psi')$ is also compact, with respect to the product topology. Hence, in order to prove our statement it will suffice to show the closure of $\mathcal{F}_T^{all}$. Let us consider a sequence $((F_i, F_i'))_{i\in\mathbb{N}}$ of P-GENEOs. Since $\mathbf{NE}(\Phi, \Psi) \times \mathbf{NE}(\Phi', \Psi')$ is compact, there is a subsequence that converges to a couple of non-expansive operators $(F, F') \in \mathbf{NE}(\Phi, \Psi) \times \mathbf{NE}(\Phi', \Psi')$. Now, from Proposition 1.30, we have that the actions of the sets S and Q are continuous choosing the pseudo-metrics D_{Aut}^{Φ} on S and D_{Aut}^{Ψ} on Q. Moreover, since (F_i, F_i') is T-equivariant for every $i \in \mathbb{N}$, we have that

$$F'(\varphi s) = \lim_{i\to\infty} F_i'(\varphi s) = \lim_{i\to\infty} F_i(\varphi)T(s) = F(\varphi)T(s)$$

for every $s \in S$ and every $\varphi \in \Phi$. Since equivariance is maintained, the couple (F, F') belongs to $\mathcal{F}_T^{all}$. Hence, $\mathcal{F}_T^{all}$ is a closed subset of a compact set and then it is also compact. □

2.1.2 Convexity

Assume that Ψ, Ψ' are convex. Let $(F_1, F_1'), \ldots, (F_n, F_n') \in \mathcal{F}_T^{all}$ and consider an n-tuple $(a_1, \ldots, a_n) \in \mathbb{R}^n$ with $a_i \geq 0$ for every $i \in \{1, \ldots, n\}$ and

$\sum_{i=1}^n a_i = 1$. We can define two operators $F_\Sigma\colon \Phi \to \Psi$ and $F'_\Sigma\colon \Phi' \to \Psi'$ as

$$F_\Sigma(\varphi) := \sum_{i=1}^n a_i F_i(\varphi), \quad F'_\Sigma(\varphi') := \sum_{i=1}^n a_i F'_i(\varphi')$$

for every $\varphi \in \Phi, \varphi' \in \Phi'$. Observe that the assumption on the convexity for Ψ, Ψ' guarantees that F_Σ and F'_Σ are well defined.

Proposition 2.12. *Under the previous assumptions, (F_Σ, F'_Σ) belongs to $\mathcal{F}_T^{all}$.*

Proof. We have to prove that (F_Σ, F'_Σ) is a P-GEO with respect to T and that is non-expansive. By hypothesis, for every $i \in \{1, \dots, n\}$ (F_i, F'_i) is a perception map, therefore:

$$\begin{aligned} F'_\Sigma(\varphi s) = \sum_{i=1}^n a_i F'_i(\varphi s) &= \sum_{i=1}^n a_i (F_i(\varphi) T(s)) \\ &= \Big(\sum_{i=1}^n a_i F_i(\varphi)\Big) T(s) \\ &= F_\Sigma(\varphi) T(s) \end{aligned}$$

for every $\varphi \in \Phi$ and every $s \in S$. Furthermore, since for every $i \in \{1, \dots, n\}$ $F_i(\Phi) \subseteq \Psi$ and Ψ is convex, also $F_\Sigma(\Phi) \subseteq \Psi$. Analogously, since Ψ' is convex, $F'_\Sigma(\Phi') \subseteq \Psi'$. Therefore (F_Σ, F'_Σ) is a P-GEO. It remains to show the of F_Σ and F'_Σ. By definition every F_i is non-expansive, then for every $\varphi_1, \varphi_2 \in \Phi$ we have that

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

Analogously, since every F'_i is non-expansive, for every $\varphi'_1, \varphi'_2 \in \Phi'$ we have that

$$\|F'_\Sigma(\varphi'_1) - F'_\Sigma(\varphi'_2)\|_\infty \leq \sum_{i=1}^n |a_i| \|\varphi'_1 - \varphi'_2\|_\infty = \|\varphi'_1 - \varphi'_2\|_\infty.$$

Therefore, we have proven that (F_Σ, F'_Σ) is a P-GEO with F_Σ and F'_Σ non-expansive. Hence it is a P-GENEO. □

Then, the following result holds:

Corollary 2.13. *If* Ψ, Ψ' *are convex, then the set* $\mathcal{F}_T^{all}$ *is convex.*

Proof. It is sufficient to apply Proposition 2.12 for $n = 2$, by setting $a_1 = t$, $a_2 = 1 - t$ for $0 \leq t \leq 1$. □

Chapter 3

Methods for building P-GENEOs

In Theorem 4 we have proved that $\mathcal{F}_T^{all}$ is compact if Φ, Ψ and Φ', Ψ' are compact. This is useful from a computational point of view since a compact space can be approximated by a finite set. In this chapter, we will illustrate general methods for building P-GENEOs.

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

Starting from a finite number of P-GENEOs, we will find some methods for building new P-GENEOs.
First of all, we can prove that the composition of two P-GENEOs is still a P-GENEO.

Proposition 3.1. *If we consider two P-GENEOs*

$$(F_1, F_1', T_1) \colon (\Phi, \Phi', S) \to (\Psi, \Psi', Q),$$
$$(F_2, F_2', T_2) \colon (\Psi, \Psi', Q) \to (\Omega, \Omega', K),$$

then the composition

$$(F, F', T) := (F_2 \circ F_1, F_2' \circ F_1', T_2 \circ T_1) \colon (\Phi, \Phi', S) \to (\Omega, \Omega', K)$$

is a P-GENEO.

Proof. We can observe that the map $T = T_1 \circ T_2$ respects the properties required in the definition of P-GENEOs. Therefore, it remains to verify that $F(\Phi) \subseteq \Omega$, $F'(\Phi') \subseteq \Omega'$ and that the properties of equivariance and non-expansiveness are maintained.

1. Since $F_1(\Phi) \subseteq \Psi$ and $F_2(\Psi) \subseteq \Omega$, then we have that

$$F(\Phi) = (F_2 \circ F_1)(\Phi) = F_2(F_1(\Phi)) \subseteq F_2(\Psi) \subseteq \Omega.$$

   Hence $F(\Phi) \subseteq \Omega$. Since $F_1'(\Phi') \subseteq \Psi'$ and $F_2'(\Psi') \subseteq \Omega'$, we can prove with the same steps that $F'(\Phi') \subseteq \Omega'$.

2. Since (F_1, F_1', T_1) and (F_2, F_2', T_2) are equivariant, then (F, F', T) is equivariant; indeed, for every $\varphi \in \Phi$ we have that

$$\begin{aligned} F'(\varphi s) &= (F_2' \circ F_1')(\varphi s) = F_2'(F_1'(\varphi s)) \\ &= F_2'(F_1(\varphi) T_1(s)) = F_2(F_1(\varphi)) T_2(T_1(s)) \\ &= (F_2 \circ F_1)(\varphi)(T_2 \circ T_1)(s) = F(\varphi) T(s). \end{aligned}$$

3. Since F_1 and F_2 are non-expansive, then F is non-expansive; indeed for every $\varphi_1, \varphi_2 \in \Phi$ we have that

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

   Analogously, since F_1' and F_2' are non-expansive, then F' is non-expansive:

$$\|F'(\varphi_1') - F'(\varphi_2')\|_\infty \le \|\varphi_1' - \varphi_2'\|_\infty.$$

Therefore, the statement is proven. □

Now, given a finite number of P-GENEOs with respect to the same fixed map T, we illustrate a general method to build a new operator as a combination of them.

Consider two sets X and Y. Consider a finite set $H_1, \dots, H_n$ of functions from $\Omega \subseteq \mathbb{R}_b^X$ to $\mathbb{R}_b^Y$ and a map $\mathcal{L}\colon \mathbb{R}^n \to \mathbb{R}$, where $\mathbb{R}^n$ is endowed with the norm $\|(x_1, \ldots, x_n)\|_\infty := \max_{1 \le i \le n} |x_i|$. We define $\mathcal{L}^*(H_1, \ldots, H_n)\colon \Omega \to \mathbb{R}_b^Y$ as

$$\mathcal{L}^*(H_1, \ldots, H_n)(\omega) := [\mathcal{L}(H_1(\omega), \ldots, H_n(\omega))],$$

for any $\omega \in \Omega$, where $[\mathcal{L}(H_1(\omega), \ldots, H_n(\omega))]\colon Y \to \mathbb{R}$ is defined by setting

$$[\mathcal{L}(H_1(\omega), \ldots, H_n(\omega))](y) := \mathcal{L}(H_1(\omega)(y), \ldots, H_n(\omega)(y))$$

for any $y \in Y$.

Consider two perception triples (Φ, Φ', S) and (Ψ, Ψ', Q) with domains X and Y, respectively, and a finite set of P-GENEOs $(F_1, F_1'), \ldots (F_n, F_n')$ between them, associated with the map $T\colon S \to Q$. Let $\mathcal{L}\colon \mathbb{R}^n \to \mathbb{R}$ be a non-expansive map. We can consider two functions $\mathcal{L}^*(F_1, \ldots, F_n)\colon \Phi \to \mathbb{R}_b^Y$ and $\mathcal{L}^*(F_1', \ldots, F_n')\colon \Phi' \to \mathbb{R}_b^Y$, defined as previously.

Proposition 3.2. *Consider a finite set of P-GENEOs $(F_1, F_1'), \ldots, (F_n, F_n')$ from (Φ, Φ', S) to (Ψ, Ψ', Q) with respect to $T\colon S \to Q$ and a non-expansive map $\mathcal{L}$ from $\mathbb{R}^n$ to $\mathbb{R}$. If*

$$\mathcal{L}^*(F_1, \ldots, F_n)(\Phi) \subseteq \Psi, \quad \mathcal{L}^*(F_1', \ldots, F_n')(\Phi') \subseteq \Psi',$$

then $(\mathcal{L}^*(F_1, \ldots, F_n), \mathcal{L}^*(F_1', \ldots, F_n'))$ is a P-GENEO from (Φ, Φ', S) to (Ψ, Ψ', Q) with respect to T.

Proof. By hypothesis, $\mathcal{L}^*(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$ and $\mathcal{L}^*(F_1', \ldots, F_n')(\Phi') \subseteq \Psi'$, so we just need to verify the properties of equivariance and non-expansiveness are maintained.

1. Since $(F_1, F_1'), \ldots, (F_n, F_n')$ are T-equivariant, then for any $\varphi \in \Phi$ and any $s \in S$ we have that:

$$\begin{aligned}\mathcal{L}^*(F_1', \ldots, F_n')(\varphi s) &= [\mathcal{L}(F_1'(\varphi s), \ldots, F_n'(\varphi s))] \\ &= [\mathcal{L}(F_1(\varphi)T(s), \ldots, F_n(\varphi)T(s))] \\ &= [\mathcal{L}(F_1(\varphi), \ldots, F_n(\varphi))]T(s) \\ &= \mathcal{L}^*(F_1, \ldots, F_n)(\varphi)T(s).\end{aligned}$$

   Therefore $(\mathcal{L}^*(F_1, \ldots, F_n), \mathcal{L}^*(F_1', \ldots, F_n'))$ is T-equivariant.

2. Since $F_1, \dots, F_n$ and $\mathcal{L}$ are non-expansive, then for any $\varphi_1, \varphi_2 \in \Phi$ we have that:

$$\begin{aligned}
&\|\mathcal{L}^*(F_1, \dots, F_n)(\varphi_1) - \mathcal{L}^*(F_1, \dots, F_n)(\varphi_2)\|_\infty \\
&= \max_{y\in Y} |[\mathcal{L}(F_1(\varphi_1), \dots, F_n(\varphi_1))](y) - [\mathcal{L}(F_1(\varphi_2), \dots, F_n(\varphi_2))](y)| \\
&= \max_{y\in Y} |\mathcal{L}(F_1(\varphi_1)(y), \dots, F_n(\varphi_1)(y)) - \mathcal{L}(F_1(\varphi_2)(y), \dots, F_n(\varphi_2)(y))| \\
&\le \max_{y\in Y} \|(F_1(\varphi_1)(y) - F_1(\varphi_2)(y), \dots, F_n(\varphi_1)(y) - F_n(\varphi_2)(y))\|_\infty \\
&= \max_{y\in Y} \max_{1\le i\le n} |F_i(\varphi_1)(y) - F_i(\varphi_2)(y)| \\
&= \max_{1\le i\le n} \|F_i(\varphi_1) - F_i(\varphi_2)\|_\infty \\
&\le \|\varphi_1 - \varphi_2\|_\infty.
\end{aligned}$$

Hence, $\mathcal{L}^*(F_1, \dots, F_n)$ is non-expansive. Analogously, since $F'_1, \dots, F'_n$ and $\mathcal{L}$ are non-expansive, then $\mathcal{L}^*(F'_1, \dots, F'_n)$ is non-expansive; indeed, following the same steps we have that:

$$\|\mathcal{L}^*(F'_1, \dots, F'_n)(\varphi'_1) - \mathcal{L}^*(F'_1, \dots, F'_n)(\varphi'_2)\|_\infty \le \|\varphi'_1 - \varphi'_2\|_\infty.$$

for any $\varphi'_1, \varphi'_2 \in \Phi'$.

Therefore $(\mathcal{L}^*(F_1, \dots, F_n), \mathcal{L}^*(F'_1, \dots, F'_n))$ is a P-GENEO from (Φ, Φ', S) to (Ψ, Ψ', Q) with respect to T. □

3.1.1 Examples of operators

The above result describes a general method to build new P-GENEOs, starting from a finite number of known P-GENEOs. In the following, we show some examples.

Maximum operator Consider a finite set $H_1, \dots, H_n$ of functions from $\Omega \subseteq \mathbb{R}_b^X$ to $\mathbb{R}_b^Y$ and the function

$$\max(H_1, \dots, H_n)(\omega) := [\max(H_1(\omega), \dots, H_n(\omega))]$$

from Ω to $\mathbb{R}_b^X$, where $[\max(H_1(\omega), \dots, H_n(\omega))]\colon Y \to \mathbb{R}$ is defined by setting

$$[\max(H_1(\omega), \dots, H_n(\omega))](y) := \max\{H_1(\omega)(y), \dots, H_n(\omega)(y))\}$$

for every $y \in Y$.

Remark 3.3. If we call $\mathcal{L}$ the maximum function $\max\colon \mathbb{R}^n \to \mathbb{R}$ that maps $x = (x_1, \ldots, x_n)$ to $\max\{x_1, \ldots, x_n\}$, we can note that the function $\max(H_1, \ldots, H_n)$ that we have just built is exactly the map $\mathcal{L}^*(H_1, \ldots, H_n)$ previously defined.

Consider two perception triples (Φ, Φ', S) and (Ψ, Ψ', Q) with domains X and Y, respectively, and a finite set of P-GENEOs $(F_1, F_1'), \ldots (F_n, F_n')$ between them, associated with the map $T\colon S \to Q$. Consider the functions $\max(F_1, \ldots, F_n)\colon \Phi \to \mathbb{R}_b^Y$ and $\max(F_1', \ldots, F_n')\colon \Phi' \to \mathbb{R}_b^Y$.

Proposition 3.4. *If* $\max(F_1, \ldots, F_n) \subseteq \Psi$ *and* $\max(F_1', \ldots, F_n') \subseteq \Psi'$*, then* $(\max(F_1, \ldots, F_n), \max(F_1', \ldots, F_n'))$ *is a P-GENEO from* (Φ, Φ', S) *to* (Ψ, Ψ', Q) *associated with the map* T*.*

In order to prove the above result, we recall the following Lemma:

Lemma 3.5. *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. We can suppose that $\max\{u_1, \ldots, u_n\} = u_1$, up to permutations of indices. If $\max\{v_1, \ldots, v_n\} = v_1$, the claim trivially follows since it holds that $|u_1 - v_1| \leq \max\{|u_1 - v_1|, \ldots, |u_n - v_n|\}$. It only remains to check the case $\max\{v_1, \ldots, v_n\} = v_i,\ i \neq 1$. We have that

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

On the other hand, we have that

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

Therefore, the statement is true. □

Using the result just shown of Lemma 3.5, we can now give the proof of Proposition 3.4:

Proof. Because of Proposition 3.2, it is sufficient to show that the maximum function $\max\colon \mathbb{R}^n \to \mathbb{R}$ that maps $x = (x_1, \ldots, x_n)$ to $\max\{x_1, \ldots, x_n\}$ is a non-expansive function. Consider two n-tuples $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ of real numbers. Then, using Lemma 3.5, we have 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 Consider two perception triples (Φ, Φ', S) and (Ψ, Ψ', Q) with domains X and Y, respectively. Let (F, F') be a P-GENEO between them, associated with the map $T\colon S \to Q$. Consider the functions $F_b\colon \Phi \to \mathbb{R}_b^Y$ and $F'_b\colon \Phi' \to \mathbb{R}_b^Y$ defined by setting

$$F_b(\varphi) := F(\varphi) - b, \quad F'_b(\varphi') := F'(\varphi') - b$$

Proposition 3.6. *If $F_b(\Phi) \subseteq \Psi$ and $F'_b(\Phi') \subseteq \Psi'$, then (F_b, F'_b) is a P-GENEO from (Φ, Φ', S) to (Ψ, Ψ', Q) with respect to T.*

Proof. The function $S_b\colon \mathbb{R} \to \mathbb{R}$ defined by setting $S_b(x) := x - b$ preserves the distances; indeed, for every $x, y \in \mathbb{R}$ we have that

$$|S_b(x) - S_b(y)| = |x - b - y + b| = |x - y|.$$

Then, because of Proposition 3.2, (F_b, F'_b) is a P-GENEO from (Φ, Φ', S) to (Ψ, Ψ', Q) with respect to T. □

Convex combination operator Consider two perception triples (Φ, Φ', S) and (Ψ, Ψ', Q) with domains X and Y, respectively, and a finite set of P-GENEOs $(F_1, F'_1), \ldots (F_n, F'_n)$ between them, associated with the map $T\colon S \to Q$. Consider an n-tuple $(a_1, \ldots, a_n) \in \mathbb{R}^n$ with $\sum_{i=1}^n |a_i| \leq 1$. We can define two functions $F_\Sigma\colon \Phi \to \mathbb{R}_b^X$ and $F'_\Sigma\colon \Phi' \to \mathbb{R}_b^X$ such that

$$F_\Sigma(\varphi) := \sum_{i=1}^n a_i F_i(\varphi), \quad F'_\Sigma(\varphi') := \sum_{i=1}^n a_i F'_i(\varphi')$$

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

Proposition 3.7. *If $F_\Sigma(\Phi) \subseteq \Psi$ and $F'_\Sigma(\Phi') \subseteq \Psi'$, then (F_Σ, F'_Σ) is a P-GENEO from (Φ, Φ', S) to (Ψ, Ψ', Q) with respect to T.*

Proof. Consider an n-tuple $(a_1, \ldots, a_n) \in \mathbb{R}^n$ with $\sum_{i=1}^n |a_i| \leq 1$. Because of Proposition 3.2, it is sufficient to show that the function $\Sigma\colon \mathbb{R}^n \to \mathbb{R}$ that maps $x = (x_1, \ldots, x_n)$ to $\sum_{i=1}^n a_i x_i$ is a non-expansive function. Consider two n-tuples $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ of real numbers. We have that:

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

Hence, Σ is a non-expansive function. □

Remark 3.8. We can observe that Proposition 2.12 is a special case of Proposition 3.7, where we consider the n-tuple $(a_1, \ldots, a_n) \in \mathbb{R}^n$ with the additional requests that $a_i \geq 0$ for every $i \in \{1, \ldots, n\}$ and $\sum_{i=1}^n a_i = 1$.

Power mean operator Before applying Proposition 3.2, we recall some definitions and properties about power means and p-norms.

Definition 3.9. Consider a real number $p > 0$. The **power mean operator** M_p is the function $M_p\colon \mathbb{R}^n \to \mathbb{R}$ such that

$$M_p(x_1, \ldots, x_n) := \Big(\frac{1}{n} \sum_{i=1}^n |x_i|^p \Big)^{\frac{1}{p}}$$

for any $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Definition 3.10. Consider a real number $p > 0$. The p**-norm** is the function $\| \cdot \|_p\colon \mathbb{R}^n \to \mathbb{R}$ such that

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

for any $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Remark 3.11. It is well known that, for $p \geq 1$, $\| \cdot \|_p$ is a norm. Moreover, if $x \in \mathbb{R}^n$ and $0 < p < q < \infty$ we have

$$\lim_{q \to \infty} \|x\|_q = \|x\|_\infty \quad \text{and} \quad \|x\|_q \leq \|x\|_p \leq n^{\frac{1}{p}-\frac{1}{q}} \|x\|_q.$$

Then, for q tending to infinity, we obtain the following inequality:

$$\|x\|_\infty \leq \|x\|_p \leq n^{\frac{1}{p}} \|x\|_\infty. \tag{3.1}$$

Now, we can consider a finite set of P-GENEOs $(F_1, F_1'), \ldots, (F_n, F_n')$ from (Φ, Φ', S) to (Ψ, Ψ', Q) with respect to $T \colon S \to Q$, where we call Y the domain of the functions in Ψ, Ψ'. Consider also a real number $p > 0$. We can define the operators $M_p(F_1, \ldots, F_n) \colon \Phi \to \mathbb{R}_b^Y$ and $M_p(F_1', \ldots, F_n') \colon \Phi' \to \mathbb{R}_b^Y$ by setting

$$M_p(F_1, \ldots, F_n)(\varphi)(y) := M_p(F_1(\varphi)(y), \ldots, F_n(\varphi)(y))$$
$$M_p(F_1', \ldots, F_n')(\varphi')(y) := M_p(F_1'(\varphi')(y), \ldots, F_n'(\varphi')(y))$$

for every $y \in Y$, every $\varphi \in \Phi$ and every $\varphi' \in \Phi'$.

Proposition 3.12. *Consider $p \geq 1$. If $M_p(F_1, \ldots, F_n)(\Phi) \subseteq \Psi$ and $M_p(F_1', \ldots, F_n')(\Phi') \subseteq \Psi'$, then $(M_p(F_1, \ldots, F_n), M_p(F_1', \ldots, F_n'))$ is a P-GENEO from (Φ, Φ', S) to (Ψ, Ψ', Q) with respect to T.*

Proof. By Proposition 3.2, it will suffice to show that M_p is a non-expansive function. Take two n-tuples $x = (x_1 \ldots, x_n), y = (y_1 \ldots, y_n) \in \mathbb{R}^n$. Since $\| \cdot \|_p$ is a norm, the reverse triangle inequality holds. Hence, using inequality (3.1), we have that:

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

Therefore, for $p \geq 1$, M_p is a non-expansive operator and because of Proposition 3.2 the statement of Proposition 3.12 is true. □

3.2 Series of P-GENEOs

In this section we will describe a method to build P-GENEOs starting from an infinite set of known P-GENEOs and studying series of P-GENEOs. Before proceeding, we recall some results about series of functions (see [27]).

Theorem 5. *Let $(a_k)_{k\in\mathbb{N}}$ be a positive real sequence such that $(a_k)_{k\in\mathbb{N}}$ is decreasing and $\lim_{k\to\infty} a_k = 0$. Let $(g_k)_{k\in\mathbb{N}}$ be a sequence of bounded functions from the topological space X to $\mathbb{C}$. If there exists a real number $M > 0$ such that*

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

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.

Now we recall a second result, which ensures that a uniformly convergent series of continuous functions is a continuous function.

Theorem 6. *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of continuous functions 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 P-GENEOs. Consider two perception triples (Φ, Φ', S) and (Ψ, Ψ', Q) and the space of P-GENEOs $\mathcal{F}_T^{all}$ between them with respect to $T\colon S \to Q$. Assume that the domain Y of Ψ, Ψ' is a compact pseudo-metric space with respect to both the pseudo-metrics D_Y^{Ψ} and $D_Y^{\Psi'}$. Let $(a_k)_{k\in\mathbb{N}}$ be a positive real sequence such that $(a_k)_{k\in\mathbb{N}}$ is decreasing, $\lim_{k\to\infty} a_k = 0$ and $\sum_{k=1}^{\infty} a_k \leq 1$. Suppose that a sequence $((F_k, F'_k))_{k\in\mathbb{N}} \in \mathcal{F}_T^{all}$ is given and that for any $\varphi \in \Phi, \varphi' \in \Phi'$, there exist $M(\varphi) > 0$ and $M'(\varphi') > 0$ such that

$$\left| \sum_{k=1}^{n} F_k(\varphi)(y) \right| \leq M(\varphi), \quad \left| \sum_{k=1}^{n} F'_k(\varphi')(y) \right| \leq M'(\varphi')$$

for every $y \in Y$ and every $n \in \mathbb{N}$. Hence, the hypothesis of Theorems 5 and 6 are satisfied and then the following operators are well-defined; we consider $F\colon \Phi \to C_b^0(Y, \mathbb{R})$ and $F'\colon \Phi' \to C_b^0(Y, \mathbb{R})$ defined by setting

$$F(\varphi) := \sum_{k=1}^{\infty} a_k F_k(\varphi), \quad F'(\varphi') := \sum_{k=1}^{\infty} a_k F'_k(\varphi'),$$

Proposition 3.13. *If $F(\Phi) \subseteq \Psi$ and $F(\Phi') \subseteq \Psi'$, then $(F, F') \in \mathcal{F}_T^{all}$.*

Proof. Consider $s \in S$. We recall that $T(s)$ is uniformly continuous from $Y_{\Psi'}$ to Y_Ψ, since Proposition 1.19 guarantees that every element of $\mathrm{Aut}_{\Psi,\Psi'}(Y)$ is non-expansive from $Y_{\Psi'}$ to Y_Ψ. Moreover, (F_k, F'_k) is T-equivariant and then we have that:

$$\begin{aligned} F'(\varphi s) &= \sum_{k=1}^{\infty} a_k F'_k(\varphi s) \\ &= \sum_{k=1}^{\infty} a_k F_k(\varphi) T(s) \\ &= \left(\sum_{k=1}^{\infty} a_k F_k(\varphi) \right) T(s) \\ &= F(\varphi) T(s) \end{aligned}$$

for any $\varphi \in \Phi$. Since F_k is non-expansive for every $k \in \mathbb{N}$ 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 $\varphi_1, \varphi_2 \in \Phi$. Analogously, since F'_k is non-expansive for every $k \in \mathbb{N}$ and $\sum_{k=1}^{\infty} a_k \leq 1$, F' is non-expansive:

$$\|F'(\varphi'_1) - F'(\varphi'_2)\|_\infty \leq \|\varphi'_1 - \varphi'_2\|_\infty$$

for any $\varphi'_1, \varphi'_2 \in \Phi'$. □

Conclusions

Since real-world data rarely follows strict mathematical symmetries due to noisy or incomplete data or to symmetry breaking features, in some cases it is desirable to consider only some transformations of the data set, ignoring others. The theory of GENEOs studies applications in which the set containing all the symmetries, for which the equivariance holds, is represented by a group. In this thesis we have proposed a generalization of some known results in the theory of GENEOs to a new mathematical framework, where the collection of all symmetries is represented by a subset of a transformation group. Therefore, we have introduced the concept of P-GENEO, replacing the group with a subset having a weaker structure. In particular, we have shown that P-GENEOs are an extension of GENEOs since a GENEO can be represented as a special case of P-GENEO.

In our mathematical setting, we have defined data sets and the set of admissible transformations. We have then defined topological structures, by defining pseudo-metrics, on the space of data and consequently on the space of P-GENEOs. Under the assumption that the function spaces are compact and convex, we have obtained the compactness and the convexity of the space of P-GENEOs. The compactness guarantees that any operator can always be approximated by a finite number of operators belonging to the same space, while the convexity allows to take convex combinations of P-GENEOs. Compactness and convexity together ensure that every strictly convex loss function on the space of P-GENEOs admits one and only one global minimum. In order to have more tools to analyze the space of P-GENEOs we also have presented some methods to build P-GENEOs starting from a finite or an infinite set of known P-GENEOs.

However, in the theory of GENEOs there are other results that could be generalized to our new mathematical framework, e.g., [25], [3], [1], [8], [11],and [15]. For example, persistent homology has been investigated as an impor-

tant tool to find a fast comparison of functions and then a fast comparison of GENEOs. Moreover, there is another method to define new GENEOs, by means of permutant measures. This technique is based on the use of a symmetric weighted average to build new GENEOs. Furthermore, a Riemannian structure on manifolds of GENEOs has been defined, allowing the use of gradient methods to find optimal operators in such manifolds. It would be interesting to try to generalize these results in the theory of P-GENEOs.
In addition, perhaps the concept of P-GENEO could be weakened further to give even more freedom of choice to the observer, depending on the application considered. So far, admissible transformations have been represented by bijections on the space X, i.e., subsets of $\mathrm{Aut}(X)$. But one could consider the more general space of all the functions (or relations) between elements of X and try to investigate how the topological structure varies and which results continue to hold. All this could form the subject of further research.

Bibliography

[1] F. Ahmad, *Compactification of perception pairs and spaces of Group Equivariant non-Expansive Operators*, arXiv preprint arXiv:2210.04043, (2022).

[2] M. Bakator and D. Radosav, *Deep learning and medical diagnosis: A review of literature*, Multimodal Technologies and Interaction, 2 (2018), p. 47.

[3] 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), pp. 423–433.

[4] G. Bocchi, S. Botteghi, M. Brasini, P. Frosini, and N. Quercioli, *On the finite representation of linear group equivariant operators via permutant measures*, Annals of Mathematics and Artificial Intelligence, (2023), pp. 1–23.

[5] G. Bocchi, P. Frosini, A. Micheletti, A. Pedretti, C. Gratteri, F. Lunghini, A. R. Beccari, and C. Talarico, *GENEOnet: A new machine learning paradigm based on Group Equivariant Non-Expansive Operators. An application to protein pocket detection*, arXiv preprint arXiv:2202.00451, (2022).

[6] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, *Geometric Deep Learning: Grids, groups, graphs, geodesics, and gauges*, arXiv preprint arXiv:2104.13478, (2021).

[7] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, *Geometric Deep Learning: Going beyond Euclidean data*, IEEE Signal Processing Magazine, 34 (2017), pp. 18–42.

[8] F. Camporesi, P. Frosini, and N. Quercioli, *On a new method to build group equivariant operators by means of permutants*, in Machine Learning and Knowledge Extraction: Second IFIP TC 5, TC 8/WG 8.4, 8.9, TC 12/WG 12.9 International Cross-Domain Conference, CD-MAKE 2018, Hamburg, Germany, August 27–30, 2018, Proceedings, Springer, 2018, pp. 265–272.

[9] Z. Cang, L. Mu, and G.-W. Wei, *Representability of algebraic topology for biomolecules in machine learning based scoring and virtual screening*, PLoS Computational Biology, 14 (2018), p. 1005929.

[10] G. Carlsson, *Topology and data*, Bulletin of the American Mathematical Society, 46 (2009), pp. 255–308.

[11] 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).

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

[13] F. Chazal and B. Michel, *An introduction to Topological Data Analysis: Fundamental and practical aspects for data scientists*, Frontiers in Artificial Intelligence, 4 (2021), p. 667963.

[14] T. Cohen and M. Welling, *Group equivariant convolutional networks*, in International Conference on Machine Learning, PMLR, 2016, pp. 2990–2999.

[15] F. Conti, P. Frosini, and N. Quercioli, *On the construction of group equivariant non-expansive operators via permutants and symmetric functions*, Frontiers in Artificial Intelligence, 5 (2022), p. 16.

[16] H. Edelsbrunner and J. L. Harer, *Computational Topology: an Introduction*, American Mathematical Society, 2022.

[17] L. Fan, T. Zhang, X. Zhao, H. Wang, and M. Zheng, *Deep topology network: A framework based on feedback adjustment learning rate for image classification*, Advanced Engineering Informatics, 42 (2019), p. 100935.

[18] M. Finzi, G. Benton, and A. G. Wilson, *Residual pathway priors for soft equivariance constraints*, in Advances in Neural Information Processing Systems, M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, eds., vol. 34, Curran Associates, Inc., 2021, pp. 30037–30049.

[19] S. Gaal, *Point Set Topology*, ISSN, Elsevier Science, 1964.

[20] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, *Deep learning for visual understanding: A review*, Neurocomputing, 187 (2016), pp. 27–48.

[21] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, *Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups*, IEEE Signal Processing Magazine, 29 (2012), pp. 82–97.

[22] Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nature, 521 (2015), pp. 436–444.

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

[24] J. Masci, E. Rodolà, D. Boscaini, M. M. Bronstein, and H. Li, *Geometric Deep Learning*, in SIGGRAPH ASIA 2016 Courses, Association for Computing Machinery, New York, NY, United States, 2016, pp. 1–50.

[25] N. Quercioli, *On the topological theory of Group Equivariant Non-Expansive Operators*, PhD thesis, Alma Mater Studiorum - Università di Bologna, Maggio 2021.

[26] D. W. Romero and S. Lohit, *Learning equivariances and partial equivariances from data*, CoRR, abs/2110.10211 (2021).

[27] W. Rudin et al., *Principles of Mathematical Analysis*, vol. 3, McGraw-hill New York, 1976.

[28] T. F. van der Ouderaa, D. W. Romero, and M. van der Wilk, *Relaxing equivariance constraints with non-stationary continuous filters*, arXiv preprint arXiv:2204.07178, (2022).

[29] R. Wang, R. Walters, and R. Yu, *Approximately equivariant networks for imperfectly symmetric dynamics*, CoRR, abs/2201.11969 (2022).

[30] M. Weiler and G. Cesa, *General E(2)-equivariant steerable CNNs*, in Advances in Neural Information Processing Systems, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, eds., vol. 32, Curran Associates, Inc., 2019.

[31] W. Wu, Z. Qi, and L. Fuxin, *Pointconv: Deep convolutional networks on 3D point clouds*, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 9621–9630.