Equivariant symmetry breaking sets

by

YuQing Xie

B.S., Massachusetts Institute of Technology (2022)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2024

Authored by: YuQing Xie
Department of Electrical Engineering and Computer Science
January 26, 2024

Certified by: Tess Smidt
Assistant Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by: Leslie A. Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Equivariant symmetry breaking sets

by

YuQing Xie

Submitted to the Department of Electrical Engineering and Computer Science on January 26, 2024 in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

ABSTRACT

Equivariant neural networks (ENNs) have been shown to be extremely useful in many applications involving some underlying symmetries. However, equivariant networks are unable to produce lower symmetry outputs given a high symmetry input. Spontaneous symmetry breaking occurs in many physical systems where we have a less symmetric stable state from an initial highly symmetric one. Hence, it is imperative that we understand how to systematically break symmetry for equivariant neural networks. In this work, we propose the first symmetry breaking framework that is fully equivariant. Our approach is general and applicable to equivariance under any group. To achieve this, we introduce the idea of symmetry breaking sets (SBS). Rather than redesign existing networks to output symmetrically degenerate sets, we design sets of symmetry breaking objects which we feed into our network based on the symmetry of our input. We show there is a natural way to define equivariance on these sets which gives an additional constraint. Minimizing the size of these sets equates to data efficiency. We show that bounding the size of these sets translates to the well studied group theory problem of finding complements of normal subgroups. We tabulate solutions to this problem for the point groups. Finally, we provide some examples of symmetry breaking to demonstrate how our approach works in practice.

Thesis supervisor: Tess Smidt
Title: Assistant Professor of Electrical Engineering and Computer Science

Acknowledgments

I am extremely grateful to my advisor Tess Smidt for her guidance and expertise. Your support and insightful feedback were invaluable throughout the research process. In addition, her genuine concern for our personal and academic development has created a supportive environment that goes beyond the scope of research.

I would also like to thank my lab-mates and fellow members of the Atomic Architects for their insightful discussions and camaraderie. I am grateful for the countless hours spent together, whether in lab, in the climbing gym, playing chess, playing soccer, running, or elsewhere. You have truly made this a fruitful and enjoyable experience.

I thank the MIT College of Computing fellowship and the National Science Foundation Graduate Research Fellowship under Grant No. (DGE-1745302) for funding this research.

Finally, I would like to thank my family and friends for their unwavering support throughout the years. Without you, this thesis would not be possible.

Contents

Title page 1

Abstract 2

Acknowledgments 3

List of Figures 7

List of Tables 9

1 Introduction 11
- 1.1 Related works . . . 12
- 1.2 Our contribution . . . 12

2 Background 14
- 2.1 Group theory . . . 14
  - 2.1.1 Concepts for understanding equivariant neural networks . . . 15
  - 2.1.2 Concepts for understanding our symmetry breaking scheme . . . 16
- 2.2 Equivariant neural networks . . . 18
- 2.3 Symmetry breaking problem . . . 19

3 Fully broken symmetry 22
- 3.1 Symmetry breaking set (SBS) . . . 22
- 3.2 Equivariant SBS . . . 23
- 3.3 Ideal case and complement of normal subgroups . . . 26
- 3.4 Nonideal equivariant SBSs . . . 27

4 Partially broken symmetry 30
- 4.1 Partial SBS . . . 30
- 4.2 Equivariant partial SBS . . . 31
- 4.3 Ideal equivariant partial SBS . . . 33
- 4.4 Nonideal equivariant partial SBS . . . 35
- 4.5 Optimality of exact partial symmetry breaking . . . 37

5 Experiments **38**
5.1 Full symmetry breaking: chiral triangular prism . . . 38
5.1.1 Obtaining an equivariant SBS . . . 39
5.1.2 Training results . . . 40
5.2 Partial symmetry breaking: octagon to rectangle . . . 41
5.2.1 Obtaining an equivariant partial SBS . . . 42
5.2.2 Training results . . . 43
5.3 $BaTiO_3$ phase transitions . . . 45

6 Conclusion **48**

Bibliography **49**

A Notation and commonly used symbols **52**

B Classification of full symmetry breaking cases for $O(3)$ **53**
B.1 Normalizer: K_h . . . 54
B.2 Normalizer: $D_{\infty h}$ . . . 54
B.3 Normalizer: $D_{(2n)h}$ . . . 54
B.4 Normalizer: I_h . . . 55
B.5 Normalizer: O_h . . . 56

C Equivariant full SBS better than exact partial SBS **57**

List of Figures

3.1 (a) Naive way to break symmetry in a triangular prism where one vector points to a vertice of a triangle and a second vector points to the lower or outer triangle. (b) A rotated version of the triangular prism in (a). Note that the same symmetry breaking objects now point to edges of the triangle rather than vertices. However, both prisms have the exact same symmetry elements. 23
3.2 Diagram of how we might structure our symmetry breaking scheme. From our data x, we may obtain its symmetry S. This S is then fed into a function σ which gives us the set of symmetry breaking objects needed. We sample a b from this set breaking the symmetry of our input and feed this b along with the input x into our equivariant function f. Finally we obtain an output y which has lower symmetry than the input x. 24
3.3 Diagram of how we break symmetry, but now we keep all possible outputs. . 25
3.4 Diagram of what happens when we act on the input with some group element g. 25

4.1 Diagram of how we perform partial symmetry breaking. Here, we need to specify not just the symmetry of our input but also the symmetries of our output. Since any of our outputs are equally valid, it only makes sense to specify the set of conjugate subgroups $\mathrm{Cl}_S(K)$ our outputs are symmetric under. 31
4.2 Diagram of how our symmetry scheme changes when we transform our input by some group element $g \in G$. 32

5.1 (a) Triangular prism with D_3 symmetry and patterned cylinder with D_{6h} symmetry. The generators are a, b, m where a is a $2\pi/6$ rotation about the blue axis, b is a π rotation about the red axis, and m is a reflection across the orange plane. (b) An ideal symmetry breaking set for the triangular prism. A complement of D_3 in D_{6h} is generated by the mirror planes shown here in yellow and purple. The vector in red is a symmetry breaking object with this complement as stabilizer. The orbit of this vector under the normalizer generates the other vectors shown in black. 39
5.2 (a) Output (red) generated by our model and symmetry breaking object (blue) given. (b) The set of all the outputs generated by our model if we feed in all symmetry breaking objects. 40

5.3 (a) Octagon with D_8 symmetry we input and patterned cylinder with $N_G(S, K) = D_{8h}$ symmetry. The generators are a^2, b, m where a^2 is a $2\pi/8$ rotation about the blue axis, b is a π rotation about the red axis, and m is a reflection across the orange plane. (b) A symmetry breaking object which generates an ideal partial SBS. Note that in addition to the D_2 symmetry, this object is also symmetric under reflections across the purple and yellow planes. 42
5.4 (a) Output (blue) of our model when we match a symmetry breaking object with a compatible rectangle. (b) Output (blue) of our model when we match a symmetry breaking object with a different compatible rectangle. (c) Output (blue) when we attempt to match the symmetry breaking object with an incompatible rectangle (green). Note the square has symmetries of both the symmetry breaking object and the target rectangle. 44
5.5 (a) Initial high symmetry crystal structure of $BaTiO_3$. (b) Target low symmetry crystal structure of $BaTiO_3$. 45
5.6 Distortions of a highly symmetric crystal structure of $BaTiO_3$ when provided with each of the possible symmetry breaking objects in our ideal equivariant partial SBS. 47

B.1 Object with symmetryD_{10h}. We can identify generator a as the 10-fold rotation about the blue axis, generator b as the 2-fold rotation about the red axis, and m as the reflection over the plane shown in orange. 55

List of Tables

A.1 Notation used throughout this paper 52
A.2 Commonly used symbols 52

B.1 Normalizers of the point groups in Schönflies notation. Note we have the equivalences $C_1 = 1$, $S_2 = C_i$, $C_{1h} = C_{1v} = C_s$, $D_1 = C_2$, $D_{1h} = C_{2v}$, $D_{1d} = C_{2h}$. 53
B.2 Groups with normalizer $K_h = O(3)$ and their complements. 54
B.3 Groups with normalizer $D_{(2n)h}$ and their complements. 55
B.4 Groups with normalizer I_h and their complements. 56
B.5 Groups with normalizer O_h and their complements. 56

Chapter 1

Introduction

Equivariant neural networks have emerged in recent years as an extremely promising class of models, especially for geometric and scientific data. In particular, such data often has well known underlying symmetries. For example, the coordinates of a molecule may be different under rotations and translations, but the molecule remains the same. Traditional neural networks must see many rotated examples to learn this symmetry. In contrast, equivariant networks already incorporate such symmetries and can focus on learning the underlying physics. Such models have been extremely successful in a large number of applications [1]–[5].

As a consequence of the built in symmetries of equivariant networks, they are unable to output a lower symmetry object given a higher symmetry input. This phenomena was first pointed out in Smidt et. al. [6]. However, many physical systems exhibit spontaneous symmetry breaking, from octahedral tilting of perovskites to the Higgs mechanism for giving particles mass [7], [8]. In such situations, we may have some high symmetry input such as an initial configuration or a Hamiltonian and require a lower symmetry output such as a final configuration or ground state. We emphasize that while individual outputs may break symmetry, the distribution of allowed outputs remains symmetric. Since the fundamental physics is symmetric, we would still like to apply equivariant networks, but our networks will only output an average.

1.1 Related works

One approach to this problem is learning to break symmetry. This is especially useful when there is a one lower symmetry output which is preferred over the others. For example, [6] showed that the gradients of the loss function can be used to learn a symmetry breaking order parameter. Another related approach is approximate and relaxed equivariant networks [9]–[12]. These networks have architectures similar to those of equivariant ones, but are not constrained to remain equivariant. Hence, they can learn the amount of symmetry they need to preserve to achieve the desired output. However, this only works when there is systematic symmetry breaking in the data. If all symmetrically related lower symmetry outputs are equally likely and reflected in the data, then they will still fail. Further, since equivariance is broken, there is no guarantee the method will work when shown data in a different orientation.

Another attempt explicitly designs a planar symmetry detection algorithm for data with mirror plane symmetry [13]. However, the methods in that paper are limiting and only works for that specific type of symmetry. Crystals for example often have much higher symmetries. Finally, Kaba and Ravanbakhsh [14] give a different definition of relaxed equivariance. They derive modified constraints linear layers would need to satisfy for this relaxed equivariance and argue such models would solve the symmetry breaking problem. However, they also mention these conditions do not reduce as easily as for the usual equivariant linear layers.

1.2 Our contribution

In this work, we propose a novel solution to the symmetry breaking problem. Our approach is similar to Smidt et. al. [6] in that we provide symmetry breaking parameters as input to the model. However, rather than learning these parameters, we show that we can sample them from a symmetry breaking set (SBS) that we design based only on the input and output symmetries. In particular, we prove bounds on the size of these SBSs, which characterizes how difficult symmetry breaking is.

Compared to existing methods, our approach has the following advantages:

1. **Equivariance:** Our framework guarantees equivariance. That we can achieve this is the key point of this work.

2. **Simple to implement:** Our approach only requires a designing a set of additional inputs into an equivariant network.

3. **Generalizability:** We emphasize that our characterization of SBSs applies to networks which are equivariant under any groups.

The rest of this thesis is organized as follows. In section 3, we first examine the easier case where we break all symmetries of our input data. We motivate the idea of a symmetry breaking set and show that imposing equivariance leads to an additional nontrivial constraint of closure under the normalizer group. The intuition is that the normalizer group characterizes all orientations of our data which do not change its symmetry. Next, we translate bounds on the size of the equivariant SBSs into the purely group theoretical problem of finding complements of normal subgroups. We have tabulated these complements in Appendix B for the point groups. In section 4, we generalize to the case where we may still share some symmetries with our input. Finally, in section 5, we introduce examples of symmetry breaking and demonstrate how our method works in practice.

Chapter 2

Background

2.1 Group theory

Group theory is the mathematical language used to describe symmetries. Here, we present a brief overview of concepts from group theory we need to both define equivariance, and to understand our proposed symmetry breaking scheme. For a more comprehensive treatment of group theory, we refer to standard textbooks [15]–[17]. We begin by defining what a group is.

Definition 2.1 (Group)**.** *Let G be a nonempty set equipped with a binary operator $\cdot : G \times G \to G$. This is a group if the follwing group axioms are satsfied*

1. *Associativity: For all $a, b, c \in G$, we have $(a \cdot b) \cdot c = a \cdot (b \cdot c)$*
2. *Identity element: There is an element $e \in G$ such that for all $g \in G$ we have $e \cdot g = g \cdot e = g$*
3. *Inverse element: For all $g \in G$, there is an inverse $g^{-1} \in G$ such that $g \cdot g^{-1} = g^{-1} \cdot g = e$ for identity e.*

Some examples of groups include the group of rotation matrices with matrix multiplication as the group operation, the group of integers under addition, and the group of positive reals under multiplication. One very important group is the group of automorphisms on a vector space. This group is denoted $GL(V)$ and we can think of it as the group of invertible matrices.

2.1.1 Concepts for understanding equivariant neural networks

While abstractly groups are interesting on their own, we care about using them to describe symmetries. Intuitively, the group elements abstractly represent the symmetry operations. In order to understand what these actions are, we need to define a group action.

Definition 2.2 (Group action)**.** *Let G be a group and Ω a set. A group action is a function $\alpha : G \times \Omega \to \Omega$ such that $\alpha(e, x) = x$ and $\alpha(g, \alpha(h, x)) = \alpha(gh, x)$ for all $g, h \in G$ and $x \in \Omega$.*

Often, we may want to relate two groups to each other. This is done using group homomorphisms, a mapping which preserves the group structure.

Definition 2.3 (Group homomorphism and isomorphism)**.** *Let G and H be groups. A group homomorphism is a function $f : G \to H$ such that $f(u \cdot v) = f(u) \cdot f(v)$ for all $u, v \in G$. A group homomorphism is an isomorphism if f is a bijection.*

Because there are many linear algebra tools for working with matrices, it is particular useful to relate arbitrary groups to groups consisting of matrices. Such a homomorphism together with the vector space the matrices act on is a group representation.

Definition 2.4 (Group representation)**.** *Let G be a group and V a vector space over a field F. A group representation is a homomorphism $\rho : G \to GL(V)$ taking elements of G to autmorphisms of V.*

Given any representation, there are often orthogonal subspaces which do not interact with each other. If this is the case, we can break our representation down into smaller pieces by restricting to these subspaces. Hence, it is useful to consider the representations which cannot be broken down. These are known as the irreducible representations (irreps) and often form the building blocks of more complex representations.

Definition 2.5 (Irreducible representation)**.** *Let G be a group, V a vector space, and $\rho : G \to GL(V)$ a representation. A representation is irreducible if there is no nontrivial proper subspace $W \subset V$ such that $\rho|_W$ is a representation of G over space W.*

There has been much work on understanding the irreps of various groups and many equivariant neural network designs use this knowledge.

2.1.2 Concepts for understanding our symmetry breaking scheme

One natural question is whether there is a subset of group elements which themselves form a group under the same group operation. Such a subset is a called a subgroup.

Definition 2.6 (Subgroup)**.** *Let G be a group and $S \subseteq G$. If S together with the group operation of G $\cdot$ satisfy the group axioms, then S is a subgroup of G which we denote as $S \leq G$.*

One particular feature of a subgroup is that we can use them to decompose our group into disjoint chunks called cosets.

Definition 2.7 (Cosets)**.** *Let G be a group and S a subgroup. The left cosets are sets obtained by multiplying S with some fixed element of G on the left. That is, the left cosets are for all $g \in G$*

$$gS = \{gs : s \in S\}.$$

We denote the set of left cosets as G/S. The right cosets are defined similarly except we multiply with a fixed element of G on the right. That is, the right cosets are for all $g \in G$

$$Sg = \{sg : s \in S\}.$$

We denote the set of right cosets as $G\backslash S$.

In general, the left and right cosets are not the same. However, for some subgroups they are the same. Those subgroups are called normal subgroups.

Definition 2.8 (Normal subgroup). *Let G be a group and N a subgroup. Then N is a normal subgroup if for all $g \in G$, we have $gNg^{-1} = N$.*

It turns out that given a normal subgroup, one can construct a group operation on the cosets. The resulting group is called a quotient group.

Definition 2.9 (Quotient group). *Let G be a group and N a normal subgroup. One can define a group operation on the cosets as $aN \cdot bN = (a \cdot b)N$. The resulting group is called the quotient group and is denoted G/N.*

For subgroups S which are not normal in G, it is often useful to consider a subgroup of G containing S where S is in fact normal. The largest such subgroup is called the normalizer.

Definition 2.10 (Normalizer). *Let G be a group and S a subgroup. The normalizer of S in G is*

$$N_G(S) = \{g : gSg^{-1} = S\}.$$

Similar to orthogonal vector spaces, one can imagine an analogous notion for groups. These are called complement subgroups.

Definition 2.11 (Complement). *Let G be a group and S a subgroup. A subgroup H is a complement of S if for all $g \in G$, we have $g = sh$ for some $s \in S$ and $h \in H$ and $S \cap H = \{e\}$.*

It turns out that if S is a normal subgroup of G and H is a complement, then H is isomorphic to the quotient group.

Finally, it is useful to define what we mean by symmetry of an object. These are all group elements which leave the object unchanged and is called the stabilizer.

Definition 2.12 (Stabilizer). *Let G be a group, Ω some set with an action of G defined on it, and $u \in \Omega$. The stabilizer of u is all elements of G which leave u invariant. That is*

$$\mathrm{Stab}_G(u) = \{g : gu = u, g \in G\}.$$

One can check that the stabilizer is indeed a subgroup. Closely related to the stabilizer is the orbit. This is all the values we get when we act with our group on some object.

Definition 2.13 (Orbit)**.** *Let G be a group, Ω some set with an action of G defined on it, and $u \in \Omega$. The orbit of u is the set of all values obtained when we act with all elements of G on it. That is,*

$$\mathrm{Orb}_G(u) = \{gu : g \in G\} = Gu.$$

It turns out one can show that the stabilizer of elements in the orbit are related. This relation turns out to be conjugation which we define below.

Definition 2.14 (Conjugate subgroups)**.** *Let S and S' be subgroups of G. We say S and S' are conjugate in G if there is some $g \in G$ such that $S = gS'g^{-1}$. We denote the set of all conjugate subgroups by*

$$\mathrm{Cl}_G(S) = \{gSg^{-1} : g \in G\}.$$

2.2 Equivariant neural networks

We are now ready to formally describe how equivariant neural networks work. Suppose our neural network maps inputs in some space X to some space Y. Now, suppose we have a group G and that there are group actions of G defined on both the input space X and output space Y. Then we can define equivariance as the following.

Definition 2.15 (Equivariance)**.** *Suppose we have spaces X and Y with group actions of G defined on them. A function $f : X \to Y$ is said to be equivariant if for all $g \in G, x \in X, y \in Y$, we have*

$$f(gx) = gf(x).$$

Intuitively, we can interpret this as rotating the input giving the same result as just rotating the output. It is easy to check that the composition of equivariant functions is still an equivariant function. Hence, equivariant neural networks are designed using a composition of equivariant layers.

There has been considerable study into how one should design equivariant layers. One approach is to modify convolutional filters by transforming them with the elements of our group [18]. This approach is known as group convolution and is based on the intuition that convolutional filters are translation equivariant. In group convolution, one interprets our data as a signal over some domain. The first layer is a lifting convolution which transforms our data into a signal over the group. The remaining layers then just convolve this signal with filters which are also signals over the group. This approach works well for finite groups where we can represent signals using just a finite set of values.

One can further use group theory tools to break down the convolutional filters into irreps. This leads to steerable convolutional networks [19]. These can be extended and used to parameterize continuous filters which can be used for infinite groups [20]. It turns out the irreps of the group are natural data types for equivariant networks. Further, we can express the convolutions as tensor products of irreps. We can think of equivariant operations as being composed of tensor products of irreps, linear mixing of irreps, and scaling by invariant quantities. Combining these, we get tensorfield networks which works on point clouds and is rotation equivariant [21]. For this project, we demonstrate our method using networks built from the `e3nn` framework for $O(3)$ equivariance [22].

2.3 Symmetry breaking problem

Consider the following simple task. Given a pair of nodes, color exactly one of them. Suppose we attempt to solve this problem with a permutation equivariant network and show it an example where the left node gets colored. The equivariant network will realize that swapping the nodes gives the same input graph. However, swapping the outputs gives an equally valid solution where the right node is colored. Hence, rather than giving exactly one node colored as output, it will output a graph where both nodes are half colored.

The reason this happens is because the output has lower symmetry than that of the input. For the input, swapping the two nodes changes nothing, however swapping the nodes for the valid outputs swaps between left and right nodes being colored. More formally, we can describe the symmetry breaking problem with the following lemma.

Lemma 2.16. *Let X be a set with a transitive group action by G defined on it. Let Y be some set with a group action of G defined on it. Let $f : X \to Y$ be an G-equivariant function. We can choose f such that $f(u) = y$ if and only if $\mathrm{Stab}_G(y) \geq \mathrm{Stab}_G(u)$. Further this f is unique.*

Proof. First suppose we did have $f(u) = y$. For any $g \in \mathrm{Stab}_G(u)$, we have by equivariance of f that

$$gy = f(gu) = f(u) = y.$$

So $g \in \mathrm{Stab}_G(y)$.

Next, suppose $\mathrm{Stab}_G(y) \geq \mathrm{Stab}_G(u)$. For any $x \in X$, there is some $r \in G$ so that $x = ru$. Let us pick exactly one such r for each X and form a set R. Hence any x is uniquely written as $x = ru$ for $r \in R$. Define

$$f(x) = f(ru) = ry.$$

We claim f is equivariant. For any $g \in G$ and $x \in X$, let $x = ru$ and $gx = r'u$ for some $r, r' \in R$. Then,

$$f(gx) = f(gru) = f(r'u) = r'y.$$

But note that $gx = r'u$ implies $r'^{-1}gx = r'^{-1}gru = u$. So $r'^{-1}gr \in \mathrm{Stab}_G(u) \leq \mathrm{Stab}_G(y)$. Hence, we also have $r'^{-1}gry = y$. So,

$$f(gx) = r'y = r'(r'^{-1}gry) = gry = gf(x).$$

Hence, f is equivariant.

Finally, for uniqueness, suppose f, f' are two equivariant functions such that $f(u) = f'(u) = y$. Then by equivariance, for any $x = gu \in X$ we have

$$f(x) = f(gu) = gy = f'(gu) = f'(x).$$

□

The interesting part of this lemma is that there is a converse. When the output symmetry is equal or higher than the input symmetry, we can construct an equivariant function. This converse turns out to be crucial for understanding why the approach we present works.

Chapter 3

Fully broken symmetry

First, we consider the case where we want to break all symmetry of our input data. Here, our desired outputs y share no symmetry with x. In other words $\text{Stab}_S(y) = \{e\}$. This will form the foundation for our analysis of the general case of partially broken symmetry. Let the symmetry group of our data x be S.

3.1 Symmetry breaking set (SBS)

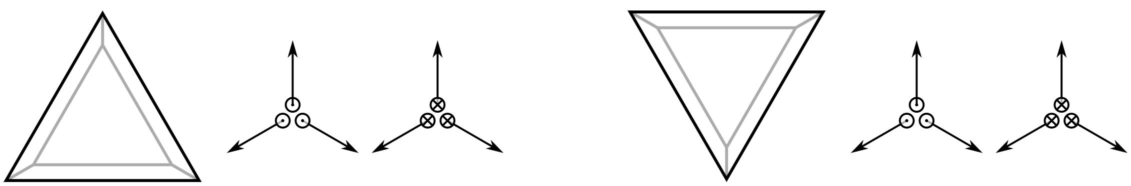
When there is symmetry breaking there are multiple equally valid symmetrically related outputs. The purpose of a symmetry breaking object is to allow an equivariant network to pick one of them. In principle, we want all symmetrically related outputs to be equally likely so it makes sense to think of a set B of symmetry breaking objects we sample from rather than a single object.

For any $s \in S$ and $b \in B$, since sb is symmetrically related to b it is natural to include it in B as well. Hence, acting with s on the elements of B should leave the set unchanged and we naturally have a group action on B. Further, note that for any $b \in B$, we must have $sb \neq b$ for any nontrivial $s \in S$. Otherwise, s would be a symmetry of the output of our network. This is exactly the definition of a free group action. Hence, we can define a symmetry breaking set as follows.

Definition 3.1 (Symmetry breaking set). *Let S be a symmetry group. Let B be a set of elements which S acts on. Then B is a symmetry breaking set (SBS) if the action of S on B is a free action.*

3.2 Equivariant SBS

However, it turns out the above definition of a SBS is not sufficient when considering equivariant networks. To illustrate the problem, consider a network which is $SO(3)$ equivariant and a triangular prism aligned so that triangular faces are parallel to the xy plane. Suppose our task was just to pick a point of the prism. A naive way to break the symmetry is to have an ordered pair of unit vectors. The first vector is in the xy plane and points towards one of the triangle vertices. The second vector is just in the z direction point up or down corresponding to the upper or lower triangle.



(a) (b)

Figure 3.1: (a) Naive way to break symmetry in a triangular prism where one vector points to a vertice of a triangle and a second vector points to the lower or outer triangle. (b) A rotated version of the triangular prism in (a). Note that the same symmetry breaking objects now point to edges of the triangle rather than vertices. However, both prisms have the exact same symmetry elements.

However, consider the triangular prism which is rotated 180° around z. Looking only at the symmetry group of the two prisms, we can check that they are exactly the same. But, the symmetry breaking objects are related differently. In the second prism, the first vector points to an edge rather than a vertex. From an equivariant point of view, we would want our symmetry breaking objects to be related to our data in the same when no matter how we rotate it. So in some sense, our choice of SBS was not equivariant.

In this case, it is simple enough to choose a canonical orientation and decide that we will rotate the original SBS by 180° in the latter case. However, our input data may be arbitrarily complicated and it may be hard to choose a canonical orientation. Further, canonicalization may introduce discontinuities. Hence, we would like to construct SBSs to be only dependent on the symmetry of our data, not how our data is represented or its specific orientation. To understand exactly what additional condition is necessary, we need to think carefully about how the symmetry breaking scheme works.

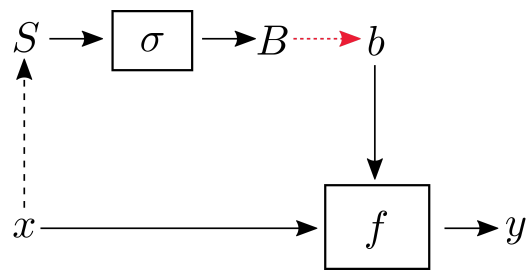


Figure 3.2: Diagram of how we might structure our symmetry breaking scheme. From our data x, we may obtain its symmetry S. This S is then fed into a function σ which gives us the set of symmetry breaking objects needed. We sample a b from this set breaking the symmetry of our input and feed this b along with the input x into our equivariant function f. Finally we obtain an output y which has lower symmetry than the input x.

Let f be our function which is equivariant under a group G. Let x be our input data. Suppose we know the symmetry S of our input. Let $\mathbf{B}$ be some set with a group action of G defined on it. We would like to obtain our set of symmetry breaking objects based on just information about the input symmetry. So suppose we have a function $\sigma : \text{Sub}(G) \to \mathcal{P}(\mathbf{B})$ that does so. This function takes in a subgroup symmetry and gives a set of symmetry breaking objects composed of elements from a set $\mathbf{B}$. Then the symmetry breaking step happens when we take a random sample b from the SBS given by the function σ. This symmetry breaking object is also fed into our equivariant function, allowing it to break symmetry. A diagram of this process is shown in Fig. 3.2.

Certainly, since we break the symmetry of our input data, we break equivariance. However, imagine we give our function all possible symmetry breaking objects and collect at the end the set Y of all outputs our model gives. This process shown in Fig. 3.3 then would not break any symmetry. The key insight now is that we can impose equivariance on this operation.

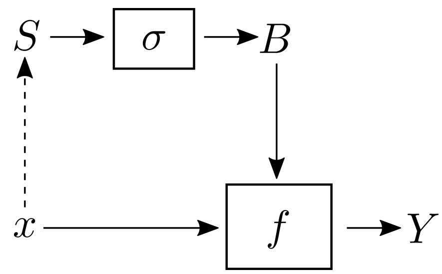


Figure 3.3: Diagram of how we break symmetry, but now we keep all possible outputs.

It is well known that the composition of equivariant functions remains equivariant. Hence, we just need to impose equivariance on σ. In order to do so, we must understand how the input and output transform. Suppose we act on our data with some group element $g \in G$. Then it becomes gx. Since S is the symmetry of our original data, we find $gx = gsx = (gsg^{-1})(gx)$ for any $s \in S$. So the symmetry of the transformed data is gSg^{-1}. Hence, the input of σ transforms as conjugation. Next, recall the output of σ is some subset B of elements of $\mathbf{B}$. Since there is a group action for G defined on $\mathbf{B}$, we can define a natural action on B by just acting on its elements and forming a new set. Fig. 3.4 shows how our procedure would change if it were equivariant and our input was rotated by some group element g.

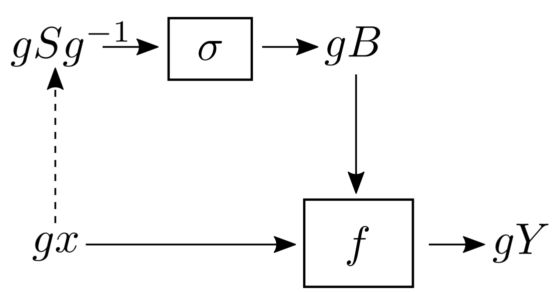


Figure 3.4: Diagram of what happens when we act on the input with some group element g.

With this understanding, we can try to characterize when we can define an equivariant σ. Note that the input of σ transforms under conjugation so we can split the domain into $\mathrm{Cl}_G(S)$, the set of conjugate subgroups of S.

Proposition 3.2. *Let G be a group and S be a subgroup of G. Let $B \in \mathcal{P}(\mathbf{B})$ be a set where there is some group action of G defined on $\mathbf{B}$. Then there exists an equivariant $\sigma|_{\mathrm{Cl}_G(S)} : \mathrm{Cl}_G(S) \to \mathcal{P}(\mathbf{B})$ such that $\sigma|_{\mathrm{Cl}_G(S)} = B$ if and only if $nB = B$ for all $n \in N_G(S)$.*

Proof. Note that $\mathrm{Cl}_G(S)$ is a set where action by conjugation is a transitive one. Also note by definition that $\mathrm{Stab}_G(S)$ for this action is precisely the definition of a normalizer $N_G(S)$. Then by Lemma 2.16, we see such a function exists if and only B is also symmetric under $N_G(S)$. □

Now that we know invariance under the normalizer is sufficient to be able to consistently define an equivariant $\sigma|_{\mathrm{Cl}}$, we can provide a proper definition for equivariant SBSs.

Definition 3.3 (Equivariant symmetry breaking sets)**.** *Let S be a subgroup symmetry of a group G and $\mathbf{B}$ be a set with an action of G defined on it. Let $B \subset \mathbf{B}$ be a SBS. Then B breaks the symmetry G-equivariantly if $\forall g \in N_G(S)$ we have $B = gB$.*

3.3 Ideal case and complement of normal subgroups

Now that we know what it means to equivariantly break a symmetry, we would like to understand how to do so efficiently. Intuitively, we expect a smaller SBS to be better and this is indeed true. If we have a larger SBS, multiple symmetry breaking objects might map to the same output so the network needs to learn that these are the same. Reducing the SBS would then decrease the equivalences our network needs to learn. In the ideal case, exactly one symmetry breaking parameter corresponds to each output. Since our outputs are related by symmetry transformations (transitive) under S, this corresponds to the equivariant SBS being transitive under S. It turns out we can equate the existence of an ideal equivariant SBS to the existence of complements of normal subgroups. A slightly weaker version of this statement can be found in Theorem 3.1.4 of [17].

Theorem 3.4. *Let G be a group and S a subgroup. Let B be a G-equivariant SBS for S. Then it is possible to choose an ideal B if and only if S has a complement in $N_G(S)$.*

Proof. Suppose B is transitive under S and pick $b \in B$. Consider the stabilizer group $\mathrm{Stab}_{N_G(S)}(b)$. For any $g \in N_G(S)$, by transitivity under S we must have $gb = sb$ for some $s \in S$. So, $s^{-1}gu = u$ implying that $h = s^{-1}g \in \mathrm{Stab}_{N_G(S)}(b)$. So we find that we can write any g as $g = sh$ for some $s \in S$ and $h \in \mathrm{Stab}_{N_G(S)}(u)$ so

$$N_G(S) = S \cdot \mathrm{Stab}_{N_G(S)}(u).$$

But note that since B is symmetry breaking, $S \cap \mathrm{Stab}_{N_G(S)}(u) = \{e\}$. Hence, $\mathrm{Stab}_{N_G(S)}(u)$ is indeed a complement.

For the converse, suppose H is a complement of S in $N_G(S)$. We claim $B = N_G(S)/H$ is the equivariant SBS we desire. Note that clearly by construction, this is closed under $N_G(S)$ so we satisfy the equivariance condition. Further, note that $\mathrm{Stab}_S(H) = \mathrm{Stab}_{N_G(S)}(H) \cap S = H \cap S = \{e\}$. Since B is transitive under $N_G(S)$, stabilizers of all other elements are obtained by conjugation and hence also trivial. Hence, it is indeed symmetry breaking. Finally, any $g \in N_G(S)$ is uniquely written as sh for some $s \in S, h \in H$ so $gH = shH = sH$. So B is transitive under S as well. □

Remark 3.5. *It turns out the complement if it exists is isomorphic to $N_G(S)/S$. Further, we can intuitively think of $N_G(S)/S$ as giving all possible orientations of our data such that its symmetry remains unchanged.*

Finding complements of normal subgroups is a well studied problem in group theory [17]. In the case of point groups, which are the finite subgroups of $O(3)$, we have explicitly tabulated in Appendix B whether we can achieve an ideal equivariant SBS.

3.4 Nonideal equivariant SBSs

In the case where we cannot achieve an ideal equivariant SBS, we would still like to characterize how efficient it is. To do this, we define what we call the degeneracy of an equivariant

SBS. In general, each orbit under S gives us one SBS which can be matched one to one to our outputs.

Definition 3.6 (Degeneracy)**.** *Let B be a G-equivariant SBS for S. We define the degeneracy to be*

$$\mathrm{Deg}_S(B) = |B/S|.$$

Note that an ideal equivariant SBS B_{ideal} (if it exists) is has exactly 1 orbit of S so $\mathrm{Deg}_S(B_{ideal}) = 1$. We would also like to understand how small we can make the degeneracy if we cannot make it 1. It turns out Theorem 3.4 allows us to convert this to a group theory problem.

Corollary 3.7. *Let G be a group and S a subgroup. Let M be such that $S \leq M \leq N_G(S)$. Let B be a G-equivariant SBS for S which is transitive under $N_G(S)$. Then it is possible to choose B such that every M-orbit is also transitive under S if and only if S has a complement in M. In particular, such a B has*

$$\mathrm{Deg}_S(B) \leq |N_G(S)/M|.$$

Proof. Suppose we have such a B and pick any $b \in B$. By transitivity of the orbit under S, we have $Mb = Sb$. Let $B' = Mb$. We can check that this is in fact an ideal M-equivariant SBS for S. That it is a symmetry breaker follows since B is symmetry breaking. That it is M-equivariant and transitive follows since $Mb = Sb$ and $N_M(S) = M$. By Theorem 3.4 this implies S has a complement in M.

Next, suppose we have a complement of S in M. By Theorem 3.4 we can construct B' which is an ideal M-equivariant SBS for S. We can lift this to a G-equivariant SBS for S by just taking $B = N_G(S)B'$.

Finally, to compute the order, we note that every S-orbit is also a M orbit. Since B is transitive under $N_G(S)$, there are at most $|N_G(S)/M|$ number of M-orbits and hence only that many S-orbits. So

$$\mathrm{Deg}_S(B) \leq |N_G(S)/M|.$$

□

In the ideal case we can make M to be $N_G(S)$ so the above formula gives an degeneracy of 1.

Chapter 4

Partially broken symmetry

We can now use our framework for full symmetry breaking to understand the case of partial symmetry breaking. In this case, our desired output may share some nontrivial subgroup symmetry $K \leq S$ with our input. Note the case of $K = 1$ corresponds to full symmetry breaking and $K = S$ corresponds to no symmetry breaking.

4.1 Partial SBS

Similar to the full symmetry breaking case, we would like to create a set of objects which we can use to break our symmetry. Now we can relax the restriction of free action. Intuitively, we want to allow our symmetry breaking objects to share symmetry with our input, as long as it is lower symmetry than our outputs. However, the symmetrically related outputs may be invariant under different subgroups of S. Recall that if some element y gets transformed to sy, its stabilizer K gets transformed to sKs^{-1}. Hence, the stabilizers of the outputs are from $\mathrm{Cl}_S(K)$, the subgroups conjugate to K under S. Based on this intuition, we can define partial SBS as follows.

Definition 4.1 (Partial SBSs). *Let S be a symmetry group and K a subgroup of S. Let P be a set of elements with an action by S. Then P is a K-partial SBS if for any $p \in P$, there* *exists some* $K' \in \mathrm{Cl}_S(K)$ *such that* $K' \geq \mathrm{Stab}_S(p)$.

Certainly, a full SBS is a partial one as well since the stabilizers of all its elements under S is the trivial group. In general, we can always break more symmetry than needed and still obtain our desired output. However, it is useful to consider the case where we only break the necessary symmetries.

Definition 4.2 (Exact partial SBS)**.** *Let S be a symmetry group and K a subgroup of S. Let P be a K-partial SBS for S. We say P is exact if for all $p \in P$, we have* $\mathrm{Stab}_S(p) \in \mathrm{Cl}_S(K)$.

4.2 Equivariant partial SBS

Similar to before, we can define equivariant partial SBSs. The idea is the same, but now we need to identify the symmetry of the input and the set of conjugate symmetries for the output. Define

$$\mathrm{SubCl}(G) = \{(S, \mathrm{Cl}_S(K)) : S \in \mathrm{Sub}(G), K \leq S\}.$$

Let $\mathbf{P}$ be a set with a group action of G defined on it. As before, the idea is that we have an function $\pi : \mathrm{SubCl}(G) \to \mathcal{P}(\mathbf{P})$ which outputs our partial SBS. The condition of equivariance for our partial SBS is imposing equivariance on π.

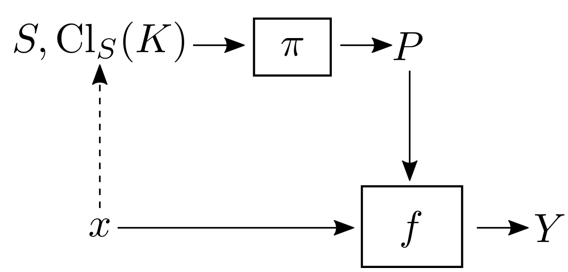


Figure 4.1: Diagram of how we perform partial symmetry breaking. Here, we need to specify not just the symmetry of our input but also the symmetries of our output. Since any of our outputs are equally valid, it only makes sense to specify the set of conjugate subgroups $\mathrm{Cl}_S(K)$ our outputs are symmetric under.

The symmetry breaking scheme is depicted in Figure 4.1. As before, we can impose equivariance on this diagram. We need to know how $\mathrm{Cl}_S(K)$ transform. Note that if our input gets acted by g, we expect the outputs to also get acted by g. Since K is the stabilizer of one of the outputs, we expect K to transform to gKg^{-1}. Hence we have the transformation

$$\mathrm{Cl}_S(K) \to \mathrm{Cl}_{gSg^{-1}}(gKg^{-1}).$$

Similar to before, by Lemma 2.16 we need the output of π to also be invariant under the stabilizer of the input. Noting that the normalizer is defined as the stabilizer of S under conjugation, we can define a generalized normalizer as the stabilizer of $S, \mathrm{Cl}_S(K)$.

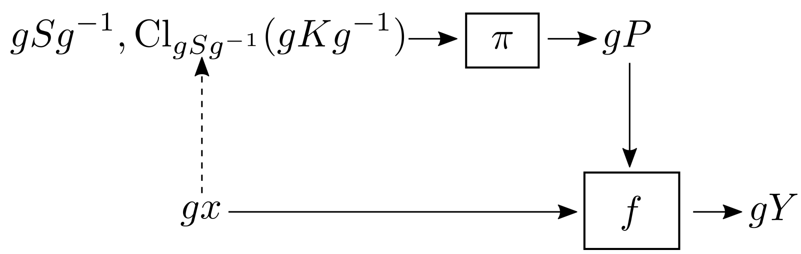


Figure 4.2: Diagram of how our symmetry scheme changes when we transform our input by some group element $g \in G$.

Definition 4.3 (Generalized normalizer)**.** *Define the generalized normalizer $N_G(S, K)$ to be*

$$N_G(S, K) = \{g : gKg^{-1} \in \mathrm{Cl}_S(K), g \in N_G(S)\}.$$

Proposition 4.4. *Let G be a group, S a subgroup of G, and K a subgroup of S. Let $\mathbf{P}$ be a set with a group action of G defined on it and $P \subset \mathbf{P}$. There exists an equivariant $\pi|_{\mathrm{Orb}_G((S,\mathrm{Cl}_S(K)))} : \mathrm{Orb}_G((S, \mathrm{Cl}_S(K))) \to \mathcal{P}(\mathbf{P})$ such that $\pi|_{\mathrm{Orb}_G((S,\mathrm{Cl}_S(K)))}((S, \mathrm{Cl}_S(K))) = P$ if and only if $N_G(S, K)$ leaves P invariant.*

Proof. By Lemma 2.16, we need P to be closed under the stabilizer of the input. But the generalized normalizer $N_G(S, K)$ is precisely this stabilizer. □

We can now define equivariant partial SBSs as follows.

Definition 4.5 (Equivariant partial SBSs)**.** *Let S be a subgroup symmetry of a group G.*

Let P be a K-partial SBS. Then P breaks the symmetry G-equivariantly if $\forall g \in N_G(S,K)$ we have $P = gP$.

Note that closure under $N_G(S,K)$ is a weaker condition than closure under $N_G(S)$. Hence any full equivariant SBS is also an equivariant K-partial SBS for any K.

4.3 Ideal equivariant partial SBS

Similar to the full symmetry breaking case, we ideally would like to have a one to one correspondence between elements in our equivariant SBS and our symmetrically related outputs. For this to happen, we clearly need our SBS to be exact and for our SBS to be transitive under S. We can generalize Theorem 3.4 to obtain a necessary and sufficient condition to have an ideal equivariant partial SBS.

Theorem 4.6. *Let G be a group and S and K be subgroups $K \leq S \leq G$. Let P be a G-equivariant K-partial SBS. Then we can choose an ideal P (exact and transitive under S) if and only if $N_S(K)/K$ has a complement in $N_{N_G(S,K)}(K)/K$.*

Proof. Let $P = Su$ where u has symmetry $\text{Stab}_S(u) = K$. We can define an action of any coset $N_G(S,K)/K$ on u as just the action of a coset representative on u. This is consistent since u is invariant under K. In particular, note that K is a normal subgroup of $N_S(K)$ so $N_S(K)/K$ is a quotient group. Let $B' = (N_S(K)/K)u$. Since u is in a K-partial SBS, we must have $su \neq u$ for any $s \in S - K$. Hence, for any coset $gK \in N_S(K)/K$, $gu \neq u$ if $g \notin K$. Therefore, B' must be a SBS for $N_S(K)/K$.

Next, consider any coset gK in $N_{N_G(S,K)}(K)/K$. Then we know $gu \in Su$ so $gu = su$ for some $s \in S$. Since K was a symmetry of u, $gKg^{-1} = sKs^{-1}$ is a symmetry of $gu = su$. So the stabilizer of su must be $sKs^{-1} = K$. Hence, s must be in $N_S(K)$. Therefore the action of gK on u gives us an element of $B' = (N_S(K)/K)u$. Hence B' is $N_{N_G(S,K)}(K)/K$-equivariant.

By Theorem 3.4, the existence of an ideal $N_{N_G(S,K)}(K)/K$-equivariant SBS for $N_S(K)/K$ implies that $N_S(K)/K$ has a complement in $N_{N_G(S,K)}(K)/K$.

For the converse direction, suppose that A is a complement of $N_S(K)/K$ in $N_{N_G(S,K)}(K)/K$. Note the elements of A are cosets of K so we can define a set of elements of $N_G(S,K)$ as

$$H = \bigcup_{C \in A} C.$$

Define $P = \mathrm{Orb}_S(H)$. We claim that P is a transitive exact equivariant partial SBS.

We first show that P is exact K-partial symmetry breaking. Consider $s \in S$. We can write

$$sH = \bigcup_{C \in A} sC = \bigcup_{C \in sA} C.$$

Now we see if $s \in K$, then since K is the identity in the quotient group $sA = A$. Hence $sH = H$ in this case. If $s \in N_S(K) - K$, then sK is not the identity in $N_S(K)/K$. But A is a complement so $sA \neq A$ implying $sH \neq H$. Finally, if $s \notin N_S(K)$ then $sK \notin N_{N_G(S,K)}(K)/K$. So $sH \not\subset N_{N_G(S,K)}(K)$. But $H \subset N_{N_G(S,K)}(K)$ so $sH \neq H$. Hence, $\mathrm{Stab}_S(H) = K$ and since the rest of P is just the orbit of H, stabilizers of the other elements are in $\mathrm{Cl}_S(K)$. Hence, P as we constructed is an exact K-partial SBS.

For equivariance consider any $n \in N_G(S,K)$ giving a coset

$$nH = \bigcup_{C \in A} nC = \bigcup_{C \in nA} C.$$

If $n \in N_{N_G(S,K)}(K)$ then since A is a complement, $(nK) = (sK)(aK)$ for some $sK \in N_S(K)/K$ and $aK \in A$, so $nA = (nK)A = (sK)(aK)A = (sK)A = sA$ for some $s \in N_S(K) \subset S$. Hence, $nH = sH$ for some $s \in S$. If $n \notin N_{N_G(S,K)}(K)$, then there is some s so that $nKn^{-1} = sKs^{-1}$. Therefore, $s^{-1}nKn^{-1}s = K$ so $s^{-1}n \in N_{N_G(S,K)}$. But we saw before that this means there is some s' such that $s^{-1}nH = s'H$. Thus, $nH = ss'H$ and $ss' \in S$. So $nH \in P$ so P is indeed closed under action by $N_G(S)$. □

4.4 Nonideal equivariant partial SBS

Similar to the full symmetry breaking case, when we cannot achieve an ideal equivariant partial SBS we want to characterize how efficient our nonideal partial SBS is. Again, the idea is that in the nonideal case, our network will need to map multiple symmetry breaking objects to the same output and we define the degeneracy of P to quantify this multiplicity.

Definition 4.7 (Degeneracy)**.** *Let G be a group, S be a subgroup, and K a subgroup of S. Let P be a G-equivariant K-partial SBS for S. Let T be a transversal of S/K. Let P_t be such that every $p \in P$ is uniquely written as $p = tp_t$ for some $t \in T$ and $p_t \in P_t$. Then we define*

$$\mathrm{Deg}_{S,K}(P) = |P_t|.$$

The intuition for this definition is that P_t is the set of objects which together with our input may get mapped to some output y by our equivariant network. In other words, we may have $f(x, P_t) = \{y\}$ for equivariant f and all other P get mapped to different symmetrically related outputs. Without loss of generality we can assume y has $\mathrm{Stab}_S(y) = K$. Then for any symmetrically related output ty (where $t \in T$), we can see from equivariance of f that $f(x, tP_t) = \{ty\}$. It is now clear that the size of P_t counts how many symmetry breaking objects must be mapped to the same output.

Note that in the case $K = 1$, then $S/K = S$ so P_t just consists of representatives from P/S. So this reduces to the degeneracy defined for full SBS. Also, note that in the ideal case, there is exactly one symmetry breaking object for each output. So degeneracy is 1 in that case.

As before, we would like to know how small we can make the degeneracy of our equivariant partial SBSs. Similar to the full symmetry breaking case, we can use Theorem 4.6 to convert this into a group theory question.

Corollary 4.8. *Let G be a group, S a subgroup, and K a subgroup of S. Let K' be a* *subgroup of K and M a subgroup of $N_G(S,K) \cap N_G(S,K')$ which contains S. Suppose P is a G-equivariant K-partial SBS for S which is transitive under $N_G(S,K)$. We can choose P such that $\mathrm{Stab}_S(p) \in \mathrm{Cl}_{N_G(S,K)}(K')$ for all p and all M-orbits in P are transitive under S if and only if $N_S(K')/K'$ has a complement in $N_M(K')/K'$. Further, such a P has*

$$\mathrm{Deg}_{S,K}(P) \leq |K/K'| \cdot |N_G(S,K)/M|.$$

Proof. Suppose we had such a P. Pick some $p \in P$ such that $\mathrm{Stab}_S(p) = K'$. Since M-orbits are transitive under S, we have $Mp = Sp$. Let $P' = Mp$. We can check then that this is a M-equivariant set. Further, since $M \subset N_G(S,K')$, we see that this is an exact K'-partial symmetry breaking set. Hence, it is an ideal M-equivariant K'-partial SBS. Also note that since $M \subset N_G(S,K')$, we have $M = N_M(S,K')$. So by Theorem 4.6, $N_S(K')/K'$ must have a complement in $N_M(K')/K'$.

Conversely, suppose $N_S(K')/K'$ has a complement in $N_M(K')/K'$. Again, we note $M = N_M(S,K')$ so by Theorem 4.6 we have an ideal M-equivariant K'-partial SBS for S. We can lift this to G-equivariance by taking the orbit under $N_G(S,K)$.

To see the order of such a P, we consider the S-orbits. Let T be a transversal of S/K. For each S-orbit, we can pick some p in that orbit so that $\mathrm{Stab}_S(p) \leq K$. We put the elements of Kp in our P_t. Within each S-orbit, any p' can be written as sp for some $s \in S$ and any s is uniquely written as $s = tk$ for some $t \in T$ and $k \in K$. So $p' = tkp$. However, note that any other s' where $p' = s'p = sp$ can be written as $s' = sk'$ for some $k' \in \mathrm{Stab}_S(p)$. Hence, p' is uniquely written as $p' = t(kp)$ since $kk'p = kp$. So each S-orbit contributes $|K/\mathrm{Stab}_S(p)|$ elements to P_t. However, since $\mathrm{Stab}_S(p) \in \mathrm{Cl}_{N_G(S,K)}(K')$, we must have $|K/\mathrm{Stab}_S(p)| = |K/K'|$. Finally, we know each S-orbit is also an M-orbit, since P is transitive under $N_G(S, K)$, there are at most $|N_G(S, K)/M|$ different S-orbits. So

$$\text{Deg}_{S,K}(P) = |P_t| \leq |K/K'| \cdot |N_G(S, K)/M|.$$

□

4.5 Optimality of exact partial symmetry breaking

Note that in the previous section, we have been very careful to allow our partial SBS to break more symmetry than needed. Intuitively, we would like to say that it is always optimal to break down exactly to the symmetry of our output. That is we only need to consider exact partial SBSs.

Certainly, ignoring any equivariance constraints, given any non-exact K-partial SBS, we can construct an exact K-partial SBS by picking an element b with $\text{Stab}_S(b) \leq K$ and identifying its orbit under K together as one partial symmetry breaking object $p = Kb$. We construct the orbit of p under action by S as our K-partial SBS.

We might expect there to be some modification of this construction to convert an non-exact equivariant K-partial SBS into an exact equivariant K-partial symmetry breaking one. Naively, we just take the orbit of the elements in the construction above under $N_G(S, K)$ to obtain G-equivariance. However, we can come up with a contrived explicit example where no exact equivariant K-partial SBS is smaller than the best equivariant full SBS.

Chapter 5

Experiments

Here we provide some example tasks where we apply our framework to full symmetry breaking and partial symmetry breaking cases. We consider the cases where we can find an ideal equivariant SBS or partial SBS. We explicitly work through how to obtain the equivariant SBS in each case.

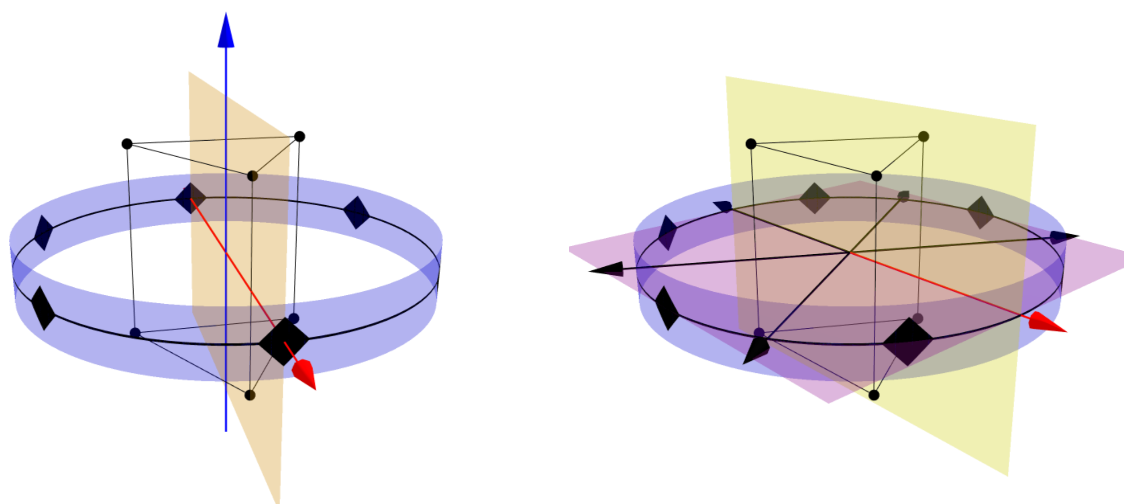
For each of our tasks, we trained a equivariant convolutional message-passing graph neural network (GNN) to output a vector pointing to a vertex of the prism. We use a modified version of the default network from the `e3nn` library [22]. In our modified version we input the graph edges instead of automatic graph generation based on a radius cutoff. Each layer consists of the equivariant 3D steerable convolutions followed by gated nonlinearities described in [23]. We use a gaussian basis for our radial network.

5.1 Full symmetry breaking: chiral triangular prism

For an example of full symmetry breaking, we consider the task of pointing to a vertex of a triangular prism, similar to that described in section 3.2. Our input is a graph with 6 nodes with edges given by the edges of the prism. We have position features at the nodes corresponding to the positions of the vertices. Since we want this to be a full-symmetry breaking task, we require chirality in our prism. In order to make it chiral, we also have pseudoscalar features of value 1 at all the vertices. We want our model to output a vector which points to the position of a vertex from the center of the prism.

For this task, the hidden features in our model are up to $l = 2$ of both parities and our convolutional filters use up to $l = 4$ spherical harmonics. For our radial network, we use a 3 layer fully connected network with 16 hidden features in each layer. If we need to break symmetry, we add our symmetry breaking object as an additional feature to all nodes of the graph.

5.1.1 Obtaining an equivariant SBS

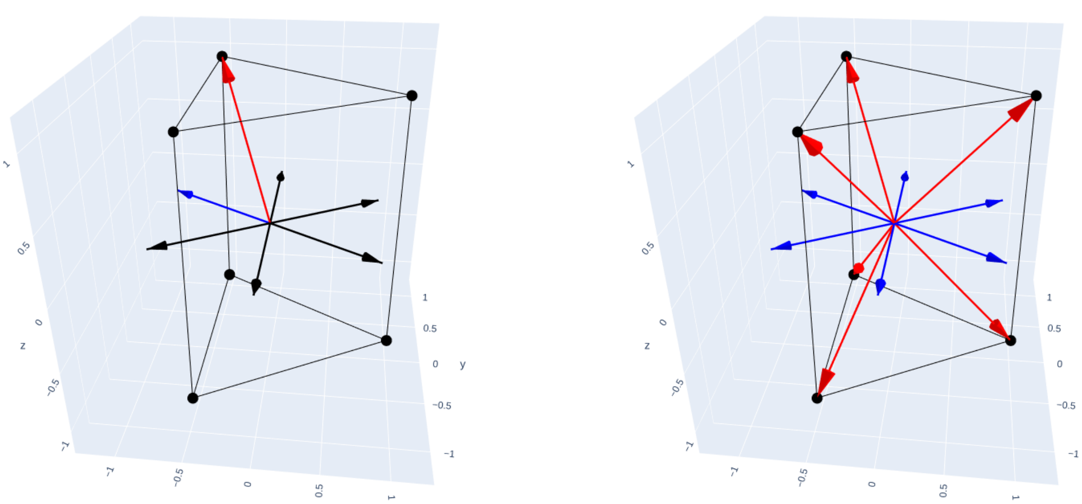


(a) (b)

Figure 5.1: (a) Triangular prism with D_3 symmetry and patterned cylinder with D_{6h} symmetry. The generators are a, b, m where a is a $2\pi/6$ rotation about the blue axis, b is a π rotation about the red axis, and m is a reflection across the orange plane. (b) An ideal symmetry breaking set for the triangular prism. A complement of D_3 in D_{6h} is generated by the mirror planes shown here in yellow and purple. The vector in red is a symmetry breaking object with this complement as stabilizer. The orbit of this vector under the normalizer generates the other vectors shown in black.

In this case, it turns out a choice of ideal equivariant SBS is the set of unit vectors parallel to an edge of the triangular faces of the prism. Table B.3 tells us D_3 is generated by a^2, b and that a complement H is generated by am, bm. By Theorem 3.4, we know that if we can pick some object v with stabilizer $\mathrm{Stab}_{D_{6h}}(v) = H$, then the orbit of v under D_3 gives an ideal equivariant SBS. In this case, one such v that works is a vector parallel to the triangular faces of the prism and one of the sides of the prism. This is shown in Figure 5.1b. Note reflection across the yellow plane corresponds to am and across the purple plane corresponds to bm. It is clear the arrow in red is stabilized by this complement. One can further check it shares no symmetries with the triangular prism. The other 5 arrows in black are the other symmetry breaking objects we obtain by taking the orbit of the red arrow under action by D_{6h}.

5.1.2 Training results



(a) (b)

Figure 5.2: (a) Output (red) generated by our model and symmetry breaking object (blue) given. (b) The set of all the outputs generated by our model if we feed in all symmetry breaking objects.

We first fix a choice of one symmetry breaking object from our equivariant SBS and one of the vertices of the triangular prism. We then give the chosen symmetry breaking object as an additional input to our equivariant GNN and train it to output a vector (odd

$l = 1$) feature pointing to our chosen point from the center. An example of the result of this training is shown in Figure 5.2a. We also observe that no matter which choices of vertex and symmetry breaking object we pick, our equivariant network is able to learn to output the vector pointing that that vertex. In practice, this means that we can choose any of our symmetry breaking objects as additional input.

Once trained on one pair of symmetry breaking object and vertex, the equivariance of our GNN means that inputting the other symmetry breaking objects in our SBS gives the other symmetrically related outputs. This is shown in Figure 5.2b.

Further, rather than picking one vertex, we also tried modifying our loss so that we compute the loss for all choices of vertex and take the minimum. Hence, our network can learn which vertex to pair with each symmetry breaking object. In this prism example, our pairing is random. This is method of taking the minimum loss is especially useful when we have multiple instances of symmetry breaking in our data.

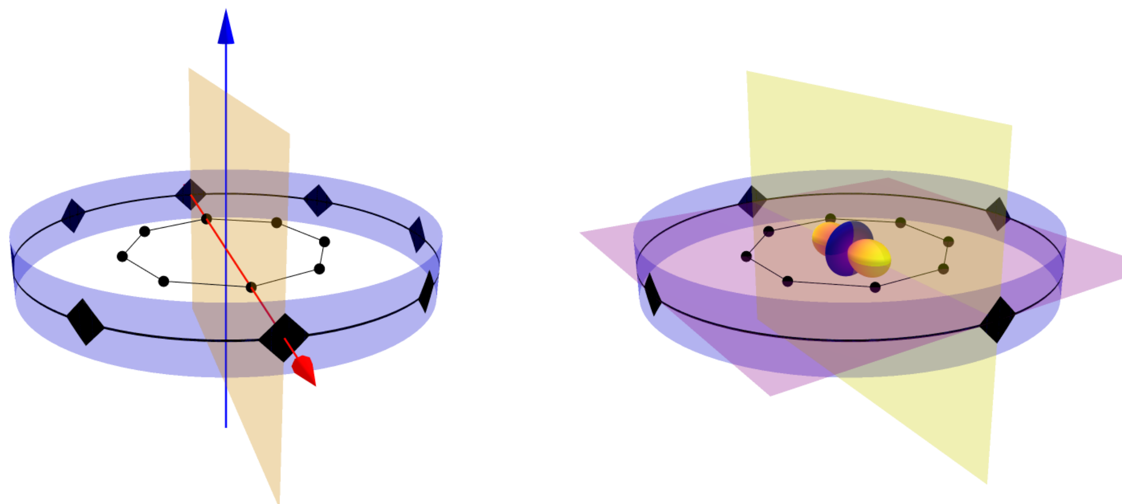
5.2 Partial symmetry breaking: octagon to rectangle

For an example of partial symmetry breaking, we consider the task of deforming an octagon to a rectangle. We choose to make our octagon chiral and impose chirality on our octagon by adding pseudoscalar features of value 1 to all vertices. We select this example this because the construction of the stabilizer for a single symmetry breaking object illustrates the general procedure. The nodes of the input graph are just the vertices of the octagon and the edges are just the sides.

We use the exact same architecture as for the triangular prism experiment. Here, we output vector ($l = 1$) features on each vertex which represent how much we should distort that vertex.

5.2.1 Obtaining an equivariant partial SBS

One choice of ideal equivariant SBS for this case consists of $l = 2$ objects aligned to be parallel to an edge of the octagon. In this scenario, we want G-equivariance for $G = O(3)$ and we have $S = D_8$ and partial symmetry $K = D_2$.



(a) (b)

Figure 5.3: (a) Octagon with D_8 symmetry we input and patterned cylinder with $N_G(S, K) = D_{8h}$ symmetry. The generators are a^2, b, m where a^2 is a $2\pi/8$ rotation about the blue axis, b is a π rotation about the red axis, and m is a reflection across the orange plane. (b) A symmetry breaking object which generates an ideal partial SBS. Note that in addition to the D_2 symmetry, this object is also symmetric under reflections across the purple and yellow planes.

The octagon has D_8 symmetry and we know from Table B.1 that its normalizer is D_{16h}. We also know from Appendix B.3 that a presentation of this normalizer is

$$\langle a, b, m | a^{16}, b^2, m^2, (ab)^2, (am)^2, (bm)^2 \rangle.$$

Note the symmetry of the octagon S is generated by a^2, b and that of a rectangle K by a^8, b. Next, we can check by brute force that $N_G(S, K)$ is generated by a^2, b, m.

We then need to compute $N_S(K)$ and $N_{N_G(S,K)}(K)$. We note from Table B.1 that $N_G(K) = D_{4h}$ and in particular, it is not hard to see the specific copy of D_{4h} is generated by a^4, b, m. We now just have $N_S(K) = S \cap N_G(K)$. Looking at the generators, we can check that m is not present in S so $N_S(K) = D_4$ and is generated by a^4, b. For $N_{N_G(S,K)}$, from the generators of $N_G(S, K)$ we can see that $N_G(K)$ is a subgroup so $N_{N_G(S,K)}(K) = N_G(K) = D_{4h}$ and is generated by a^4, b, m.

Finally, from Theorem 4.6, we need to look at the quotient groups $N_S(K)/K$ and $N_{N_G(S,K)}(K)/K$. For the latter, we can set the cosets

$$X = \{a^4, a^{12}, a^4b, a^{12}b\} \qquad Y = \{m, a^8m, bm, a^8bm\}$$

which we can check generate the quotient group. In particular we have relations $X^2 = Y^2 = (XY)^2 = 1$ so a presentation is given by just

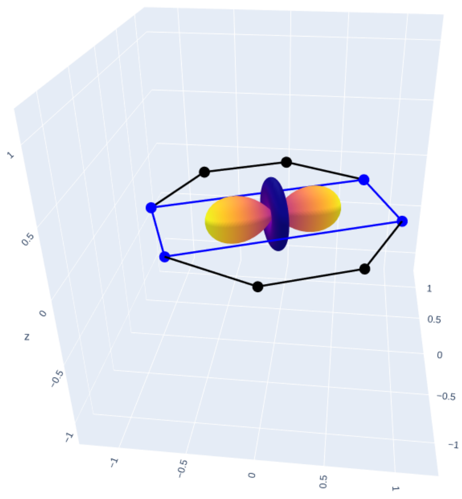
$$\langle X, Y | X^2, Y^2, (XY)^2 \rangle.$$

For $N_S(K)/K$ which is a subgroup of this, we can see that it is just generated by X. But it is easy to see that the quotient group generated by Y forms a complement.

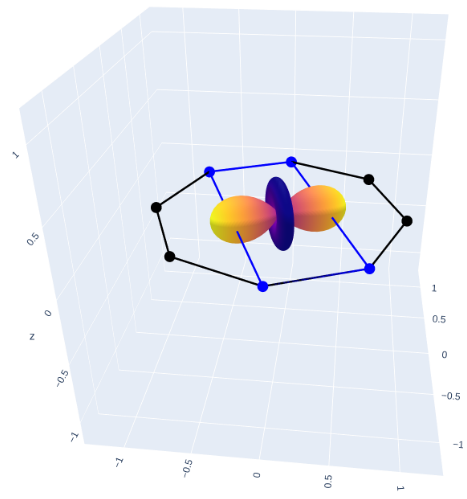
Following the argument in Theorem 4.6, we see that we need a symmetry breaking object with stabilizer generated by group elements in the cosets of the complement of $N_S(K)/K$. Since here the complement is generated by coset Y, we need the stabilizer to be generated by m, a^8m, bm, a^8bm. It is not hard to check we can simplify the list of generators to a^8, b, m. We note that an $l = 2$ irrep with even parity works in this case.

5.2.2 Training results

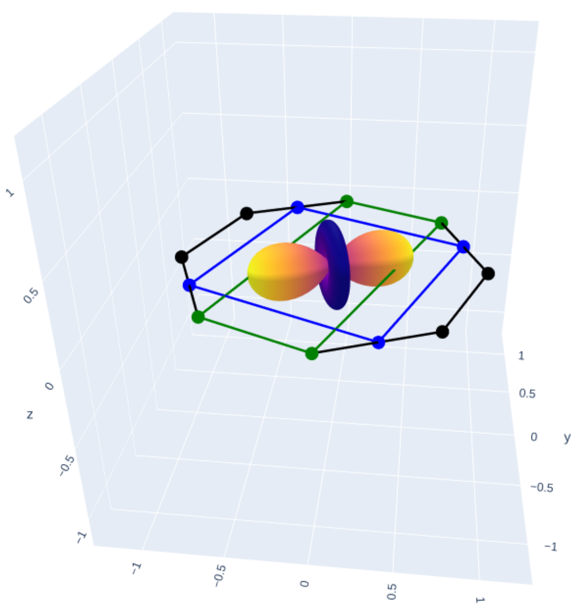
Similar to the prism case, we try training by matching a specific symmetry breaking object to a rectangle. When the symmetry breaking object and rectangle are compatible (share the same symmetries), then our model has no problem learning to deform the octagon



(a)



(b)



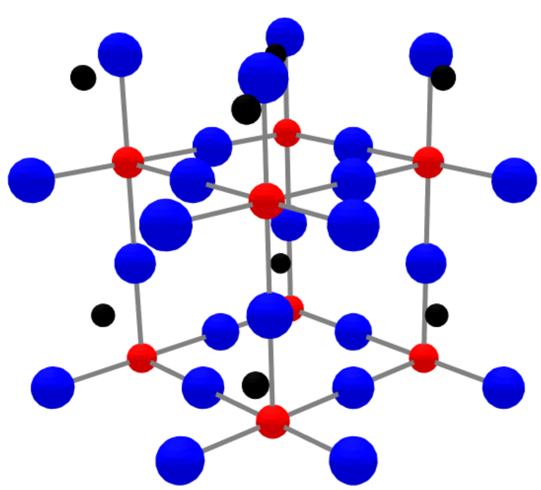
(c)

Figure 5.4: (a) Output (blue) of our model when we match a symmetry breaking object with a compatible rectangle. (b) Output (blue) of our model when we match a symmetry breaking object with a different compatible rectangle. (c) Output (blue) when we attempt to match the symmetry breaking object with an incompatible rectangle (green). Note the square has symmetries of both the symmetry breaking object and the target rectangle.

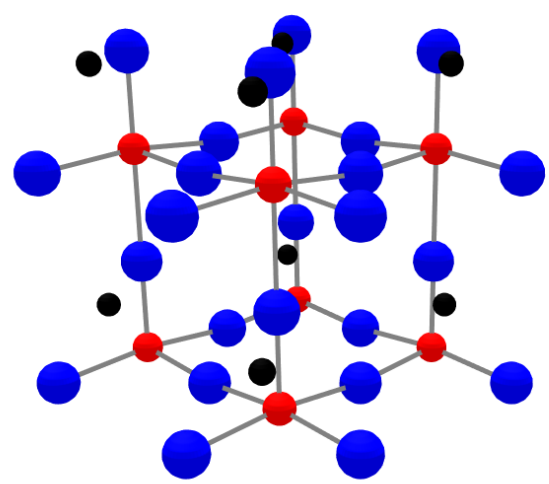
into the rectangle. This is shown in Figures 5.4a and 5.4b. An interesting failure case occurs when we try to match a symmetry breaking object and rectangle with incompatible symmetries. This is shown in Figure 5.4c. Here, the D_2 symmetry of the rectangle and of the symmetry breaking object are misaligned. As a result, our model predicts an output which has symmetry of D_4 which is the group generated when we include the symmetry elements of both the target rectangle and the symmetry breaking object. Hence, the resulting shape is a square.

As with the triangular prism case, we also tried letting the model choose which rectangle to deform to given a symmetry breaking object. In this case, our model computes loss separately for all 4 possible rectangles and takes the minimum. We note that for a given symmetry breaking object, 2 of the possible rectangles are symmetrically compatible while 2 are not. Over 200 random initializations, we find roughly 30% of the time our model attempts to match symmetrically incompatible symmetry breaking objects to a rectangle. This is better than the 50% we would expect if it matches pairs randomly.

5.3 $BaTiO_3$ phase transitions



(a)



(b)

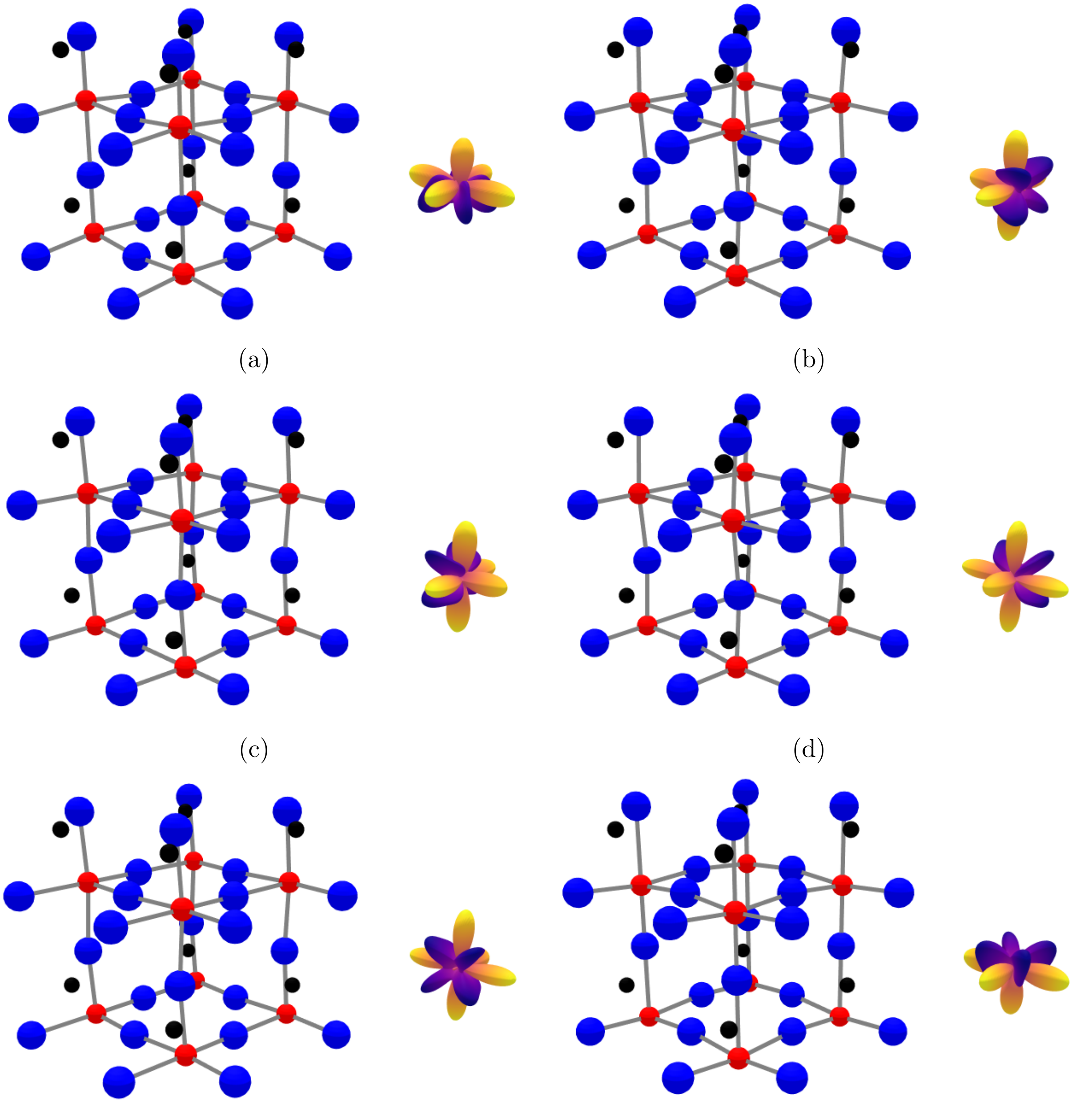
Figure 5.5: (a) Initial high symmetry crystal structure of $BaTiO_3$. (b) Target low symmetry crystal structure of $BaTiO_3$.

Finally, we demonstrate our framework on a more realistic example. For this, we examine the crystal structure of barium titanate ($BaTiO_3$). In particular, as we decrease temeperature, there is a phase transition from a high symmetry state with O_h symmetry to a lower symmetry state with C_{4v} symmetry at 403K [24], [25]. The high and low symmetry states are shown in Figures 5.5a and 5.5b respectively. For this task, we seek to deform a high symmetry state into the lower symmetry one. Data for the high and low symmetry $BaTiO_3$ crystals are obtained from materials project database [26]. For our demonstration, we focus on breaking point group symmetries. Hence, we set the unit cell of both crystals to be a cube with side length 4Å, which is close to the real unit cells.

The model used for this task has a similar architecture as the previous ones, with the modification of incorporating periodic boundary conditions because we are modelling a crystal. Similar to the octagon distortion task, we output vector ($l = 1$) features at each node which tells us how much to distort the corresponding atom.

It turns out that any object with C_{4v} symmetry works for generating an ideal equivariant partial SBS. This is because O_h has itself as normalizer in $O(3)$ so the symmetry completely determines orientation. A simple choice consists of a spherical harmonic with appropriate nonzero $l = 1$ and $l = 4$ terms.

The result of our training is shown in Figure 5.6. In Figure 5.6a, we have the symmetry breaking parameter which our model matches to the observed tilted output. We see that feeding the rest of the symmetry breaking parameters lets our model recover the other distortions.



(a) (b) (c) (d) (e) (f)

Figure 5.6: Distortions of a highly symmetric crystal structure of $BaTiO_3$ when provided with each of the possible symmetry breaking objects in our ideal equivariant partial SBS.

Chapter 6

Conclusion

We propose the idea of equivariant symmetry breaking sets which allows equivariant networks to sample or generate all possible symmetrically related outputs given a highly symmetric input. We show that bounding the size of these sets is intimately connected to a well studied group theory problem and tabulate solutions for the ideal case for the point groups. We demonstrate how our symmetry breaking framework on toy example problems.

One future direction is to include translations and tabulate complements for the space groups in their respective normalizers. This would be particularly useful for crystallography applications. Another direction is to automate finding stabilizers for partial symmetry breaking objects. Finally, our method assumes we can efficiently detect the symmetry of our input and outputs. Designing fast symmetry detection algorithms would also be extremely beneficial.

Bibliography

[1] I. Batatia, D. P. Kovacs, G. Simm, C. Ortner, and G. Csányi, "Mace: Higher order equivariant message passing neural networks for fast and accurate force fields," *Advances in Neural Information Processing Systems*, vol. 35, pp. 11 423–11 436, 2022.

[2] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky, "E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials," *Nature communications*, vol. 13, no. 1, p. 2453, 2022.

[3] O.-E. Ganea, X. Huang, C. Bunne, Y. Bian, R. Barzilay, T. Jaakkola, and A. Krause, "Independent se (3)-equivariant models for end-to-end rigid protein docking," *arXiv preprint arXiv:2111.07786*, 2021.

[4] Y.-L. Liao and T. Smidt, "Equiformer: Equivariant graph attention transformer for 3d atomistic graphs," *arXiv preprint arXiv:2206.11990*, 2022.

[5] D. Wang, J. Y. Park, N. Sortur, L. L. Wong, R. Walters, and R. Platt, "The surprising effectiveness of equivariant models in domains with latent symmetry," *arXiv preprint arXiv:2211.09231*, 2022.

[6] T. E. Smidt, M. Geiger, and B. K. Miller, "Finding symmetry breaking order parameters with euclidean neural networks," *Physical Review Research*, vol. 3, no. 1, p. L012002, 2021.

[7] A. Beekman, L. Rademaker, and J. van Wezel, "An introduction to spontaneous symmetry breaking," *SciPost Physics Lecture Notes*, p. 011, 2019.

[8] P. M. Woodward, "Octahedral tilting in perovskites. i. geometrical considerations," *Acta Crystallographica Section B: Structural Science*, vol. 53, no. 1, pp. 32–43, 1997.

[9] N. Huang, R. Levie, and S. Villar, "Approximately equivariant graph networks," *arXiv preprint arXiv:2308.10436*, 2023.

[10] T. van der Ouderaa, D. W. Romero, and M. van der Wilk, "Relaxing equivariance constraints with non-stationary continuous filters," *Advances in Neural Information Processing Systems*, vol. 35, pp. 33 818–33 830, 2022.

[11] R. Wang, R. Walters, and T. E. Smidt, "Relaxed octahedral group convolution for learning symmetry breaking in 3d physical systems," *arXiv preprint arXiv:2310.02299*, 2023.

[12] R. Wang, R. Walters, and R. Yu, “Approximately equivariant networks for imperfectly symmetric dynamics,” in *International Conference on Machine Learning*, PMLR, 2022, pp. 23 078–23 091.

[13] S. Balachandar, A. Poulenard, C. Deng, and L. Guibas, “Breaking the symmetry: Resolving symmetry ambiguities in equivariant neural networks,” *arXiv preprint arXiv:2210.16646*, 2022.

[14] S.-O. Kaba and S. Ravanbakhsh, “Symmetry breaking and equivariant neural networks,” *arXiv preprint arXiv:2312.09016*, 2023.

[15] M. S. Dresselhaus, G. Dresselhaus, and A. Jorio, *Group theory: application to the physics of condensed matter*. Springer Science & Business Media, 2007.

[16] D. S. Dummit and R. M. Foote, *Abstract algebra*. Wiley Hoboken, 2004, vol. 3.

[17] H. Kurzweil and B. Stellmacher, *The theory of finite groups: an introduction*. Springer, 2004, vol. 1.

[18] T. Cohen and M. Welling, “Group equivariant convolutional networks,” in *International conference on machine learning*, PMLR, 2016, pp. 2990–2999.

[19] T. S. Cohen and M. Welling, “Steerable cnns,” *arXiv preprint arXiv:1612.08498*, 2016.

[20] T. S. Cohen, M. Geiger, J. Köhler, and M. Welling, “Spherical cnns,” *arXiv preprint arXiv:1801.10130*, 2018.

[21] N. Thomas, T. Smidt, S. Kearnes, L. Yang, L. Li, K. Kohlhoff, and P. Riley, “Tensor field networks: Rotation-and translation-equivariant neural networks for 3d point clouds,” *arXiv preprint arXiv:1802.08219*, 2018.

[22] M. Geiger and T. Smidt, “E3nn: Euclidean neural networks,” *arXiv preprint arXiv:2207.09453*, 2022.

[23] M. Weiler, M. Geiger, M. Welling, W. Boomsma, and T. S. Cohen, “3d steerable cnns: Learning rotationally equivariant features in volumetric data,” *Advances in Neural Information Processing Systems*, vol. 31, 2018.

[24] H. F. Kay and P. Vousden, “Xcv. symmetry changes in barium titanate at low temperatures and their relation to its ferroelectric properties,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 40, no. 309, pp. 1019–1040, 1949.

[25] M. C. Oliveira, R. A. Ribeiro, E. Longo, M. R. Bomio, F. V. Motta, and S. R. de Lazaro, “Temperature dependence on phase evolution in the batio3 polytypes studied using ab initio calculations,” *International Journal of Quantum Chemistry*, vol. 120, no. 1, e26054, 2020.

[26] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, *et al.*, “Commentary: The materials project: A materials genome approach to accelerating materials innovation,” *APL materials*, vol. 1, no. 1, 2013.

[27] E. Koch and W. Fischer, “Normalizers of point groups,” *International Tables for Crystallography*, pp. 904–905, 2006.

[28] *GAP – Groups, Algorithms, and Programming, Version 4.12.2*, The GAP Group, 2022. URL: https://www.gap-system.org.

Appendix A

Notation and commonly used symbols

Here, we present the notation we use throughout this paper.

Table A.1: Notation used throughout this paper

$\mathrm{Stab}_G(x)$	Stabilizer of an element x under a group G
$N_G(S)$	Normalizer of group S in group G
$\mathrm{Cl}_G(S)$	Set of groups obtained by conjugating group S with elements in G
$\mathrm{Orb}_G(x)$	Orbit of an element x under action by elements of group G
$\mathcal{P}(X)$	Set of all subsets of X
G/S	Set of left cosets. If S is a normal subgroup, this also denotes the quotient group
$S \leq G$	If S and G are groups, this denotes that S is a subgroup of G
$f\|_X$	Function f with domain restricted to X

Table A.2: Commonly used symbols

G	Group our network is equivariant under
e	Identity element of a group
x	Input
y	Output
S	Symmetry of our input, more precisely $\mathrm{Stab}_G(x)$
K	Symmetry of our output, more precisely $\mathrm{Stab}_S(y)$

Appendix B

Classification of full symmetry breaking cases for $O(3)$

Here we tabulate the cases for full symmetry breaking for the finite subgroups of $O(3)$. These are the point groups and the normalizers are tabulated in the International Tables for Crystallography in Hermann–Mauguin notation [27]. We have translated these to Schönflies notation in Table B.1.

Table B.1: Normalizers of the point groups in Schönflies notation. Note we have the equivalences $C_1 = 1$, $S_2 = C_i$, $C_{1h} = C_{1v} = C_s$, $D_1 = C_2$, $D_{1h} = C_{2v}$, $D_{1d} = C_{2h}$.

Normalizer:	Groups:
K_h	1, C_i
$D_{\infty h}$	C_n, S_{2n} $\forall n \geq 2$; C_{nh}, C_s
$D_{(2n)h}$	C_{nv}, D_{nd}, D_{nh} $\forall n \geq 2$; D_n $\forall n \geq 3$
I_h	I, I_h
O_h	D_2, D_{2h}, T, T_d, T_h, O, O_h

In the following subsections we do casework by normalizers. For each, subgroup with a given normalizer, we give a valid complement by name if it exists. In some normalizers, the name of a subgroup is not sufficient to identify it. This is because there are multiple copies of subgroups with that name in the normalizer. In such cases, we must identify which copy of the subgroup we care about. To do so, we give the normalizers in terms of a group presentation found with the help of [28]. Group presentations are essentially a set of generators and relations among the generators. We can then specify any specific subgroups of the normalizer using the generators of the normalizer.

B.1 Normalizer: K_h

All the groups with this normalizer do have complements. Note that in Schönflies notation, K_h is just the entire group $O(3)$. The only subgroups with $O(3)$ as normalizer are the trivial group C_1 and inversion C_i. Clearly for the trivial group the complement is $O(3)$. For inversion, the complement is just $SO(3)$.

Table B.2: Groups with normalizer $K_h = O(3)$ and their complements.

Group	Complement
1	$K_h = O(3)$
C_i	$K = SO(3)$

B.2 Normalizer: $D_{\infty h}$

Unfortunately, none of the groups in this case have complements.

B.3 Normalizer: $D_{(2n)h}$

All groups with this normalizer do have complements. We list the subgroup and its complement in Table B.3. One presentation of $D_{(2n)h}$ is

$$\langle a, b, m | a^{2n}, b^2, m^2, (ab)^2, (am)^2, (bm)^2 \rangle.$$

Figure B.1 depicts an example of a D_{10h} object. The element a correspond to a $2\pi/10$ rotation about the blue vertical axis, b corresponds to a π rotation about the red axis, and

m corresponds to a reflection across the mirror plane shown in orange.

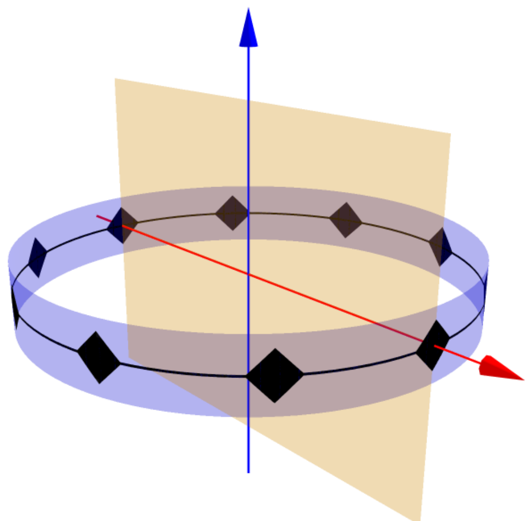


Figure B.1: Object with symmetryD_{10h}. We can identify generator a as the 10-fold rotation about the blue axis, generator b as the 2-fold rotation about the red axis, and m as the reflection over the plane shown in orange.

Table B.3: Groups with normalizer $D_{(2n)h}$ and their complements.

Group	Generators of group	Complement	Generators of a complement
C_{nv}	a^2, m	C_{2v}	am, bm
D_{nd}	a^2, abm, m	C_s	bm
D_{nh}	a^2, b, m	C_s	am
D_n	a^2, b	C_{2v}	am, bm

B.4 Normalizer: I_h

This case is simple, we either have I or I_h. Clearly we just need to add inversion to get a complement in the former case and in the latter case we can just take the trivial group.

Table B.4: Groups with normalizer I_h and their complements.

Group	Complement
I	C_i
I_h	1

B.5 Normalizer: O_h

All subgroups in this case have complements as well. One presentation of O_h is

$$\langle a, b, i | a^4, b^4, i^2, (aba)^2, (ab)^3, iaia^{-1}, ibib^{-1} \rangle.$$

Here, a and b are $\pi/2$ rotations about perpendicular axes and i is just inversion.

Table B.5: Groups with normalizer O_h and their complements.

Group	Generators of group	Complement	Generators of a complement
D_2	a^2, b^2	D_{3d}	ab, ba^2, i
D_{2h}	a^2, b^2, i	D_3	ab, ba^2
T	ab, ba	S_4	a^2b, i
T_d	ab, ba, ai	C_2	a^2b
T_h	ab, ba, i	C_2	a^2b
O	a, b	C_i	i
O_h	a, b, i	1	ϕ

Appendix C

Equivariant full SBS better than exact partial SBS

We provide an outline of the construction of the counterexample. It is easiest to explain this by introducing the concept of a wreath product on groups.

Definition C.1 (Wreath product)**.** *Let H be a group with a group action on some set Ω. Let A be another group. We can define a direct product group indexed by Ω as the set of sequences $(a_\omega)_{\omega\in\Omega}$ where $a_\omega \in A$. The action of H on Ω induces a semidirect product by reindexing. In particular, for all $h \in H$ and sequences in A^Ω we define*

$$h \cdot (a_\omega)_{\omega\in\Omega} = (a_{h^{-1}\omega})_{\omega\in\Omega}.$$

The resulting group is the unrestricted wreath product and denoted as $A \operatorname{Wr}_\Omega H$.

If rather than a direct product group A^Ω, we restrict ourselves to a direct sum where all but finitely many elements in our sequence is not the identity, then we get the restricted wreath product denoted as $A \operatorname{wr}_\Omega H$.

Note that the direct sum and direct product are the same for finite Ω so the restricted and unrestricted wreath products also coincide in those cases.

Consider the space $\Omega = \{1, -1\}$ and an action of D_4 on Ω corresponding to the A_2 representation. Intuitively, if we think of D_4 as the rotational symmetries of a square in the xy-plane, this corresponds to how the z coordinate transforms by flipping signs. Define a group G' as $G' = C_2 \text{ wr}_\Omega D_4$. This is a group of order 32 and is `SmallGroup(32,28)` in the Small Groups library [28]. One presentation of this group is

$$\langle a, b, c | a^2, b^4, (ab)^4, c^2, bcb^{-1}c, (ac)^4 \rangle. \tag{C.1}$$

In this presentation, we can interpret a, b as generators of D_4 and c as the generator of C_2.

Consider the group $G = G' \times G'$ defined using the direct product. We can write generators of G as $a_1, b_1, c_1, a_2, b_2, c_2$ corresponding to two copies of those in the presentation given in (C.1). Define S as the subgroup generated by $a_1, b_1^2, c_1, a_2, b_2^2, c_2$ and K as the subgroup generated by c_1c_2.

We can check that $N_G(S) = N_G(S, K) = G$. It is also not hard to check that a_1b_1, a_2b_2 generate a complement for S in G. Hence, by Theorem 3.4, we know that an ideal G-equivariant SBS is possible for S. Hence, we know the size of the equivariant full SBS is $|S| = 256$.

Next, suppose we wanted a G-equivariant exact partial SBS. We can always generate this partial SBS by taking the orbit of some element p where $\text{Stab}_S(p) = K$ under action by $N_G(S, K) = G$. I claim we must also have $\text{Stab}_G(p) = K$. Suppose not, then there must be some $g \in \text{Stab}_G(p)$ such that $g \notin S$. However, we can check through casework or brute force that for all such g, either $gKg^{-1} \neq K$ or $g^2 \in S - K$. But would mean that there are elemeents not in K which stabilize p so $\text{Stab}_S(p) \neq K$, a contradiction. Hence, no such g can exist so $\text{Stab}_G(p) = K$. Finally, by Orbit-Stabilizer theorem, this means the set must have size $|P| = |\text{Orb}_G(p)| = |G|/|\text{Stab}_G(p)| = (8^2 \cdot 2^4)/2 = 512$. This is larger than the equivariant full SBS.