Brauer's Group Equivariant Neural Networks

Edward Pearce–Crump [1]

Abstract

We provide a full characterisation of all of the possible group equivariant neural networks whose layers are some tensor power of $\mathbb{R}^n$ for three symmetry groups that are missing from the machine learning literature: $O(n)$, the orthogonal group; $SO(n)$, the special orthogonal group; and $Sp(n)$, the symplectic group. In particular, we find a spanning set of matrices for the learnable, linear, equivariant layer functions between such tensor power spaces in the standard basis of $\mathbb{R}^n$ when the group is $O(n)$ or $SO(n)$, and in the symplectic basis of $\mathbb{R}^n$ when the group is $Sp(n)$.

1. Introduction

Finding neural network architectures that are equivariant to a symmetry group has been an active area of research ever since it was first shown how convolutional neural networks, which are equivariant to translations, could be used to learn from images. Unlike with multilayer perceptrons, however, the requirement for the overall network to be equivariant to the symmetry group typically restricts the form of the network itself. Moreover, since these networks exhibit parameter sharing within each layer, ordinarily far fewer parameters appear in these networks than in multilayer perceptrons. This usually results in simpler, more interpretable models that generalise better to unseen data.

Symmetry groups naturally appear in problems coming from physics, where the data that is generated by a physical process often comes with a certain type of symmetry that is baked into the data itself. The data is typically high dimensional, and it can often be represented in the form of a high order tensor so that complex relationships can be captured between different features in the data. Consequently, it is important to be able to construct neural networks that can learn efficiently from such data.

[1]Department of Computing, Imperial College London, United Kingdom. Correspondence to: Edward Pearce–Crump <ep1011@ic.ac.uk>.

Proceedings of the 40^{th} International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023.

There are two approaches that are typically used for constructing group equivariant neural networks. The first employs a universal approximation theorem to learn functions that are approximately equivariant, such as in (Kumagai & Sannai, 2020). The second involves decomposing tensor product representations of the symmetry group in question into irreducible representations. For example, neural networks that are equivariant to the special orthogonal group $SO(3)$ (Kondor et al., 2018), the special Euclidean group $SE(3)$ (Weiler et al., 2018), and the proper orthochronous Lorentz group $SO^+(1,3)$ (Bogatskiy et al., 2020) all use irreducible decompositions and the resulting change of basis transformations into Fourier space in their implementations. However, for most groups, finding this irreducible decomposition is not trivial, since the relevant Clebsch–Gordan coefficients are typically unknown. Furthermore, even if such a decomposition can be found, the resulting neural networks are often inefficient since forward and backward Fourier transforms are usually required to perform the calculations, which come with a high computational cost.

In this paper, we take an entirely different approach, one which results in a full characterisation of all of the possible group equivariant neural networks whose layers are some tensor power of $\mathbb{R}^n$ for the following three symmetry groups: $O(n)$, the orthogonal group; $SO(n)$, the special orthogonal group; and $Sp(n)$, the symplectic group. Our approach is motivated by a mathematical concept that, to the best of our knowledge, has not appeared in any of the machine learning literature to date, other than in (Pearce-Crump, 2022). In that paper, they use that the symmetric group is in Schur–Weyl duality with the partition algebra to provide a full characterisation of all of the possible permutation equivariant neural networks whose layers are some tensor power of $\mathbb{R}^n$. In this paper, we use as our motivation the results given in *On Algebras Which are Connected with the Semisimple Continuous Groups* (Brauer, 1937). Brauer showed that the orthogonal group is in Schur–Weyl duality with an algebra called the Brauer algebra; that the symplectic group is also in Schur–Weyl duality with the Brauer algebra, and that the special orthogonal group is in Schur–Weyl duality with an algebra which we have termed the Brauer–Grood algebra. By adapting the combinatorial diagrams that form a basis for these algebras, we are able to find a spanning set of matrices for the learnable, linear, equivariant layer functions between tensor power spaces of $\mathbb{R}^n$ in the standard basis of $\mathbb{R}^n$ when the group is $O(n)$ or $SO(n)$, and in the symplectic basis of $\mathbb{R}^n$ when the group is $Sp(n)$. In doing so, we avoid having to calculate any irreducible decompositions for the tensor power spaces, and therefore avoid having to perform any Fourier transforms to change the basis accordingly.

The main contributions of this paper are as follows:

1. We are the first to show how the combinatorics underlying the Brauer and Brauer–Grood vector spaces, adapted from the Schur–Weyl dualities established by Brauer (1937), provides the theoretical background for constructing group equivariant neural networks for the orthogonal, special orthogonal, and symplectic groups when the layers are some tensor power of $\mathbb{R}^n$.
2. We find a spanning set of matrices for the learnable, linear, equivariant layer functions between such tensor power spaces in the standard basis of $\mathbb{R}^n$ when the group is $O(n)$ or $SO(n)$, and in the symplectic basis of $\mathbb{R}^n$ when the group is $Sp(n)$.
3. We generalise our diagrammatical approach to show how to construct neural networks that are equivariant to local symmetries.
4. We suggest that Schur–Weyl duality is a powerful mathematical concept that could be used to characterise other group equivariant neural networks beyond those considered in this paper.

2. Preliminaries

We choose our field of scalars to be $\mathbb{R}$ throughout. Tensor products are also taken over $\mathbb{R}$, unless otherwise stated. Also, we let $[n]$ represent the set $\{1, \dots, n\}$.

Recall that a representation of a group G is a choice of vector space V over $\mathbb{R}$ and a group homomorphism

$$\rho : G \to GL(V) \tag{1}$$

We choose to focus on finite-dimensional vector spaces V that are some tensor power of $\mathbb{R}^n$ in this paper.

We often abuse our terminology by calling V a representation of G, even though the representation is technically the homomorphism ρ. When the homomorphism ρ needs to be emphasised alongside its vector space V, we will use the notation (V, ρ).

3. Group Equivariant Neural Networks

Group equivariant neural networks are constructed by alternately composing linear and non-linear G-equivariant maps between representations of a group G. The following is based on the material presented in (Lim & Nelson, 2022).

We first define *G-equivariance*:

Definition 3.1. Suppose that (V, ρ_V) and (W, ρ_W) are two representations of a group G.

A map $\phi : V \to W$ is said to be G-equivariant if, for all $g \in G$ and $v \in V$,

$$\phi(\rho_V(g)[v]) = \rho_W(g)[\phi(v)] \tag{2}$$

The set of all *linear* G-equivariant maps between V and W is denoted by $\mathrm{Hom}_G(V, W)$. When $V = W$, we write this set as $\mathrm{End}_G(V)$. It can be shown that $\mathrm{Hom}_G(V, W)$ is a vector space over $\mathbb{R}$, and that $\mathrm{End}_G(V)$ is an algebra over $\mathbb{R}$. See (Segal, 2014) for more details.

A special case of G-equivariance is *G-invariance*:

Definition 3.2. The map ϕ given in Definition 3.1 is said to be G-invariant if ρ_W is defined to be the 1-dimensional trivial representation of G. As a result, $W = \mathbb{R}$.

We can now define the type of neural network that is the focus of this paper:

Definition 3.3. An L-layer G-equivariant neural network f_{NN} is a composition of *layer functions*

$$f_{NN} \coloneqq f_L \circ \ldots \circ f_l \circ \ldots \circ f_1 \tag{3}$$

such that the l^{th} layer function is a map of representations of G

$$f_l : (V_{l-1}, \rho_{l-1}) \to (V_l, \rho_l) \tag{4}$$

that is itself a composition

$$f_l \coloneqq \sigma_l \circ \phi_l \tag{5}$$

of a learnable, linear, G-equivariant function $\phi_l : (V_{l-1}, \rho_{l-1}) \to (V_l, \rho_l)$ together with a fixed, non-linear activation function $\sigma_l : (V_l, \rho_l) \to (V_l, \rho_l)$ such that

1. σ_l is a G-equivariant map, as in (2), and
2. σ_l acts pointwise (after a basis has been chosen for each copy of V_l in σ_l.)

We focus on the learnable, linear, G-equivariant functions in this paper because the non-linear functions are fixed.

Remark 3.4. The entire neural network f_{NN} is itself a G-equivariant function because it can be shown that the composition of any number of G-equivariant functions is itself G-equivariant.

Remark 3.5. One way of making a neural network of the form given in Definition 3.3 G-invariant is by choosing the representation in the final layer to be the 1-dimensional trivial representation of G.

4. The groups $G = O(n)$, $SO(n)$, and $Sp(n)$

We consider throughout the real vector space $\mathbb{R}^n$.

Let $GL(n)$ be the group of invertible linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$. If we pick a basis for each copy of $\mathbb{R}^n$, then for each linear map in $GL(n)$ we obtain its matrix representation in the bases of $\mathbb{R}^n$ that were chosen. Let $SL(n)$ be the subgroup of $GL(n)$ consisting of all invertible linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$ whose determinant is $+1$.

We can associate to $\mathbb{R}^n$ one of the following two bilinear forms:

1. a non-degenerate, symmetric bilinear form $(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$.
2. a non-degenerate, skew-symmetric bilinear form $\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. In this case, n must be even, say $n = 2m$, as a result of applying Jacobi's Theorem. See page 6 of Goodman and Wallach (2009) for more details.

Then we can define the groups $O(n)$, $SO(n)$, and $Sp(n)$ as follows:

1. $O(n) \coloneqq \left\{ g \in GL(n) \;\middle|\; \begin{array}{c} (gx, gy) = (x, y) \\ \text{for all } x, y \in \mathbb{R}^n \end{array} \right\}$
2. $SO(n) \coloneqq O(n) \cap SL(n)$
3. $Sp(n) \coloneqq \left\{ g \in GL(n) \;\middle|\; \begin{array}{c} \langle gx, gy\rangle = \langle x, y\rangle \\ \text{for all } x, y \in \mathbb{R}^n \end{array} \right\}$, for $n = 2m$.

It can be shown that each of these groups are subgroups of $GL(n)$.

There are special bases of $\mathbb{R}^n$ with respect to each of the forms given above.

Firstly, for the form $(\cdot,\cdot)$, by Lemma 1.1.2 on page 4 of Goodman and Wallach (2009), we may assume that there is an ordered basis

$$B \coloneqq \{e_1, e_2, \dots, e_n\} \tag{6}$$

of $\mathbb{R}^n$, where e_i has a 1 in the i^{th} position, and a 0 elsewhere, which satisfies the relations

$$(e_i, e_j) = \delta_{i,j} \tag{7}$$

with respect to the form $(\cdot,\cdot)$. The basis B is called the *standard basis* for $\mathbb{R}^n$, and by (7), it is an orthonormal basis of $\mathbb{R}^n$. It is clear that the matrix representation of the form $(\cdot,\cdot)$ in the basis B is the $n \times n$ identity matrix.

Hence the form $(\cdot,\cdot)$ in the basis B is the Euclidean inner product

$$(x, y) = x^\top y \text{ for all } x, y \in \mathbb{R}^n \tag{8}$$

where x is the column vector $(x_1, x_2, \dots, x_n)^\top$ and y is the column vector $(y_1, y_2, \dots, y_n)^\top$ when expressed in the basis B.

Secondly, for the form $\langle\cdot,\cdot\rangle$, where $n = 2m$, by Lemma 1.1.5 on page 7 of Goodman and Wallach (2009), we may assume that there is an ordered basis

$$\widetilde{B} \coloneqq \{e_1, e_{1'}, \dots, e_m, e_{m'}\} \tag{9}$$

of $\mathbb{R}^n$, where the i^{th} basis vector in the set has a 1 in the i^{th} position and a 0 elsewhere, which satisfies the relations

$$\langle e_\alpha, e_\beta\rangle = \langle e_{\alpha'}, e_{\beta'}\rangle = 0 \tag{10}$$

$$\langle e_\alpha, e_{\beta'}\rangle = -\langle e_{\alpha'}, e_\beta\rangle = \delta_{\alpha,\beta} \tag{11}$$

with respect to the form $\langle\cdot,\cdot\rangle$. The basis $\widetilde{B}$ is called the *symplectic basis* for $\mathbb{R}^n$.

Hence the form $\langle\cdot,\cdot\rangle$ in the basis $\widetilde{B}$ is the skew product

$$\langle x, y\rangle = \sum_{r=1}^{m} (x_r y_{r'} - x_{r'} y_r) = \sum_{i,j} \epsilon_{i,j} x_i y_j \tag{12}$$

for all $x, y \in \mathbb{R}^n$, where x is the column vector $(x_1, x_{1'}, \dots, x_m, x_{m'})^\top$ expressed in the basis $\widetilde{B}$, y is the column vector $(y_1, y_{1'}, \dots, y_m, y_{m'})^\top$ expressed in the basis $\widetilde{B}$, and

$$\epsilon_{\alpha,\beta} = \epsilon_{\alpha',\beta'} = 0 \tag{13}$$

$$\epsilon_{\alpha,\beta'} = -\epsilon_{\alpha',\beta} = \delta_{\alpha,\beta} \tag{14}$$

5. The space $(\mathbb{R}^n)^{\otimes k}$ as a representation of G

Let G be any of the groups $O(n)$, $SO(n)$, and $Sp(n)$ (where $n = 2m$ for $Sp(n)$).

Then, for any positive integer k, the space $(\mathbb{R}^n)^{\otimes k}$ is a representation of G, denoted by ρ_k, where

$$\rho_k(g)(v_1 \otimes \cdots \otimes v_k) \coloneqq gv_1 \otimes \cdots \otimes gv_k \tag{15}$$

for all $g \in G$ and for all vectors $v_i \in \mathbb{R}^n$. We call $(\mathbb{R}^n)^{\otimes k}$ the k-order tensor power space of $\mathbb{R}^n$.

Moreover, each form $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ induces a non-degenerate bilinear form $(\mathbb{R}^n)^{\otimes k} \times (\mathbb{R}^n)^{\otimes k} \to \mathbb{R}$, given by

$$(v_1 \otimes \cdots \otimes v_k, w_1 \otimes \cdots \otimes w_k) \coloneqq \prod_{r=1}^{k} (v_r, w_r) \tag{16}$$

for the symmetric case, and

$$\langle v_1 \otimes \cdots \otimes v_k, w_1 \otimes \cdots \otimes w_k\rangle \coloneqq \prod_{r=1}^{k} \langle v_r, w_r\rangle \tag{17}$$

for the skew-symmetric case.

Consequently, there is a standard basis for $(\mathbb{R}^n)^{\otimes k}$ that is induced from the standard basis for $\mathbb{R}^n$ for the symmetric case, and, similarly, there is a symplectic basis for $(\mathbb{R}^n)^{\otimes k}$ that is induced from the symplectic basis for $\mathbb{R}^n$ for the skew-symmetric case.

Our goal is to characterise all of the possible learnable, linear G-equivariant layer functions between any two tensor power spaces of $\mathbb{R}^n$. In doing so, we will be able to characterise and implement all of the possible G-equivariant neural networks whose layers are a tensor power space of $\mathbb{R}^n$.

Specifically, we want to find, ideally, a basis, or, at the very least, a spanning set, of matrices for $\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$ when either the standard basis (for $G = O(n), SO(n)$) or the symplectic basis (for $G = Sp(n), n = 2m$) is chosen for $\mathbb{R}^n$.

Note that the G-invariance case is encapsulated within this, since this occurs when $l = 0$.

Remark 5.1. We assume throughout most of the paper that the feature dimension for all of our representations is one. This is because the group G does not act on the feature space. We relax this assumption in Section 7.

Also, the layer functions under consideration do not take into account any bias terms, but we will show in Section 7 that these can be easily introduced.

6. A Spanning Set for $\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$

Brauer's (1937) paper *On Algebras Which are Connected with the Semisimple Continuous Groups* focused, in part, on calculating a spanning set of matrices for $\mathrm{End}_G((\mathbb{R}^n)^{\otimes k})$ in the standard basis of $\mathbb{R}^n$ for the groups $G = O(n)$ and $SO(n)$, and in the symplectic basis of $\mathbb{R}^n$ for $Sp(n)$.

Brauer achieved this by applying the First Fundamental Theorem of Invariant Theory for each of the groups in question (see the Technical Appendix) to a spanning set of invariants, one for each group G, that he showed are in bijective correspondence with the spanning set of matrices for $\mathrm{End}_G((\mathbb{R}^n)^{\otimes k})$.

In particular, for each group G, he associated a diagram with each element of the spanning set of invariants, which we describe in further detail below. In doing so, he constructed a bijective correspondence between a set of such diagrams and a spanning set of matrices for $\mathrm{End}_G((\mathbb{R}^n)^{\otimes k})$ in the standard/symplectic basis of $\mathbb{R}^n$.

We want to find, instead, a spanning set of matrices for $\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$ in the standard/symplectic basis of $\mathbb{R}^n$. The results that we describe in the following will contain, as a special case, Brauer's results, since when $l = k$,

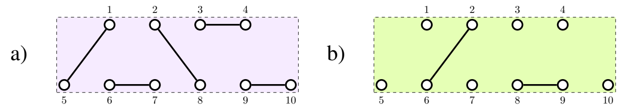


Figure 1. a) The diagram d_β corresponding to the $(6, 4)$–Brauer partition β given in (19). b) The diagram d_α corresponding to the $(4 + 6)\backslash 6$–partition α given in (20).

we see that

$$\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l}) = \mathrm{End}_G((\mathbb{R}^n)^{\otimes k}) \tag{18}$$

Brauer considered two types of diagrams coming from certain set partitions, whose definitions we adapt for our purposes. The first is defined as follows:

Definition 6.1. For any $k, l \in \mathbb{Z}_{\geq 0}$, a (k, l)–Brauer partition β is a partitioning of the set $[l + k]$ into a disjoint union of pairs. We call each pair a block. Clearly, if $l + k$ is odd, then no such partitions exist. Therefore, in the following, we assume that $l + k$ is even. By convention, if $k = l = 0$, then there is just one $(0, 0)$–Brauer partition.

We can represent each (k, l)–Brauer partition β by a diagram d_β, called a (k, l)–Brauer diagram, consisting of two rows of vertices and edges between vertices such that there are 1) l vertices in the top row, labelled left to right by $1, \dots, l$; 2) k vertices in the bottom row, labelled left to right by $l + 1, \dots, l + k$; and 3) the edges between vertices correspond to the blocks of β. In particular, this means that each vertex is incident to exactly one edge; hence there are precisely $\frac{l+k}{2}$ edges in total in d_β.

It is clear that the number of (k, l)–Brauer diagrams, if $l + k$ is even, is $(l + k - 1)!! := (l + k - 1)(l + k - 3) \cdots 5 \cdot 3 \cdot 1$, and is 0 otherwise.

For example, we see that

$$\beta := \{1, 5 \mid 2, 8 \mid 3, 4 \mid 6, 7 \mid 9, 10\} \tag{19}$$

is a valid $(6, 4)$–Brauer partition. Figure 1a) shows the $(6, 4)$–Brauer diagram d_β corresponding to β.

The second type of diagram is what we have decided to call an $(l + k)\backslash n$–diagram (pronounced "l *plus* k *without* n"). These diagrams were originally hinted at by Brauer (1937) in the case where $l = k$ and were looked at in greater detail by Grood (1999), but again only in the case where $l = k$. In their paper, they called these diagrams $k\backslash m$–diagrams, since they only considered the situation where $l = k$ and $n = 2m$. We will see that the definition below is equivalent to their definition in this case. Our naming convention for the diagrams makes more sense since they are a generalisation of those considered by Brauer and Grood.

Definition 6.2. For any k, l and $n \in \mathbb{Z}_{\geq 0}$, an $(l + k)\backslash n$–partition is a partitioning of the set $[l + k]$ with some n

elements removed into a disjoint union of pairs. Once again, each pair is called a block.

An $(l+k)\backslash n$–diagram is the representation of an $(l+k)\backslash n$–partition in its diagram form, constructed in a similar way to the (k,l)–Brauer diagrams above, where there is still a top row consisting of l vertices and a bottom row consisting of k vertices, but now there are only $\frac{l+k-n}{2}$ edges between pairs of vertices. In this case, the n vertices removed from the set $[l+k]$ will not be incident to any edge, and an edge exists between any two vertices that are in the same pair. We call the n vertices whose labels have been removed from $[l+k]$ *free* vertices.

It is clear that if $n > l+k$, then no $(l+k)\backslash n$–diagrams exist. Also, no such diagrams exist if n is odd and $l+k$ is even, or if n is even and $l+k$ is odd.

Otherwise, that is, if $n \leq l+k$ and either n is even and $l+k$ is even, or n is odd and $l+k$ is odd, then the number of $(l+k)\backslash n$–diagrams is $\binom{l+k}{n}(l+k-n-1)!!$, since there are $\binom{l+k}{n}$ ways to pick n free vertices, and $(l+k-n-1)!!$ ways to pair up the remaining $l+k-n$ vertices.

For example, we see that, if $k=6$ and $l=4$, then

$$\alpha := \{2, 6 \mid 8, 9\} \tag{20}$$

is a $(4+6)\backslash 6$–partition. Figure 1b) shows the $(4+6)\backslash 6$–diagram d_α corresponding to α.

Remark 6.3. We choose throughout to focus on (k,l)–Brauer and $(l+k)\backslash n$*–diagrams* over their equivalent set partition form. This is because both the diagrams and the matrices that they correspond to have matching shapes. In fact, it will become clear that, using these diagrams, we can view the matrix multiplication of a spanning set element in $\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$ with an input vector in $(\mathbb{R}^n)^{\otimes k}$ as a process represented by the corresponding diagram, where the input vector is passed into the bottom row of the diagram, and an output vector is returned from the top row.

From the two types of diagrams defined above, we form the following two vector spaces.

Definition 6.4. We define the Brauer vector space, $B_k^l(n)$, which exists for any integer $n \in \mathbb{Z}_{\geq 1}$ and for any $k, l \in \mathbb{Z}_{\geq 0}$, as follows. Let $B_0^0(n) := \mathbb{R}$. Otherwise, define $B_k^l(n)$ to be the $\mathbb{R}$-linear span of the set of all (k,l)–Brauer diagrams.

Definition 6.5. We define the Brauer–Grood vector space, $D_k^l(n)$, which exists for any integer $n \in \mathbb{Z}_{\geq 1}$ and for any $k, l \in \mathbb{Z}_{\geq 0}$, as follows. Let $D_0^0(n) := \mathbb{R}$. Otherwise, define $D_k^l(n)$ to be the $\mathbb{R}$-linear span of the set of all (k,l)–Brauer diagrams together with the set of all $(l+k)\backslash n$–diagrams.

Clearly, if $n > l+k$, or if n is odd and $l+k$ is even, or if n is even and $l+k$ is odd, then $D_k^l(n) = B_k^l(n)$.

With these two vector spaces, we are now able to find a spanning set of matrices for $\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$ in the standard basis (for $G = O(n), SO(n)$) or the symplectic basis (for $G = Sp(n), n = 2m$) of $\mathbb{R}^n$.

In order to explicitly state what these spanning sets are, we note that, for any $k, l \in \mathbb{Z}_{\geq 0}$, as a result of picking the standard/symplectic basis for $\mathbb{R}^n$, the vector space $\mathrm{Hom}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$ has a standard basis of matrix units

$$\{E_{I,J}\}_{I \in [n]^l, J \in [n]^k} \tag{21}$$

where I is a tuple $(i_1, i_2, \dots, i_l) \in [n]^l$, J is a tuple $(j_1, j_2, \dots, j_k) \in [n]^k$ and $E_{I,J}$ has a 1 in the (I,J) position and is 0 elsewhere. If one or both of k, l is equal to 0, then we replace the tuple that indexes the matrix by a 1. For example, when $k=0$ and $l \in \mathbb{Z}_{\geq 1}$, (21) becomes $\{E_{I,1}\}_{I \in [n]^l}$.

We obtain the following results, which are given in the following three theorems.

Theorem 6.6 (Spanning set when $G = O(n)$)**.** *For any $k, l \in \mathbb{Z}_{\geq 0}$ and any $n \in \mathbb{Z}_{\geq 1}$, there is a surjection of vector spaces*

$$\Phi_{k,n}^l : B_k^l(n) \to \mathrm{Hom}_{O(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l}) \tag{22}$$

which is defined as follows.

If $l+k$ is odd, then we map the empty set onto the empty set. Hence, in this case, $\mathrm{Hom}_{O(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l}) = \varnothing$.

Otherwise, for any $k, l \in \mathbb{Z}_{\geq 0}$ and for each (k,l)–Brauer diagram d_β, associate the indices $i_1, i_2, \dots, i_l$ with the vertices in the top row of d_β, and $j_1, j_2, \dots, j_k$ with the vertices in the bottom row of d_β. Then, for any $n \in \mathbb{Z}_{\geq 1}$, define

$$E_\beta := \sum_{I \in [n]^l, J \in [n]^k} \delta_{r_1,u_1}\delta_{r_2,u_2} \dots \delta_{r_{\frac{l+k}{2}},u_{\frac{l+k}{2}}} E_{I,J} \tag{23}$$

where $r_1, u_1, \dots, r_{\frac{l+k}{2}}, u_{\frac{l+k}{2}}$ is any permutation of the indices $i_1, i_2, \dots, i_l, j_1, j_2, \dots, j_k$ such that the vertices corresponding to r_p, u_p are in the same block of β.

The adapted version of Brauer's Invariant Argument, given in the Technical Appendix, shows that (23) defines a bijective correspondence between the set of all (k,l)–Brauer diagrams and a spanning set for $\mathrm{Hom}_{O(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$.

Consequently, when $l+k$ is even, the surjection of vector spaces given in (22) is defined by

$$d_\beta \mapsto E_\beta \tag{24}$$

for all (k,l)–Brauer diagrams d_β, and is extended linearly on the basis of such diagrams for $B_k^l(n)$.

Hence the set

$$\{E_\beta \mid d_\beta \text{ is a } (k,l)\text{–Brauer diagram}\} \tag{25}$$

is a spanning set for $\mathrm{Hom}_{O(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$ *in the standard basis of* $\mathbb{R}^n$*, of size* 0 *when* $l+k$ *is odd, and of size* $(l+k-1)!!$ *when* $l+k$ *is even.*

Lehrer and Zhang (2012) showed that when $2n \geq l+k$*,* $\Phi_{k,n}^l$ *is an isomorphism of vector spaces, and so the set (25) forms a basis of* $\mathrm{Hom}_{O(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$ *in this case.*

Theorem 6.7 (Spanning set when $G = Sp(n), n = 2m$)**.** *For any* $k, l \in \mathbb{Z}_{\geq 0}$ *and any* $n \in \mathbb{Z}_{\geq 2}$ *such that* $n = 2m$*, there is a surjection of vector spaces*

$$X_{k,n}^l : B_k^l(n) \to \mathrm{Hom}_{Sp(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l}) \tag{26}$$

which is defined as follows.

If $l+k$ *is odd, then we map the empty set onto the empty set. Hence, in this case,* $\mathrm{Hom}_{Sp(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l}) = \varnothing$*.*

Otherwise, for any $k, l \in \mathbb{Z}_{\geq 0}$ *and for each* (k,l)*–Brauer diagram* d_β*, associate the indices* $i_1, i_2, \dots, i_l$ *with the vertices in the top row of* d_β*, and* $j_1, j_2, \dots, j_k$ *with the vertices in the bottom row of* d_β*. Then, for any* $n \in \mathbb{Z}_{\geq 2}$*, define*

$$F_\beta \coloneqq \sum_{I,J} \gamma_{r_1,u_1}\gamma_{r_2,u_2}\dots\gamma_{r_{\frac{l+k}{2}},u_{\frac{l+k}{2}}} E_{I,J} \tag{27}$$

where the indices i_p, j_p *range over* $1, 1', \dots, m, m'$*, where* $r_1, u_1, \dots, r_{\frac{l+k}{2}}, u_{\frac{l+k}{2}}$ *is any permutation of the indices* $i_1, i_2, \dots, i_l, j_1, j_2, \dots, j_k$ *such that the vertices corresponding to* r_p, u_p *are in the same block of* β*, and*

$$\gamma_{r_p,u_p} \coloneqq \begin{cases} \delta_{r_p,u_p} & \text{if the vertices corresponding to } r_p, u_p \text{ are in different rows of } d_\beta \\ \epsilon_{r_p,u_p} & \text{if the vertices corresponding to } r_p, u_p \text{ are in the same row of } d_\beta \end{cases} \tag{28}$$

where ϵ_{r_p,u_p} *was defined in (13) and (14).*

The adapted version of Brauer's Invariant Argument, given in the Technical Appendix, shows that (27) defines a bijective correspondence between the set of all (k,l)*–Brauer diagrams and a spanning set for* $\mathrm{Hom}_{Sp(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$*.*

Consequently, when $l+k$ *is even, the surjection of vector spaces given in (26) is defined by*

$$d_\beta \mapsto F_\beta \tag{29}$$

for all (k,l)*–Brauer diagrams* d_β*, and is extended linearly on the basis of such diagrams for* $B_k^l(n)$*.*

Hence the set

$$\{F_\beta \mid d_\beta \text{ is a } (k,l)\text{–Brauer diagram}\} \tag{30}$$

is a spanning set for $\mathrm{Hom}_{Sp(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$*, for* $n = 2m$*, in the symplectic basis of* $\mathbb{R}^n$*, of size* 0 *when* $l+k$ *is odd, and of size* $(l+k-1)!!$ *when* $l+k$ *is even.*

Lehrer and Zhang (2012) showed that when $n \geq l+k$*,* $X_{k,n}^l$ *is an isomorphism of vector spaces, and so the set (30) forms a basis of* $\mathrm{Hom}_{Sp(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$*, for* $n = 2m$*, in this case.*

Theorem 6.8 (Spanning set when $G = SO(n)$)**.** *For any* $k, l \in \mathbb{Z}_{\geq 0}$ *and any* $n \in \mathbb{Z}_{\geq 1}$*, we construct a surjection of vector spaces*

$$\Psi_{k,n}^l : D_k^l(n) \to \mathrm{Hom}_{SO(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l}) \tag{31}$$

as follows.

If $n > l+k$*, or if* n *is odd and* $l+k$ *is even, or if* n *is even and* $l+k$ *is odd, then we saw that* $D_k^l(n) = B_k^l(n)$*. Hence, in these cases,* $\Psi_{k,n}^l = \Phi_{k,n}^l$*, and so* $\mathrm{Hom}_{SO(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l}) = \mathrm{Hom}_{O(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$*.*

Otherwise, that is, if $n \leq l+k$*, and either* n *is even and* $l+k$ *is even, or* n *is odd and* $l+k$ *is odd, there exist* $(l+k)\backslash n$*–diagrams.*

For each such diagram d_α*, again associate the indices* $i_1, i_2, \dots, i_l$ *with the vertices in the top row of* d_α*, and* $j_1, j_2, \dots, j_k$ *with the vertices in the bottom row of* d_α*. Suppose that there are* s *free vertices in the top row. Then there are* $n-s$ *free vertices in the bottom row. Relabel the* s *free indices in the top row (from left to right) by* $t_1, \dots, t_s$*, and the* $n-s$ *free indices in the bottom row (from left to right) by* $b_1, \dots, b_{n-s}$*.*

Then, define $\chi\begin{pmatrix} 1 & 2 & \cdots & s & s+1 & \cdots & n \\ t_1 & t_2 & \cdots & t_s & b_1 & \cdots & b_{n-s}\end{pmatrix}$ *as follows: it is* 0 *if the elements* $t_1, \dots, t_s, b_1, \dots, b_{n-s}$ *are not distinct, otherwise, it is* $\mathrm{sgn}\begin{pmatrix} 1 & 2 & \cdots & s & s+1 & \cdots & n \\ t_1 & t_2 & \cdots & t_s & b_1 & \cdots & b_{n-s}\end{pmatrix}$*, considered as a permutation of* $[n]$*.*

As a result, for any $n \in \mathbb{Z}_{\geq 1}$*, define* H_α *to be*

$$\sum_{I \in [n]^l, J \in [n]^k} \chi\begin{pmatrix} 1 & 2 & \cdots & s & s+1 & \cdots & n \\ t_1 & t_2 & \cdots & t_s & b_1 & \cdots & b_{n-s}\end{pmatrix} \delta(r,u) E_{I,J} \tag{32}$$

where

$$\delta(r,u) \coloneqq \delta_{r_1,u_1}\delta_{r_2,u_2}\dots\delta_{r_{\frac{l+k-n}{2}},u_{\frac{l+k-n}{2}}} \tag{33}$$

Here, $r_1, u_1, \dots, r_{\frac{l+k-n}{2}}, u_{\frac{l+k-n}{2}}$ *is any permutation of the indices*

$$\{i_1, \dots, i_l, j_1, \dots, j_k\} \backslash \{t_1, \dots, t_s, b_1, \dots, b_{n-s}\} \tag{34}$$

such that the vertices corresponding to r_p, u_p *are in the same block of* α*.*

The adapted version of Brauer's Invariant Argument, given in the Technical Appendix, shows that (23) and (32) defines a bijective correspondence between the set of all (k,l)*–Brauer diagrams together with the set* *of all $(l+k)\backslash n$–diagrams, and a spanning set for* $\mathrm{Hom}_{SO(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$.

Consequently, when $n \leq l+k$ and n is even and $l+k$ is even, the surjection of vector spaces (31) is given by

$$d_\beta \mapsto E_\beta \tag{35}$$

if d_β is a (k,l)–Brauer diagram, where E_β was defined in Theorem 6.6, and by

$$d_\alpha \mapsto H_\alpha \tag{36}$$

if d_α is an $(l+k)\backslash n$–diagram, and is extended linearly on the basis of such diagrams for $D_k^l(n)$.

When $n \leq l+k$ and n is odd and $l+k$ is odd, the surjection of vector spaces (31) is given solely by (36), since no (k,l)–Brauer diagrams exist in this case.

Hence, in all cases, the set

$$\{E_\beta\}_\beta \cup \{H_\alpha\}_\alpha \tag{37}$$

where d_β is a (k,l)–Brauer diagram and d_α is an $(l+k)\backslash n$–diagram, is a spanning set for $\mathrm{Hom}_{SO(n)}((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$ *in the standard basis of* $\mathbb{R}^n$.

7. Adding Features and Biases

7.1. Features

In Section 6, we made the assumption that the feature dimension for all of the layers appearing in the neural network was one. This simplified the analysis for the results seen in that section. All of these results can be adapted for the case where the feature dimension of the layers is greater than 1.

Suppose that an r-order tensor has a feature space of dimension d_r. We now wish to find a spanning set for

$$\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k} \otimes \mathbb{R}^{d_k}, (\mathbb{R}^n)^{\otimes l} \otimes \mathbb{R}^{d_l}) \tag{38}$$

in the standard basis of $\mathbb{R}^n$ for $G = O(n)$ and $SO(n)$, and in the symplectic basis of $\mathbb{R}^n$ for $G = Sp(n)$, where $n = 2m$.

A spanning set in each case can be found by making the following substitutions in Theorems 6.6, 6.7, and 6.8 of Section 6, where now $i \in [d_l]$ and $j \in [d_k]$:

- replace $E_{I,J}$ by $E_{I,i,J,j}$, $E_{I,1}$ by $E_{I,i,1,j}$, $E_{1,J}$ by $E_{1,i,J,j}$, and $E_{1,1}$ by $E_{1,i,1,j}$, and
- relabel E_β by $E_{\beta,i,j}$, F_β by $F_{\beta,i,j}$, and H_α by $H_{\alpha,i,j}$.

Consequently, a spanning set for (38) in the standard/symplectic basis of $\mathbb{R}^n$, is given by

- $\{E_{\beta,i,j}\}_{\beta,i,j}$ for $G = O(n)$,
- $\{F_{\beta,i,j}\}_{\beta,i,j}$ for $G = Sp(n)$ (where $n = 2m$), and
- $\{E_{\beta,i,j}\}_{\beta,i,j} \cup \{H_{\alpha,i,j}\}_{\alpha,i,j}$ for $G = SO(n)$,

where d_α is an $(l+k)\backslash n$–diagram, d_β is a (k,l)–Brauer diagram, $i \in [d_l]$, and $j \in [d_k]$.

7.2. Biases

Including bias terms in the layer functions of a G-equivariant neural network is harder, but it can be done. For the learnable linear layers of the form $\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$, Pearce–Crump (2022) shows that the G-equivariance of the bias function, $\beta : ((\mathbb{R}^n)^{\otimes k}, \rho_k) \rightarrow ((\mathbb{R}^n)^{\otimes l}, \rho_l)$, needs to satisfy

$$c = \rho_l(g)c \tag{39}$$

for all $g \in G$ and $c \in (\mathbb{R}^n)^{\otimes l}$.

Since any $c \in (\mathbb{R}^n)^{\otimes l}$ satisfying (39) can be viewed as an element of $\mathrm{Hom}_G(\mathbb{R}, (\mathbb{R}^n)^{\otimes l})$, to find the matrix form of c, all we need to do is to find a spanning set for $\mathrm{Hom}_G(\mathbb{R}, (\mathbb{R}^n)^{\otimes l})$.

But this is simply a matter of applying the results of Section 6, namely Theorem 6.6 for $G = O(n)$, Theorem 6.7 for $G = Sp(n)$, with $n = 2m$, and Theorem 6.8 for $G = SO(n)$, setting $k = 0$.

8. Equivariance to Local Symmetries

We can extend our results to looking at linear layer functions that are equivariant to a direct product of groups. In this scenario, the data is given on a collection of some p sets of sizes $n_1, \dots, n_p$, and we require equivariance to the group $G(n_i)$ for the data given on the i^{th} set. Neural networks that are constructed using these layer functions are said to be equivariant to local symmetries.

Specifically, we wish to find a spanning set for

$$\mathrm{Hom}_{G(n_1) \times \cdots \times G(n_p)}(V, W) \tag{40}$$

where

$$V := (\mathbb{R}^{n_1})^{\otimes k_1} \boxtimes \cdots \boxtimes (\mathbb{R}^{n_p})^{\otimes k_p} \tag{41}$$

$$W := (\mathbb{R}^{n_1})^{\otimes l_1} \boxtimes \cdots \boxtimes (\mathbb{R}^{n_p})^{\otimes l_p} \tag{42}$$

and $\boxtimes$ is the external tensor product.

The Hom-space given in (40) is isomorphic to

$$\bigotimes_{r=1}^{p} \mathrm{Hom}_{G(n_r)}((\mathbb{R}^{n_r})^{\otimes k_r}, (\mathbb{R}^{n_r})^{\otimes l_r}) \tag{43}$$

Consequently, we can construct a surjection of vector spaces, denoted by $\bigotimes_{r=1}^{p} \Theta_{k_r,n_r}^{l_r}$, from $\bigotimes_{r=1}^{p} A_{k_r}^{l_r}(n_r)$ to the Hom–space given in (43), where

- if $G(n_r) = O(n_r)$, then $\Theta_{k_r,n_r}^{l_r} = \Phi_{k_r,n_r}^{l_r}$ and $A_{k_r}^{l_r}(n_r) = B_{k_r}^{l_r}(n_r)$,
- if $G(n_r) = Sp(n_r)$, then $\Theta_{k_r,n_r}^{l_r} = X_{k_r,n_r}^{l_r}$ and $A_{k_r}^{l_r}(n_r) = B_{k_r}^{l_r}(n_r)$ (here $n_r = 2m_r$), and
- if $G(n_r) = SO(n_r)$, then $\Theta_{k_r,n_r}^{l_r} = \Psi_{k_r,n_r}^{l_r}$ and $A_{k_r}^{l_r}(n_r) = D_{k_r}^{l_r}(n_r)$.

As a result, a spanning set for (40) can be found by placing each possible basis diagram for each of the vector spaces $A_{k_r}^{l_r}(n_r)$ side by side, taking the image of each under its map $\Theta_{k_r,n_r}^{l_r}$, and then calculating the Kronecker product of the resulting matrices.

Features and biases can be added in exactly the same way as discussed in Section 7.

9. Related Work

The combinatorial representation theory of the Brauer algebra was developed by Brauer (1937) for the purpose of understanding the centraliser algebras of the groups $O(n)$, $SO(n)$ and $Sp(n)$. Brown published two papers (1955; 1956) on the Brauer algebra, showing that it is semisimple if and only if $n \geq k-1$. Weyl (1946) had previously shown that the Brauer algebra was semisimple if $n \geq 2k$. After Brown's papers, the Brauer algebra was largely forgotten about until Hanlon and Wales (1989) provided an isomorphism between two versions of the Brauer algebra – these versions share a common basis but have different algebra products defined on them. Grood (1999) investigated the representation theory of what we have termed the Brauer–Grood algebra. Lehrer and Zhang (2012) studied the kernel of the surjection of algebras given in Theorems 6.6 and 6.7 when $l = k$, showing that, in each case, it is a two-sided ideal generated by a single element of the Brauer algebra.

With regard to the machine learning literature, Maron et al.'s paper (2019) is the closest to ours in terms of how it motivated our research idea. They characterised all of the learnable, linear, equivariant layer functions when the layers are some tensor power of $\mathbb{R}^n$ for the symmetric group S_n in the practical cases (specifically, when $n \geq k + l$). Pearce–Crump (2022) used the Schur–Weyl duality between the symmetric group and the partition algebra to provide a full characterisation for these layer functions for all values of n and for all orders of the tensor power spaces involved, showing, in particular, that the dimension of the space of layer functions is not independent of n. Finzi et al. (2021) were the first to recognise that the dimension is not independent of n, using a numerical algorithm to calculate the correct values for small values of n, k and l. Their numerical algorithm also enabled them to find a basis for the learnable, layer, equivariant functions for the groups that are the focus of our study, namely $O(n)$, $Sp(n)$ and $SO(n)$, but only for small values of n, k and l, since their algorithm runs out of memory on higher values. In this paper, whilst we have not found a basis in all cases, we have provided a spanning set and an analytic solution for all values of n, k and l, which will make it possible to implement group equivariant neural networks for any such values of n, k and l for the three groups in question. In writing up this paper, we came across a paper by Villar et al. (2021), in which they focus on designing group equivariant neural networks for $O(n)$ and $SO(n)$, amongst others. They also use the First Fundamental Theorem of Invariant Theory for $O(n)$ and $SO(n)$, but only to characterise all invariant scalar functions $(\mathbb{R}^n)^{\otimes k} \rightarrow \mathbb{R}$ and all equivariant vector functions $(\mathbb{R}^n)^{\otimes k} \rightarrow \mathbb{R}^n$ as a sum involving $O(n)$ or $SO(n)$ invariant scalar functions. They then use multilayer perceptrons to learn these invariant scalar functions. Our method, by contrast, characterises a wider selection of functions, since we study the linear, learnable, equivariant functions between layers that are any tensor power of $\mathbb{R}^n$ for $O(n)$ and $SO(n)$ (as well as for $Sp(n)$), and we give the exact matrix form of these functions in the standard basis of $\mathbb{R}^n$, meaning that the group equivariant neural network architectures that we provide are exact and do not require any approximations via multilayer perceptrons.

10. Limitations, Discussion, and Future Work

We believe that our approach for characterising all of the possible $O(n)$, $Sp(n)$ and $SO(n)$ equivariant neural networks whose layers are some tensor power of $\mathbb{R}^n$ is promising in terms of the possible practical applications. A number of papers have appeared recently in the literature where the authors tried to learn group equivariant functions – for the groups given in this paper – on tensor power spaces of $\mathbb{R}^n$. However, they were not especially successful in their attempts. We believe that this is a consequence of them using architectures that approximate the group equivariance property of the functions that they wish to learn, rather than guaranteeing it exactly. The results in this paper directly address this problem. In Finkelshtein et al. (2022), the authors created a tensor product neural network which was approximately equivariant to $O(3) \times S_n$ in order to learn from point cloud data. By combining the results of Pearce–Crump (2022) for S_n with the results in this paper for $O(3)$, we would be able to replace the linear layers in their architecture with exact, parameterizable matrices for the tensor product spaces that are guaranteed to be equivariant to $O(3) \times S_n$. We believe that this could improve the final outcome. In addition, in Villar et al. (2021), the authors explore two numerical experiments involving tensor power representations: the first is an $O(5)$-invariant task from an order 2 tensor power of $\mathbb{R}^5$ to $\mathbb{R}$, and the second is an $S_5 \times O(3)$ task where the $O(3)$ component is equivariant on order 5 tensors of $\mathbb{R}^3$. The authors, however, use standard multi-layer perceptrons to learn the functions. We think that they could improve upon their results by using the linear layers that are characterised in this paper. Furthermore, we are of the opinion that by characterising the equivariant neural networks for these groups, we have made it possible for other researchers in the machine learning community to find further uses for these neural networks.

We are aware that equivariance to the symplectic group $Sp(n)$ does not commonly appear in the machine learning literature. However, as stated in Appendix E.6 of Finzi et al. (2021), $Sp(n)$ equivariance is especially relevant in the context of Hamiltonian mechanics and classical physics. Section 7.2 of Finzi et al. (2021) points to the paper by Greydanus et al. (2019), where the authors look to learn the Hamiltonian of a system coming from Hamiltonian mechanics. In particular, the time evolution of the system is expressed in terms of the symplectic basis. The paper by Villar et al. (2021) also highlights how there are many symmetries in physics that are relevant for the machine learning community, including "symplectic symmetry that permits reinterpretations of positions and momenta".

We appreciate that given the current state of hardware, there will be some computational limitations when implementing the neural networks that appear in this paper in practice, and some engineering may be required to obtain the necessary scale. In particular, it is not a trivial task to store the high order tensors that appear in such neural networks. This was demonstrated by Kondor et al. (2018), where the authors needed to develop custom CUDA kernels in order to implement their tensor product based neural networks. In saying that, however, we feel that given the ever-increasing availability of computing power, we should see higher-order group equivariant neural networks appear more often in practice. We also note that while the tensor power spaces increase exponentially in dimension as we increase their order, the dimension of the space of equivariant maps between such tensor spaces is typically much smaller, and the matrices themselves are often sparse. Hence, while there may be some technical difficulties in storing such matrices, it should be possible to do so with the current computing power that is available.

We recognise, however, that further work is required in order to demonstrate fully the practical applicability of our results. In particular, we need to conduct practical experiments to assess how these neural networks perform when the order of the tensor power spaces is increased, and we need to show that these neural networks provide sufficient practical advantages over the existing approaches that use irreducible decompositions.

11. Conclusion

We are the first to show how the combinatorics underlying the Brauer and Brauer–Grood vector spaces provides the theoretical background for constructing group equivariant neural networks for the orthogonal, special orthogonal, and symplectic groups when the layers are some tensor power of $\mathbb{R}^n$. We looked at the problem of calculating the matrix form of the linear layer functions between such spaces in the standard/symplectic basis for $\mathbb{R}^n$. We recognised that a solution to this problem would provide a powerful method for constructing group equivariant neural networks for the three groups in question since we could avoid having to solve the much more difficult problem of decomposing the tensor power spaces of $\mathbb{R}^n$ into their irreducible representations and then avoid having to find the change of basis matrix that would be needed to perform the layer mappings.

We saw how a basis of diagrams for the Brauer and Brauer–Grood vector spaces enabled us to find a spanning set of matrices for the layer functions themselves in the standard/symplectic basis for $\mathbb{R}^n$ for each of the three groups in question, and how each diagram provided the parameter sharing scheme for its image in the spanning set. As a result, we were able to characterise all of the possible group equivariant neural networks whose layers are some tensor power of $\mathbb{R}^n$ for each of the three groups in question. We were also able to generalise this diagrammatic approach to layer functions that were equivariant to local symmetries.

As discussed in the Introduction, our results were motivated by Brauer (1937) who showed that there exists a Schur–Weyl duality for each of the three groups in question with an algebra of diagrams. Consequently, this leads to the following question, which is one for future research: what other Schur–Weyl dualities exist between a group and an algebra of diagrams that would enable us to understand the structure of neural networks that are equivariant to the group?

Acknowledgements

The author would like to thank his PhD supervisor Professor William J. Knottenbelt for being generous with his time throughout the author's period of research prior to the publication of this paper.

This work was funded by the Doctoral Scholarship for Applied Research which was awarded to the author under Imperial College London's Department of Computing Applied Research scheme. This work will form part of the author's PhD thesis at Imperial College London.

References

Barcelo, H. and Ram, A. Combinatorial Representation Theory, 1997. arXiv:math/9707221.

Benkart, G. Commuting Actions – a Tale of Two Groups. In *Lie Algebras and Their Representations: A Symposium on Lie Algebras and Representation Theory, January 23 - 27, 1995, Seoul National University, Seoul, Korea*, volume 194, pp. 1–46, 1996. URL dx.doi.og/10.1090/conm/194/02387.

Benkart, G. Connecting the McKay correspondence and Schur–Weyl duality. In *Proceedings of the International Congress of Mathematicians, August 13 - 21, 2014, Seoul, Korea*, volume I, pp. 633–656, 2014.

Benkart, G. and Halverson, T. Partition algebras $P_k(n)$ with $2k > n$ and the fundamental theorems of invariant theory for the symmetric group S_n. *J. London Math. Soc*, 99 (2):194–224, 2019a. URL https://doi.org/10.1112/jlms.12175.

Benkart, G. and Halverson, T. Partition Algebras and the Invariant Theory of the Symmetric Group. In Barcelo, H., Karaali, G., and Orellana, R. (eds.), *Recent Trends in Algebraic Combinatorics*, volume 16 of *Association for Women in Mathematics Series*, pp. 1–41. Springer, 2019b. URL https://doi.org/10.1007/978-3-030-05141-9.

Benkart, G. and Moon, D. Tensor Product Representations of Temperley–Lieb Algebras and Chebyshev Polynomials. In *The Tenth International Conference on Representations of Algebras and Related Topics, July 15 – August 10, 2002, Fields Institute, Toronto, Canada*, volume 45, pp. 57–80, 2005.

Bogatskiy, A., Anderson, B., Offermann, J. T., Roussi, M., Miller, D. W., and Kondor, R. Lorentz Group Equivariant Neural Network for Particle Physics. In *Proceedings of the 37th International Conference on Machine Learning*, ICML'20. JMLR.org, 2020. URL https://proceedings.mlr.press/v119/bogatskiy20a/bogatskiy20a.pdf.

Brauer, R. On Algebras Which Are Connected with the Semisimple Continuous Groups. *Ann. Math.*, 38:857–872, 1937. ISSN 0003486X. URL http://www.jstor.org/stable/1968843.

Brown, W. P. An Algebra Related to the Orthogonal Group. *Michigan Mathematical Journal*, 3(1):1 – 22, 1955. URL https://doi.org/10.1307/mmj/1031710528.

Brown, W. P. The Semisimplicity of the Brauer Algebra. *Annals of Mathematics*, 63(2):324–335, 1956. URL https://www.jstor.org/stable/1969613.

Cohen, T. and Welling, M. Group Equivariant Convolutional Networks. In Balcan, M. F. and Weinberger, K. Q. (eds.), *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 2990–2999, New York, New York, USA, 20–22 Jun 2016. PMLR. URL https://proceedings.mlr.press/v48/cohenc16.html.

Cohen, T. S., Geiger, M., Köhler, J., and Welling, M. Spherical CNNs. In *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=Hkbd5xZRb.

Colmenarejo, L., Orellana, R., Saliola, F., Schilling, A., and Zabrocki, M. An Insertion Algorithm on Multiset Partitions with Applications to Diagram Algebras. *Journal of Algebra*, 557:97–128, 2020. ISSN 0021-8693. URL https://doi.org/10.1016/j.jalgebra.2020.04.010.

Finkelshtein, B., Baskin, C., Maron, H., and Dym, N. A Simple and Universal Rotation Equivariant Point-Cloud Network. In Cloninger, A., Doster, T., Emerson, T., Kaul, M., Ktena, I., Kvinge, H., Miolane, N., Rieck, B., Tymochko, S., and Wolf, G. (eds.), *Proceedings of Topological, Algebraic, and Geometric Learning Workshops 2022*, volume 196 of *Proceedings of Machine Learning Research*, pp. 107–115. PMLR, 25 Feb–22 Jul 2022. URL https://proceedings.mlr.press/v196/finkelshtein22a.html.

Finzi, M., Welling, M., and Wilson, A. G. G. A Practical Method for Constructing Equivariant Multilayer Perceptrons for Arbitrary Matrix Groups. In Meila, M. and Zhang, T. (eds.), *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 3318–3328. PMLR, 18–24 Jul 2021. URL https://proceedings.mlr.press/v139/finzi21a.html.

Goodman, R. and Wallach, N. R. *Symmetry, Representations and Invariants*. Springer, 2009.

Greydanus, S., Dzamba, M., and Yosinski, J. Hamiltonian Neural Networks. In Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/file/26cd8ecadce0d4efd6cc8a8725cbd1f8-Paper.pdf.

Grood, C. Brauer Algebras and Centralizer Algebras for SO(2n,C). *Journal of Algebra*, 222(2):678–707,

1999. ISSN 0021-8693. URL https://doi.org/10.1006/jabr.1999.8069.

Halverson, T. and Jacobson, T. N. Set-Partition Tableaux and Representations of Diagram Algebras. *Algebraic Combinatorics*, 3(2):509–538, 2020. URL https://doi.org/10.5802/alco.102.

Halverson, T. and Ram, A. Gems from the Work of Georgia Benkart. *Notices of the American Mathematical Society*, 69(3):375–384, March 2022. URL https://doi.org/10.1090/noti2447.

Hanlon, P. and Wales, D. On the Decomposition of Brauer's Centralizer Algebras. *Journal of Algebra*, 121(2):409–445, 1989. ISSN 0021-8693. URL https://doi.org/10.1016/0021-8693(89)90076-8.

Hartford, J. S., Graham, D. R., Leyton-Brown, K., and Ravanbakhsh, S. Deep Models of Interactions Across Sets. In *Proceedings of the 35th International Conference on Machine Learning*, pp. 1914–1923. PMLR, 2018. URL http://proceedings.mlr.press/v80/hartford18a.html.

Kondor, R., Lin, Z., and Trivedi, S. Clebsch–Gordan Nets: a Fully Fourier Space Spherical Convolutional Neural Network. In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL https://proceedings.neurips.cc/paper/2018/file/a3fc981af450752046be179185ebc8b5-Paper.pdf.

Kumagai, W. and Sannai, A. Universal Approximation Theorem for Equivariant Maps by Group CNNs, 2020. arXiv:2012.13882.

Lehrer, G. I. and Zhang, R. B. The Brauer category and invariant theory. *Journal of the European Mathematical Society*, 17:2311–2351, 2012. URL https://ems.press/content/serial-article-files/7287.

Lim, L.-H. and Nelson, B. J. What is an equivariant neural network?, 2022. arXiv:2205.07362.

Maron, H., Ben-Hamu, H., Shamir, N., and Lipman, Y. Invariant and Equivariant Graph Networks. In *International Conference on Learning Representations*, 2019. URL https://openreview.net/forum?id=Syx72jC9tm.

Pan, H. and Kondor, R. Permutation Equivariant Layers for Higher Order Interactions. In Camps-Valls, G., Ruiz, F. J. R., and Valera, I. (eds.), *Proceedings of The 25th International Conference on Artificial Intelligence andl Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pp. 5987–6001. PMLR, 28–30 Mar 2022. URL https://proceedings.mlr.press/v151/pan22a.html.

Pearce-Crump, E. Connecting Permutation Equivariant Neural Networks and Partition Diagrams. arXiv:2212.08648, 2022.

Ryan, K. *Representation-theoretic approaches to several problems in probability*. PhD thesis, Queen Mary University of London, 2021. URL https://qmro.qmul.ac.uk/xmlui/handle/123456789/77236.

Segal, E. Group Representation Theory. Course Notes for Imperial College London, 2014. URL https://www.homepages.ucl.ac.uk/~ucaheps/papers/Group%20Representation%20theory%202014.pdf.

Villar, S., Hogg, D. W., Storey-Fisher, K., Yao, W., and Blum-Smith, B. Scalars are universal: Equivariant machine learning, structured like classical physics. In Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, 2021. URL https://openreview.net/forum?id=ba27-RzNaIv.

Weiler, M., Geiger, M., Welling, M., Boomsma, W., and Cohen, T. 3D Steerable CNNs: Learning Rotationally Equivariant Features in Volumetric Data. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, pp. 10402–10413. Curran Associates Inc., 2018. URL https://dl.acm.org/doi/pdf/10.5555/3327546.3327700.

Wenzl, H. On the Structure of Brauer's Centralizer Algebras. *Annals of Mathematics*, 128(1):173–193, 1988. URL http://www.jstor.org/stable/1971466.

Weyl, H. *The Classical Groups: Their Invariants and Representations*. Princeton University Press, 1946.

Zaheer, M., Kottur, S., Ravanbakhsh, S., Poczos, B., Salakhutdinov, R. R., and Smola, A. J. Deep Sets. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R. (eds.), *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL https://proceedings.neurips.cc/paper/2017/file/f22e4747da1aa27e363d86d40ff442fe-Paper.pdf.

A. Brauer's Invariant Argument, adapted for $\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$

A.1. Some Preliminary Material

We consider throughout the real vector space $\mathbb{R}^n$.

Let $GL(n)$ be the group of invertible linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$. Let G be any subgroup of $GL(n)$.

Recall that the vector space $\mathbb{R}^n$ has, associated with it, its dual vector space, $(\mathbb{R}^n)^*$. Let $B \coloneqq \{b_i \mid i \in [n]\}$ be any basis of $\mathbb{R}^n$. It has associated with it the dual basis $B^* \coloneqq \{b_i^* \mid i \in [n]\}$, a basis of $(\mathbb{R}^n)^*$, such that $b_i^*(b_j) = \delta_{i,j}$.

In particular, coordinates on $\mathbb{R}^n$ with respect to the basis B are linear functions, that is, elements of $(\mathbb{R}^n)^*$. Indeed, if $v = \sum_j x_j b_j$, then the coordinate function x_j can be identified with the dual basis vector b_j^* since

$$b_j^*(v) = b_j^*(\sum_i x_i b_i) = \sum_i x_i b_j^*(b_i) = x_j \tag{44}$$

Since G is a subgroup of $GL(n)$, we see that $\mathbb{R}^n$ is a representation of G, which we denote by ρ_1 in the following. In fact, for all $f \in G$, we have that $\rho_1(f) = f$.

Consequently, if v is any vector in $\mathbb{R}^n$, and f is any element of G, then the linear transformation

$$\rho_1(f) = f : v \to f(v) \tag{45}$$

can be expressed in its matrix representation, choosing B as the basis for each copy of $\mathbb{R}^n$, as the matrix $a(f) = (a_{i,j})$ such that

$$y_i = \sum_j a_{i,j} x_j \tag{46}$$

where, in the basis B,

$$v = \sum_j x_j b_j \text{ and } f(v) = \sum_j y_j b_j \tag{47}$$

for some coefficients $x_j, y_j \in \mathbb{R}$.

We will sometimes express (46) in the form $y = a(f)x$, where x, y are column vectors such that the i^{th} component of each vector is x_i and y_i, respectively.

As $(\mathbb{R}^n, \rho_1)$ is a representation of G, we can define another representation of G, called the contragredient representation, on the dual space $(\mathbb{R}^n)^*$, as follows

$$(\rho_1^{-1})^\top : G \to GL((\mathbb{R}^n)^*) \tag{48}$$

where, for all $f \in G$,

$$(\rho_1^{-1})^\top(f) : (\mathbb{R}^n)^* \to (\mathbb{R}^n)^* \tag{49}$$

is defined by

$$(\rho_1^{-1})^\top(f)[u] : v \mapsto u(\rho_1^{-1}(f)(v)) \tag{50}$$

One way of understanding the contragredient representation $((\mathbb{R}^n)^*, (\rho_1^{-1})^\top)$ of G is as the action on $(\mathbb{R}^n)^*$ such that all pairings between $(\mathbb{R}^n)^*$ and $\mathbb{R}^n$ remain invariant under their respective actions. Indeed, if $u \in (\mathbb{R}^n)^*$, and $v \in \mathbb{R}^n$, then we see that, for all $f \in G$

$$v \mapsto \rho_1(f)(v) \tag{51}$$

and

$$u \mapsto (\rho_1^{-1})^\top(f)[u] \tag{52}$$

and so

$$u(v) \mapsto (\rho_1^{-1})^\top(f)[u](\rho_1(f)(v)) = u(\rho_1^{-1}(f)(\rho_1(f)(v))) = u(v) \tag{53}$$

Hence, expressing $u \in (\mathbb{R}^n)^*$ in the dual basis B^* as

$$u = \sum_j p_j b_j^* \tag{54}$$

and $(\rho_1^{-1})^\top(f)[u] \in (\mathbb{R}^n)^*$ as

$$(\rho_1^{-1})^\top(f)[u] = \sum_j q_j b_j^* \tag{55}$$

we see that, for any $f \in G$, the linear transformation

$$(\rho_1^{-1})^\top(f) : u \to (\rho_1^{-1})^\top(f)[u] \tag{56}$$

can be expressed in its matrix representation as

$$pa(f)^{-1} = q \tag{57}$$

as a result of (53), and so

$$p = qa(f) \tag{58}$$

that is,

$$p_i = \sum_j q_j a_{j,i} \tag{59}$$

In (58), p, q are row vectors such that the i^{th} component of each is p_i and q_i, respectively.

We also have that $(\mathbb{R}^n)^{\otimes k}$ is a representation of G, which we denote by ρ_k. In particular, for all $f \in G$, we see that $\rho_k(f) = \rho_1(f)^{\otimes k} = f^{\otimes k}$.

Consequently, if v is any vector in $(\mathbb{R}^n)^{\otimes k}$, and f is any element of G, then the linear transformation

$$\rho_k(f) = f^{\otimes k} : v \to f^{\otimes k}(v) \tag{60}$$

can be expressed in its matrix representation, choosing B as the basis for each copy of $\mathbb{R}^n$, as the matrix $A(f) = (A_{I,J})$, over all tuples $I := (i_1, i_2, \dots, i_k), J := (j_1, j_2, \dots, j_k) \in [n]^k$, such that

$$y_I = \sum_J A_{I,J} x_J \tag{61}$$

where

$$A_{I,J} = \prod_{r=1}^{k} a_{i_r, j_r} \tag{62}$$

and

$$v = \sum_J x_J b_J \text{ and } f(v) = \sum_J y_J b_J \tag{63}$$

for some coefficients $x_J, y_J \in \mathbb{R}$ in the basis $\{b_J \mid J \in [n]^k\}$ of $(\mathbb{R}^n)^{\otimes k}$, where

$$b_J := b_{j_1} \otimes b_{j_2} \otimes \dots \otimes b_{j_k} \tag{64}$$

As before, since $((\mathbb{R}^n)^{\otimes k}, \rho_k)$ is a representation of G, we obtain the contragredient representation, $(((\mathbb{R}^n)^*)^{\otimes k}, (\rho_k^{-1})^\top)$

$$(\rho_k^{-1})^\top : G \to GL(((\mathbb{R}^n)^*)^{\otimes k}) \tag{65}$$

where

$$(\rho_k^{-1})^\top(f) : ((\mathbb{R}^n)^*)^{\otimes k} \to ((\mathbb{R}^n)^*)^{\otimes k} \tag{66}$$

is defined by

$$(\rho_k^{-1})^\top(f)[u] : v \mapsto u(\rho_k^{-1}(f)(v)) \tag{67}$$

In particular, we see that $(\rho_k^{-1})^\top = ((\rho_1^{-1})^\top)^{\otimes k}$.

Hence, expressing $u \in ((\mathbb{R}^n)^*)^{\otimes k}$ in the dual basis $\{b_J^* \mid J \in [n]^k\}$ of $(\mathbb{R}^n)^{\otimes k}$, where

$$b_J^* := b_{j_1}^* \otimes b_{j_2}^* \otimes \dots \otimes b_{j_k}^* \tag{68}$$

as

$$u = \sum_J p_J b_J^* \tag{69}$$

and $(\rho_k^{-1})^\top(f)[u] \in ((\mathbb{R}^n)^*)^{\otimes k}$ as

$$(\rho_k^{-1})^\top(f)[u] = \sum_J q_J b_J^* \tag{70}$$

we see that, for any $f \in G$, the linear transformation

$$(\rho_k^{-1})^\top(f) : u \to (\rho_k^{-1})^\top(f)[u] \tag{71}$$

can be expressed in its matrix representation form as

$$p_I = \sum_J q_J A_{J,I} \tag{72}$$

where

$$A_{J,I} = \prod_{r=1}^{k} a_{j_r, i_r} \tag{73}$$

A.2. Brauer's Invariant Argument

We adapt the argument given in (Brauer, 1937) to construct an invariant of G that is in bijective correspondence with an element of $\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$.

A linear map $\phi : (\mathbb{R}^n)^{\otimes k} \to (\mathbb{R}^n)^{\otimes l}$ is an element of $\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$ if and only if

$$\phi \circ \rho_k(f) = \rho_l(f) \circ \phi \tag{74}$$

for all $f \in G$.

Choosing the basis B as the basis for each copy of $\mathbb{R}^n$, the matrix representation of (74) is

$$C(\phi)K(f) = L(f)C(\phi) \tag{75}$$

where $K(f) = (K_{I,J})$ and $L(f) = (L_{I,J})$ are as in (60) and $C(\phi) = (C_{I,J})$.

Hence, (75) gives

$$\sum_{J \in [n]^k} C_{I,J} K_{J,M} = \sum_{J \in [n]^l} L_{I,J} C_{J,M} \tag{76}$$

where $I \in [n]^l$ and $M \in [n]^k$.

Brauer's trick is as follows.

Let $v(1), v(2), \dots, v(k)$ be elements of $\mathbb{R}^n$, and suppose that they are all mapped by the same transformation $\rho_1(f)$, for some $f \in G$. Then, by (47), in the basis B of $\mathbb{R}^n$, we see that

$$v(r) = \sum_j x_j(r) b_j \tag{77}$$

for all $r \in [k]$, and so, by (46), we have that

$$y_i(r) = \sum_j a_{i,j} x_j(r) \tag{78}$$

for all $r \in [k]$.

Then, considering the tensor product $v(1) \otimes v(2) \otimes \cdots \otimes v(k)$, an element of $(\mathbb{R}^n)^{\otimes k}$, and considering its transformation under $\rho_k(f)$, for the same $f \in G$, we see that the coefficient of the basis vector b_J for $v(1) \otimes v(2) \otimes \cdots \otimes v(k)$, as in (63), is

$$\prod_{r=1}^{k} x_{j_r}(r) \tag{79}$$

and the coefficient of the basis vector b_I for $\rho_k(f)[v(1) \otimes v(2) \otimes \cdots \otimes v(k)]$ is

$$\prod_{r=1}^{k} y_{i_r}(r) \tag{80}$$

and that (62) holds, namely that

$$K_{I,J} = \prod_{r=1}^{k} a_{i_r,j_r} \tag{81}$$

Hence, by (61), we have that

$$\prod_{r=1}^{k} y_{i_r}(r) = \sum_{J \in [n]^k} K_{I,J} \prod_{r=1}^{k} x_{j_r}(r) \tag{82}$$

Also, let $u(1), u(2), \dots, u(l)$ be elements of $(\mathbb{R}^n)^*$, the dual space of $\mathbb{R}^n$. Then, by (54), in the dual basis B^*, we have that

$$u(t) = \sum_j p_j(t) b_j^* \tag{83}$$

for all $t \in [l]$, and so, by (59), we see that

$$p_i(t) = \sum_j q_j(t) a_{j,i} \tag{84}$$

for all $t \in [l]$.

Then, considering the tensor product $u(1) \otimes u(2) \otimes \cdots \otimes u(l)$, an element of $((\mathbb{R}^n)^*)^{\otimes l}$, and considering its transformation under $(\rho_l^{-1})^\top(f)$, for the same $f \in G$, we see that the coefficient of the basis vector b_I^* for $u(1) \otimes u(2) \otimes \cdots \otimes u(l)$, as in (69), is

$$\prod_{t=1}^{l} p_{i_t}(t) \tag{85}$$

and the coefficient of the basis vector b_J^* for $(\rho_l^{-1})^\top(f)[u(1) \otimes u(2) \otimes \cdots \otimes u(l)]$ is

$$\prod_{t=1}^{l} q_{j_t}(t) \tag{86}$$

and that (73) holds, namely that

$$L_{J,I} = \prod_{t=1}^{l} a_{j_t,i_t} \tag{87}$$

Hence, by (72), we have that

$$\prod_{t=1}^{l} p_{i_t}(t) = \sum_{J \in [n]^l} \prod_{t=1}^{l} q_{j_t}(t) L_{J,I} \tag{88}$$

Multiplying both sides of (76) by

$$\prod_{t=1}^{l} q_{i_t}(t) \prod_{r=1}^{k} x_{m_r}(r) \tag{89}$$

adding over all tuples $I \in [n]^l, M \in [n]^k$, and applying (82) and (88) gives us, on the LHS

$$\sum_{I \in [n]^l, J \in [n]^k} C_{I,J} \prod_{t=1}^{l} q_{i_t}(t) \prod_{r=1}^{k} y_{j_r}(r) \tag{90}$$

and on the RHS

$$\sum_{J \in [n]^l, M \in [n]^k} C_{J,M} \prod_{t=1}^{l} p_{j_t}(t) \prod_{r=1}^{k} x_{m_r}(r) \tag{91}$$

This means that

$$\sum_{I \in [n]^l, J \in [n]^k} C_{I,J} \prod_{t=1}^{l} p_{i_t}(t) \prod_{r=1}^{k} x_{j_r}(r) \tag{92}$$

is an invariant for the group G, that is, it is a linear transformation

$$((\mathbb{R}^n)^*)^{\otimes l} \otimes (\mathbb{R}^n)^{\otimes k} \to \mathbb{R} \tag{93}$$

which maps an element of the form

$$u(1) \otimes u(2) \otimes \cdots \otimes u(l) \otimes v(1) \otimes v(2) \otimes \cdots \otimes v(k) \tag{94}$$

to (92) that is invariant under the action of G.

We see that each stage of the argument, from (74) to (94), is an if and only if statement, since any invariant of G which is linear in any subset $\{v(1), v(2), \dots, v(k)\}$ of k vectors of $\mathbb{R}^n$ and any subset $\{u(1), u(2), \dots, u(l)\}$ of l vectors in $(\mathbb{R}^n)^*$, and is a homogeneous function of their union, must be of the form (92) since any invariant of these $l + k$ elements must be a linear combination of the elements $\prod_{t=1}^{l} p_{i_t}(t) \prod_{r=1}^{k} x_{j_r}(r)$ where $x_{j_r}(r)$ is the j_r^{th} coefficient of $v(r)$ when expressed in some basis B of $\mathbb{R}^n$, and $p_{i_t}(t)$ is the i_t^{th} coefficient of $u(t)$ when expressed in its dual basis B^*. Hence, starting from (92) and running the argument in reverse gives (74), and therefore shows that it is an if and only if statement.

In particular, this means that each element of $\text{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$, having chosen the basis B for each copy of $\mathbb{R}^n$, is in bijective correspondence with an invariant

$$((\mathbb{R}^n)^*)^{\otimes l} \otimes (\mathbb{R}^n)^{\otimes k} \to \mathbb{R} \tag{95}$$

of G of the form (92), as desired.

B. First Fundamental Theorems for $O(n)$, $Sp(n)$ and $SO(n)$

We state, without proof, the First Fundamental Theorems for $O(n)$, $Sp(n)$ and $SO(n)$. See (Goodman & Wallach, 2009) for more details.

Theorem B.1 (First Fundamental Theorem for $O(n)$)**.** *Let $n \in \mathbb{Z}_{\geq 1}$, and suppose that the real vector space $\mathbb{R}^n$ has associated with it a non-degenerate, symmetric, bilinear form $(\cdot,\cdot)$, as in Section 4.*

Let us choose the standard basis for $\mathbb{R}^n$, so that $(\cdot,\cdot)$ becomes the Euclidean inner product on $\mathbb{R}^n$, as defined in (8).

If $f : (\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ is a polynomial function on elements in $(\mathbb{R}^n)^{\otimes(l+k)}$ of the form

$$u(1) \otimes u(2) \otimes \cdots \otimes u(l) \otimes v(1) \otimes v(2) \otimes \cdots \otimes v(k) \tag{96}$$

that is $O(n)$-invariant, then it must be a polynomial of the Euclidean inner products

$$(u(i), u(j)), (u(i), v(j)), (v(i), v(j)) \tag{97}$$

Theorem B.2 (First Fundamental Theorem for $Sp(n)$)**.** *Let $n \in \mathbb{Z}_{\geq 2}$ be even, and suppose that the real vector space $\mathbb{R}^n$ has associated with it a non-degenerate, skew-symmetric, bilinear form $\langle \cdot,\cdot \rangle$, as in Section 4.*

Let us choose the symplectic basis for $\mathbb{R}^n$, so that $\langle \cdot,\cdot \rangle$ becomes the form given in (12).

Note that, in this basis, the non-degenerate, symmetric, bilinear form $(\cdot,\cdot)$ which we can also associate with $\mathbb{R}^n$, becomes the Euclidean inner product on $\mathbb{R}^n$, as defined in (8), since the symplectic basis is standard with respect to $(\cdot,\cdot)$. If $f : (\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ is a polynomial function on elements in $(\mathbb{R}^n)^{\otimes(l+k)}$ of the form

$$u(1) \otimes u(2) \otimes \cdots \otimes u(l) \otimes v(1) \otimes v(2) \otimes \cdots \otimes v(k) \tag{98}$$

that is $Sp(n)$-invariant, then it must be a polynomial of the Euclidean inner products

$$(u(i), v(j)) \tag{99}$$

together with the skew products

$$\langle u(i), u(j) \rangle, \langle v(i), v(j) \rangle \tag{100}$$

such that $i < j$ in (100).

Theorem B.3 (First Fundamental Theorem for $SO(n)$)**.** *Let $n \in \mathbb{Z}_{\geq 1}$, and suppose that the real vector space $\mathbb{R}^n$ has associated with it a non-degenerate, symmetric, bilinear form $(\cdot,\cdot)$, as in Section 4.*

Let us choose the standard basis for $\mathbb{R}^n$, so that $(\cdot,\cdot)$ becomes the Euclidean inner product on $\mathbb{R}^n$, as defined in (8).

If $f : (\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ is a polynomial function on elements in $(\mathbb{R}^n)^{\otimes(l+k)}$ of the form

$$u(1) \otimes u(2) \otimes \cdots \otimes u(l) \otimes v(1) \otimes v(2) \otimes \cdots \otimes v(k) \tag{101}$$

that is $SO(n)$-invariant, then it must be a polynomial of the Euclidean inner products

$$(u(i), u(j)), (u(i), v(j)), (v(i), v(j)) \tag{102}$$

together with the $n \times n$ subdeterminants of the $n \times (l+k)$ matrix M, which is the matrix having as its columns:

$$M := \begin{pmatrix} | & | & & | & | & | & & | \\ u(1) & u(2) & \dots & u(l) & v(1) & v(2) & \dots & v(k) \\ | & | & & | & | & | & & | \end{pmatrix} \tag{103}$$

C. Bijective Correspondence between the Brauer and Brauer–Grood vector spaces and the Invariants for $O(n)$, $Sp(n)$ and $SO(n)$

We now provide a proof of the results given in Theorems 6.6, 6.7, and 6.8.

It can be shown that $\mathbb{R}^n$, as a representation of each of the groups $G = O(n)$, $Sp(n)$ and $SO(n)$ (for $Sp(n)$, $n = 2m$), is isomorphic to its dual space $(\mathbb{R}^n)^*$, by using the appropriate bilinear form that is used to define each of the groups in question. See (Goodman & Wallach, 2009) for more details.

Hence, for each group G, we can apply its First Fundamental Theorem to the invariant given in (92), now considered as a function $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$, noting that each term of the polynomial (92) contains each of the vectors $u(1), u(2), \dots, u(l), v(1), v(2), \dots, v(k)$ exactly once.

Consequently, we obtain a spanning set of invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ for each of $O(n)$, $Sp(n)$ and $SO(n)$ (for $Sp(n)$, $n = 2m$).

Theorem C.1 (Spanning Set of Invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ for $O(n)$)**.** *The functions*

$$(z(1), z(2))(z(3), z(4)) \dots (z(l+k-1), z(l+k)) \tag{104}$$

where $z(1), \dots, z(l+k)$ is a permutation of $u(1), u(2), \dots, u(l), v(1), v(2), \dots, v(k)$ form a spanning set of invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ for $O(n)$.

Clearly, functions of the form (104) can only be formed when $l+k$ is even, hence there are no invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ for $O(n)$ when $l+k$ is odd.

Theorem C.2 (Spanning Set of Invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ for $Sp(n)$, $n = 2m$)**.** *The functions*

$$[z(1), z(2)][z(3), z(4)] \dots [z(l+k-1), z(l+k)] \tag{105}$$

where $z(1), \dots, z(l+k)$ is a permutation of $u(1), u(2), \dots, u(l), v(1), v(2), \dots, v(k)$ and

$$[z(i), z(i+1)] := \begin{cases} (z(i), z(i+1)) & \text{if } z(i) = u(j) \text{ and } z(i+1) = v(m) \text{, or } z(i) = v(m) \text{ and } z(i+1) = u(j) \text{, for some } j \in [l],\ m \in [k] \\ \langle z(i), z(i+1) \rangle & \text{otherwise.} \end{cases} \tag{106}$$

form a spanning set of invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ for $Sp(n)$, with $n = 2m$.

Clearly, functions of the form (105) can only be formed when $l+k$ is even, hence there are no invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ for $Sp(n)$ when $l+k$ is odd.

Theorem C.3 (Spanning Set of Invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ for $SO(n)$)**.** *Functions of the form (104) together with functions of the form*

$$\det(z(1), \ldots, z(n))(z(n+1), z(n+2)) \ldots (z(l+k-1), z(l+k)) \tag{107}$$

where $z(1), \ldots, z(l+k)$ is a permutation of $u(1), u(2), \ldots, u(l), v(1), v(2), \ldots, v(k)$ form a spanning set of invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ for $SO(n)$.

Clearly, functions of the form (107) can only be formed when $n \leq l+k$, and either when n is odd and $l+k$ is odd, or when n is even and $l+k$ is even.

We can now construct a bijective correspondence between each of the functions (104), (105), and (107), and either a (k,l)–Brauer diagram or an $(l+k)\backslash n$–diagram, as follows.

Indeed, consider the spanning set of invariants of the form (104). Then we can associate a (k,l)–Brauer diagram with each element of the set by labelling the top l vertices by $u(1), u(2), \ldots, u(l)$ and the bottom k vertices by $v(1), v(2), \ldots, v(k)$, and drawing an edge between two vertices if and only if they appear in the same inner product in (104).

We do a similar thing for the spanning set of invariants of the form (105), associating a (k,l)–Brauer diagram with each element of the set, labelling the vertices in the same mamner, and drawing an edge between two vertices if and only if they appear in the same inner or skew product in (105).

Finally, consider the functions of the form (107). Then we can associate an $(l+k)\backslash n$–diagram with each element of the set by labelling the top l vertices by $u(1), u(2), \ldots, u(l)$ and the bottom k vertices by $v(1), v(2), \ldots, v(k)$, leaving the vertices $z(1), \ldots, z(n)$ free and drawing an edge between the other vertices if and only if they appear in the same inner product in (107).

As a result, we have constructed a bijective correspondence between a spanning set of invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ for $O(n)$ with the set of (k,l)–Brauer diagrams whose span is the Brauer vector space $B_k^l(n)$, for $SO(n)$ with the set of (k,l)–Brauer diagrams together with the set of $(l+k)\backslash n$–diagrams whose span is the Brauer–Grood vector space $D_k^l(n)$, and for $Sp(n)$ ($n = 2m$) with the set of (k,l)–Brauer diagrams whose span is the Brauer vector space $B_k^l(n)$.

Consequently, for each group G, the bijective correspondence (92) that exists between the spannning set of invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ (given in Theorems C.1, C.2, and C.3) and a spanning set of matrices for $\mathrm{Hom}_G((\mathbb{R}^n)^{\otimes k}, (\mathbb{R}^n)^{\otimes l})$ in the standard/symplectic basis of $\mathbb{R}^n$, and the bijective correspondence that exists between the spannning set of invariants $(\mathbb{R}^n)^{\otimes(l+k)} \to \mathbb{R}$ and the set of diagrams that span either the Brauer vector space $B_k^l(n)$, for $O(n)$ and $Sp(n)$, or the Brauer–Grood vector space $D_k^l(n)$, for $SO(n)$, together prove the results given in Theorems 6.6, 6.7, and 6.8.

D. Examples

In the following, in order to save space, we use some temporary notation to denote a sum of a number of weight parameters. We represent a sum of weight parameters, where the sum is over some index set A, by a single element that is indexed by the set of indices itself, that is

$$\lambda_A \coloneqq \sum_{i \in A} \lambda_i \tag{108}$$

For example, $\lambda_{8,11,12}$ represents the sum $\lambda_8 + \lambda_{11} + \lambda_{12}$.

D.1. $O(n)$

D.1.1. A Basis of $\mathrm{End}_{O(2)}((\mathbb{R}^2)^{\otimes 2})$

We consider the surjective map

$$\Phi_{2,2}^2 : B_2^2(2) \to \mathrm{End}_{O(2)}((\mathbb{R}^2)^{\otimes 2}) \tag{109}$$

and apply Theorem 6.6, noting that $l+k$ is even. Also, as $2n \geq l+k$, this map is an isomorphism of vector spaces, hence the images of the basis diagrams of $B_2^2(2)$ forms a basis of $\mathrm{End}_{O(2)}((\mathbb{R}^2)^{\otimes 2})$.

Figure 2 shows how to find the basis of $\mathrm{End}_{O(2)}((\mathbb{R}^2)^{\otimes 2})$ from the basis of $B_2^2(2)$.

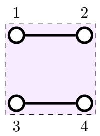
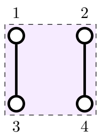
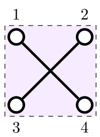
Basis Diagram d_β	Matrix Entries	Basis Element of $\text{End}_{O(2)}((\mathbb{R}^2)^{\otimes 2})$
	$(\delta_{i_1,i_2}\delta_{j_1,j_2})$	$\begin{array}{c} \\ 1,1 \\ 1,2 \\ 2,1 \\ 2,2 \end{array}\begin{array}{cccc} 1,1 & 1,2 & 2,1 & 2,2 \\ \hline 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{array}$
	$(\delta_{i_1,j_1}\delta_{i_2,j_2})$	$\begin{array}{c} \\ 1,1 \\ 1,2 \\ 2,1 \\ 2,2 \end{array}\begin{array}{cccc} 1,1 & 1,2 & 2,1 & 2,2 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}$
	$(\delta_{i_1,j_2}\delta_{i_2,j_1})$	$\begin{array}{c} \\ 1,1 \\ 1,2 \\ 2,1 \\ 2,2 \end{array}\begin{array}{cccc} 1,1 & 1,2 & 2,1 & 2,2 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}$

Figure 2. The images under $\Phi^2_{2,2}$ of the basis diagrams of $B^2_2(2)$ make up a basis of $\text{End}_{O(2)}((\mathbb{R}^2)^{\otimes 2})$.

Hence, any element of $\text{End}_{O(2)}((\mathbb{R}^2)^{\otimes 2})$, in the basis of matrix units of $\text{End}((\mathbb{R}^2)^{\otimes 2})$, is of the form

$$\begin{array}{c} \\ 1,1 \\ 1,2 \\ 2,1 \\ 2,2 \end{array}\begin{array}{cc|cc} 1,1 & 1,2 & 2,1 & 2,2 \\ \lambda_{1,2,3} & 0 & 0 & \lambda_1 \\ 0 & \lambda_2 & \lambda_3 & 0 \\ \hline 0 & \lambda_3 & \lambda_2 & 0 \\ \lambda_1 & 0 & 0 & \lambda_{1,2,3} \end{array} \tag{110}$$

for scalars $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$.

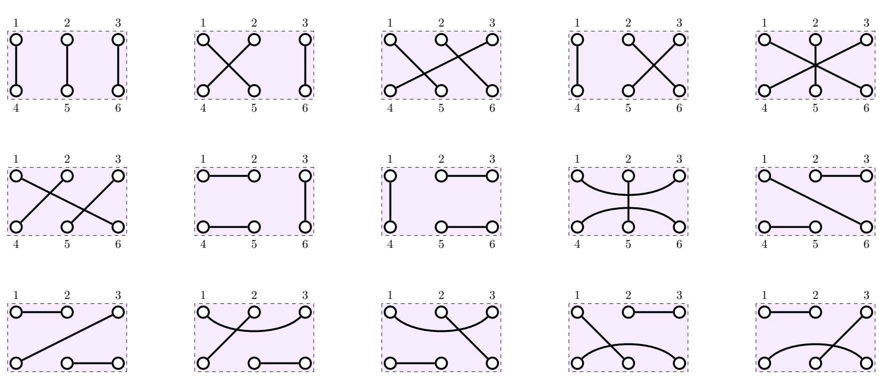
D.1.2. A Basis of $\text{End}_{O(3)}((\mathbb{R}^3)^{\otimes 3})$

We consider the surjective map

$$\Phi^3_{3,3} : B^3_3(3) \to \text{End}_{O(3)}((\mathbb{R}^3)^{\otimes 3}) \tag{111}$$

and apply Theorem 6.6, noting that $l + k$ is even. Also, as $2n \geq l + k$, this map is an isomorphism of vector spaces, hence the images of the basis diagrams of $B^3_3(3)$ forms a basis of $\text{End}_{O(3)}((\mathbb{R}^3)^{\otimes 3})$.

The basis diagrams of $B^3_3(3)$ are



Taking their images under $\Phi_{3,3}^3$, we see that any element of $\text{End}_{O(3)}((\mathbb{R}^3)^{\otimes 3})$, in the basis of matrix units of $\text{End}((\mathbb{R}^3)^{\otimes 3})$, is of the form

$$
\begin{array}{c|ccc|ccc|ccc|ccc|ccc|ccc|ccc|ccc|ccc|}
 & 1,1,1 & 1,1,2 & 1,1,3 & 1,2,1 & 1,2,2 & 1,2,3 & 1,3,1 & 1,3,2 & 1,3,3 & 2,1,1 & 2,1,2 & 2,1,3 & 2,2,1 & 2,2,2 & 2,2,3 & 2,3,1 & 2,3,2 & 2,3,3 & 3,1,1 & 3,1,2 & 3,1,3 & 3,2,1 & 3,2,2 & 3,2,3 & 3,3,1 & 3,3,2 & 3,3,3 \\
1,1,1 & \lambda_{1,\ldots,15} & 0 & 0 & 0 & \lambda_{8,11,12} & 0 & 0 & 0 & \lambda_{8,11,12} & 0 & \lambda_{9,14,15} & 0 & \lambda_{7,10,13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{9,14,15} & 0 & 0 & 0 & \lambda_{7,10,13} & 0 & 0 \\
1,1,2 & 0 & \lambda_{1,2,7} & 0 & \lambda_{4,6,15} & 0 & 0 & 0 & 0 & 0 & \lambda_{3,5,11} & 0 & 0 & 0 & \lambda_{7,11,15} & 0 & 0 & 0 & \lambda_{11} & 0 & 0 & 0 & 0 & 0 & \lambda_{15} & 0 & \lambda_7 & 0 \\
1,1,3 & 0 & 0 & \lambda_{1,2,7} & 0 & 0 & 0 & \lambda_{4,6,15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_7 & 0 & \lambda_{15} & 0 & \lambda_{3,5,11} & 0 & 0 & 0 & \lambda_{11} & 0 & 0 & 0 & \lambda_{7,11,15} \\
\hline
1,2,1 & 0 & \lambda_{3,4,13} & 0 & \lambda_{1,5,9} & 0 & 0 & 0 & 0 & 0 & \lambda_{2,6,12} & 0 & 0 & 0 & \lambda_{9,12,13} & 0 & 0 & 0 & \lambda_{12} & 0 & 0 & 0 & 0 & 0 & \lambda_9 & 0 & \lambda_{13} & 0 \\
1,2,2 & \lambda_{8,10,14} & 0 & 0 & 0 & \lambda_{1,4,8} & 0 & 0 & 0 & \lambda_8 & 0 & \lambda_{2,3,14} & 0 & \lambda_{5,6,10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{14} & 0 & 0 & 0 & \lambda_{10} & 0 & 0 \\
1,2,3 & 0 & 0 & 0 & 0 & 0 & \lambda_1 & 0 & \lambda_4 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & \lambda_6 & 0 & 0 & 0 & \lambda_3 & 0 & \lambda_5 & 0 & 0 & 0 & 0 & 0 \\
\hline
1,3,1 & 0 & 0 & \lambda_{3,4,13} & 0 & 0 & 0 & \lambda_{1,5,9} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{13} & 0 & \lambda_9 & 0 & \lambda_{2,6,12} & 0 & 0 & 0 & \lambda_{12} & 0 & 0 & 0 & \lambda_{9,12,13} \\
1,3,2 & 0 & 0 & 0 & 0 & 0 & \lambda_4 & 0 & \lambda_1 & 0 & 0 & 0 & \lambda_3 & 0 & 0 & 0 & \lambda_5 & 0 & 0 & 0 & \lambda_2 & 0 & \lambda_6 & 0 & 0 & 0 & 0 & 0 \\
1,3,3 & \lambda_{8,10,14} & 0 & 0 & 0 & \lambda_8 & 0 & 0 & 0 & \lambda_{1,4,8} & 0 & \lambda_{14} & 0 & \lambda_{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{2,3,14} & 0 & 0 & 0 & \lambda_{5,6,10} & 0 & 0 \\
\hline
2,1,1 & 0 & \lambda_{5,6,10} & 0 & \lambda_{2,3,14} & 0 & 0 & 0 & 0 & 0 & \lambda_{1,4,8} & 0 & 0 & 0 & \lambda_{8,10,14} & 0 & 0 & 0 & \lambda_8 & 0 & 0 & 0 & 0 & 0 & \lambda_{14} & 0 & \lambda_{10} & 0 \\
2,1,2 & \lambda_{9,12,13} & 0 & 0 & 0 & \lambda_{2,6,12} & 0 & 0 & 0 & \lambda_{12} & 0 & \lambda_{1,5,9} & 0 & \lambda_{3,4,13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_9 & 0 & 0 & 0 & \lambda_{13} & 0 & 0 \\
2,1,3 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & 0 & \lambda_6 & 0 & 0 & 0 & \lambda_1 & 0 & 0 & 0 & \lambda_4 & 0 & 0 & 0 & \lambda_5 & 0 & \lambda_3 & 0 & 0 & 0 & 0 & 0 \\
\hline
2,2,1 & \lambda_{7,11,15} & 0 & 0 & 0 & \lambda_{3,5,11} & 0 & 0 & 0 & \lambda_{11} & 0 & \lambda_{4,6,15} & 0 & \lambda_{1,2,7} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{15} & 0 & 0 & 0 & \lambda_7 & 0 & 0 \\
2,2,2 & 0 & \lambda_{7,10,13} & 0 & \lambda_{9,14,15} & 0 & 0 & 0 & 0 & 0 & \lambda_{8,11,12} & 0 & 0 & 0 & \lambda_{1,\ldots,15} & 0 & 0 & 0 & \lambda_{8,11,12} & 0 & 0 & 0 & 0 & 0 & \lambda_{9,14,15} & 0 & \lambda_{7,10,13} & 0 \\
2,2,3 & 0 & 0 & \lambda_7 & 0 & 0 & 0 & \lambda_{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{1,2,7} & 0 & \lambda_{4,6,15} & 0 & \lambda_{11} & 0 & 0 & 0 & \lambda_{3,5,11} & 0 & 0 & 0 & \lambda_{7,11,15} \\
\hline
2,3,1 & 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 & \lambda_5 & 0 & 0 & 0 & \lambda_4 & 0 & 0 & 0 & \lambda_1 & 0 & 0 & 0 & \lambda_6 & 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 \\
2,3,2 & 0 & 0 & \lambda_{13} & 0 & 0 & 0 & \lambda_9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{3,4,13} & 0 & \lambda_{1,5,9} & 0 & \lambda_{12} & 0 & 0 & 0 & \lambda_{2,6,12} & 0 & 0 & 0 & \lambda_{9,12,13} \\
2,3,3 & 0 & \lambda_{10} & 0 & \lambda_{14} & 0 & 0 & 0 & 0 & 0 & \lambda_8 & 0 & 0 & 0 & \lambda_{8,10,14} & 0 & 0 & 0 & \lambda_{1,4,8} & 0 & 0 & 0 & 0 & 0 & \lambda_{2,3,14} & 0 & \lambda_{5,6,10} & 0 \\
\hline
3,1,1 & 0 & 0 & \lambda_{5,6,10} & 0 & 0 & 0 & \lambda_{2,3,14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{10} & 0 & \lambda_{14} & 0 & \lambda_{1,4,8} & 0 & 0 & 0 & \lambda_8 & 0 & 0 & 0 & \lambda_{8,10,14} \\
3,1,2 & 0 & 0 & 0 & 0 & 0 & \lambda_6 & 0 & \lambda_2 & 0 & 0 & 0 & \lambda_5 & 0 & 0 & 0 & \lambda_3 & 0 & 0 & 0 & \lambda_1 & 0 & \lambda_4 & 0 & 0 & 0 & 0 & 0 \\
3,1,3 & \lambda_{9,12,13} & 0 & 0 & 0 & \lambda_{12} & 0 & 0 & 0 & \lambda_{2,6,12} & 0 & \lambda_9 & 0 & \lambda_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{1,5,9} & 0 & 0 & 0 & \lambda_{3,4,13} & 0 & 0 \\
\hline
3,2,1 & 0 & 0 & 0 & 0 & 0 & \lambda_5 & 0 & \lambda_3 & 0 & 0 & 0 & \lambda_6 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & \lambda_4 & 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 \\
3,2,2 & 0 & 0 & \lambda_{10} & 0 & 0 & 0 & \lambda_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{5,6,10} & 0 & \lambda_{2,3,14} & 0 & \lambda_8 & 0 & 0 & 0 & \lambda_{1,4,8} & 0 & 0 & 0 & \lambda_{8,10,14} \\
3,2,3 & 0 & \lambda_{13} & 0 & \lambda_9 & 0 & 0 & 0 & 0 & 0 & \lambda_{12} & 0 & 0 & 0 & \lambda_{9,12,13} & 0 & 0 & 0 & \lambda_{2,6,12} & 0 & 0 & 0 & 0 & 0 & \lambda_{1,5,9} & 0 & \lambda_{3,4,13} & 0 \\
\hline
3,3,1 & \lambda_{7,11,15} & 0 & 0 & 0 & \lambda_{11} & 0 & 0 & 0 & \lambda_{3,5,11} & 0 & \lambda_{15} & 0 & \lambda_7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{4,6,15} & 0 & 0 & 0 & \lambda_{1,2,7} & 0 & 0 \\
3,3,2 & 0 & \lambda_7 & 0 & \lambda_{15} & 0 & 0 & 0 & 0 & 0 & \lambda_{11} & 0 & 0 & 0 & \lambda_{7,11,15} & 0 & 0 & 0 & \lambda_{3,5,11} & 0 & 0 & 0 & 0 & 0 & \lambda_{4,6,15} & 0 & \lambda_{1,2,7} & 0 \\
3,3,3 & 0 & 0 & \lambda_{7,10,13} & 0 & 0 & 0 & \lambda_{9,14,15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{7,10,13} & 0 & \lambda_{9,14,15} & 0 & \lambda_{8,11,12} & 0 & 0 & 0 & \lambda_{8,11,12} & 0 & 0 & 0 & \lambda_{1,\ldots,15}
\end{array}
\tag{112}
$$

for scalars $\lambda_1, \ldots, \lambda_{15} \in \mathbb{R}$.

Notice that $\text{End}_{O(3)}((\mathbb{R}^3)^{\otimes 3})$ is a 15-dimensional vector space living inside a 729-dimensional vector space, $\text{End}((\mathbb{R}^3)^{\otimes 3})$.

D.2. $Sp(n)$

We saw that this is similar to $O(n)$, except we replace δ by ϵ if there is an edge between two vertices that are in the same row.

D.2.1. A Spanning Set for $\text{Hom}_{Sp(2)}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2)$

For the surjective map

$$X_{3,2}^1 : B_3^1(2) \rightarrow \text{Hom}_{Sp(2)}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2) \tag{113}$$

we apply Theorem 6.7, noting that $l + k = 4$, which is even.

Figure 3 shows how to find a spanning set for $\text{Hom}_{Sp(2)}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2)$.

This means that any element of $\text{Hom}_{Sp(2)}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2)$, in the basis of matrix units of $\text{Hom}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2)$, is of the form

$$
\begin{array}{c|cc|cc|cc|cc|}
 & 1,1,1 & 1,1,1' & 1,1',1 & 1,1',1' & 1',1,1 & 1',1,1' & 1',1',1 & 1',1',1' \\
1 & 0 & \lambda_1 + \lambda_2 & -\lambda_1 + \lambda_3 & 0 & -\lambda_2 - \lambda_3 & 0 & 0 & 0 \\
1' & 0 & 0 & 0 & \lambda_2 + \lambda_3 & 0 & \lambda_1 - \lambda_3 & -\lambda_1 - \lambda_2 & 0
\end{array}
\tag{114}
$$

for scalars $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$.

D.3. $SO(n)$

D.3.1. A Spanning Set for $\text{End}_{SO(3)}((\mathbb{R}^3)^{\otimes 3})$

We apply Theorem 6.8.

As $n \leq l + k$, and n is odd and $l + k$ is even, we see that $\text{End}_{SO(3)}((\mathbb{R}^3)^{\otimes 3}) = \text{End}_{O(3)}((\mathbb{R}^3)^{\otimes 3})$ and so any element of $\text{End}_{SO(3)}((\mathbb{R}^3)^{\otimes 3})$, in the basis of matrix units of $\text{End}((\mathbb{R}^3)^{\otimes 3})$, is of the form (112).

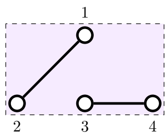
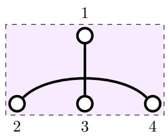
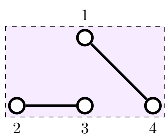
Basis Diagram d_β	Matrix Entries	Spanning Set Element of $\mathrm{Hom}_{Sp(2)}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2)$
	$(\delta_{i_1,j_1}\epsilon_{j_2,j_3})$	$\begin{array}{c} \\ 1 \\ 1' \end{array} \begin{array}{cc|cc|cc|cc} 1,1,1 & 1,1,1' & 1,1',1 & 1,1',1' & 1',1,1 & 1',1,1' & 1',1',1 & 1',1',1' \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{array}$
	$(\delta_{i_1,j_2}\epsilon_{j_1,j_3})$	$\begin{array}{c} \\ 1 \\ 1' \end{array} \begin{array}{cc|cc|cc|cc} 1,1,1 & 1,1,1' & 1,1',1 & 1,1',1' & 1',1,1 & 1',1,1' & 1',1',1 & 1',1',1' \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \end{array}$
	$(\delta_{i_1,j_3}\epsilon_{j_1,j_2})$	$\begin{array}{c} \\ 1 \\ 1' \end{array} \begin{array}{cc|cc|cc|cc} 1,1,1 & 1,1,1' & 1,1',1 & 1,1',1' & 1',1,1 & 1',1,1' & 1',1',1 & 1',1',1' \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \end{array}$

Figure 3. The images under $X^1_{3,2}$ of the basis diagrams of $B^1_3(2)$ make up a spanning set for $\mathrm{Hom}_{Sp(2)}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2)$.

D.3.2. A Spanning Set for $\mathrm{Hom}_{SO(2)}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2)$

Again, we apply Theorem 6.8.

As $n \leq l + k$, and n is even and $l + k$ is even, we see that three $(3, 1)$–Brauer diagrams and six $(1 + 3)\backslash 2$–diagrams make up a basis of $D^1_3(2)$. Their images under

$$\Psi^1_{3,2} : D^1_3(2) \rightarrow \mathrm{Hom}_{SO(2)}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2) \tag{115}$$

forms a spanning set of $\mathrm{Hom}_{SO(2)}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2)$.

Calculating the images of the three $(3, 1)$–Brauer diagrams is the same as for the $O(n)$ case. Figure 4 shows how to find the images of the six $(1 + 3)\backslash 2$–diagrams.

D.4. Local Symmetry Example

As an example of the result given in Section 8, suppose that we want to find a spanning set for

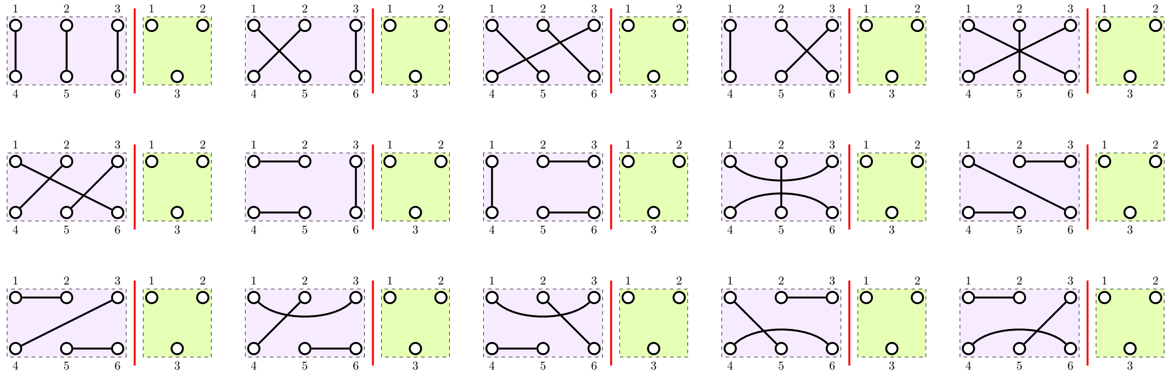
$$\mathrm{Hom}_{SO(3)\times SO(3)}((\mathbb{R}^3)^{\otimes 3} \boxtimes \mathbb{R}^3, (\mathbb{R}^3)^{\otimes 3} \boxtimes (\mathbb{R}^3)^{\otimes 2}) \tag{116}$$

By (43), we know that (116) is isomorphic to

$$\mathrm{End}_{SO(3)}((\mathbb{R}^3)^{\otimes 3}) \otimes \mathrm{Hom}_{SO(3)}(\mathbb{R}^3, (\mathbb{R}^3)^{\otimes 2}) \tag{117}$$

Hence, to find a spanning set, all we need to do is find the images of the basis elements of $D^3_3(3) \otimes D^2_1(3)$ under $\Psi^3_{3,3} \otimes \Psi^2_{1,3}$ and take the Kronecker product of the resulting matrices.

To save space, we will only show what the basis elements of $D^3_3(3) \otimes D^2_1(3)$ are. By Theorem 6.8, the basis elements are



where we have used a red demarcation line to separate the vertices of the respective diagrams. Note that no edge can cross this red line.

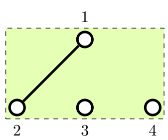
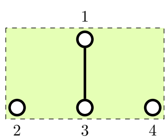
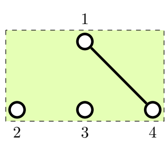
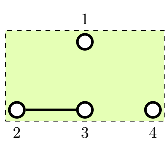
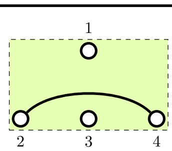
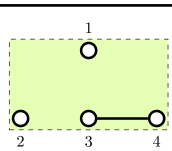
Basis Diagram d_α	Matrix Entries	Spanning Set Element of $\mathrm{Hom}_{SO(2)}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2)$
	$\left(\chi\begin{pmatrix}1 & 2\\ j_2 & j_3\end{pmatrix}\delta_{i_1,j_1}\right)$	$\begin{array}{c\|cc\|cc\|cc\|cc} & 1,1,1 & 1,1,2 & 1,2,1 & 1,2,2 & 2,1,1 & 2,1,2 & 2,2,1 & 2,2,2 \\ 1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \end{array}$
	$\left(\chi\begin{pmatrix}1 & 2\\ j_1 & j_3\end{pmatrix}\delta_{i_1,j_2}\right)$	$\begin{array}{c\|cc\|cc\|cc\|cc} & 1,1,1 & 1,1,2 & 1,2,1 & 1,2,2 & 2,1,1 & 2,1,2 & 2,2,1 & 2,2,2 \\ 1 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \end{array}$
	$\left(\chi\begin{pmatrix}1 & 2\\ j_1 & j_2\end{pmatrix}\delta_{i_1,j_3}\right)$	$\begin{array}{c\|cc\|cc\|cc\|cc} & 1,1,1 & 1,1,2 & 1,2,1 & 1,2,2 & 2,1,1 & 2,1,2 & 2,2,1 & 2,2,2 \\ 1 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \end{array}$
	$\left(\chi\begin{pmatrix}1 & 2\\ i_1 & j_3\end{pmatrix}\delta_{j_1,j_2}\right)$	$\begin{array}{c\|cc\|cc\|cc\|cc} & 1,1,1 & 1,1,2 & 1,2,1 & 1,2,2 & 2,1,1 & 2,1,2 & 2,2,1 & 2,2,2 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 2 & -1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{array}$
	$\left(\chi\begin{pmatrix}1 & 2\\ i_1 & j_2\end{pmatrix}\delta_{j_1,j_3}\right)$	$\begin{array}{c\|cc\|cc\|cc\|cc} & 1,1,1 & 1,1,2 & 1,2,1 & 1,2,2 & 2,1,1 & 2,1,2 & 2,2,1 & 2,2,2 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 2 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{array}$
	$\left(\chi\begin{pmatrix}1 & 2\\ i_1 & j_1\end{pmatrix}\delta_{j_2,j_3}\right)$	$\begin{array}{c\|cc\|cc\|cc\|cc} & 1,1,1 & 1,1,2 & 1,2,1 & 1,2,2 & 2,1,1 & 2,1,2 & 2,2,1 & 2,2,2 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 2 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{array}$

Figure 4. The images under $\Psi_{3,2}^1$ of the six $(1+3)\backslash 2$–diagrams in $D_3^1(2)$, together with the images under $\Psi_{3,2}^1$ of the three $(3,1)$–Brauer diagrams in $D_3^1(2)$ (not shown), make up a spanning set for $\mathrm{Hom}_{SO(2)}((\mathbb{R}^2)^{\otimes 3}, \mathbb{R}^2)$.