EQUIVARIANT NEURAL NETWORKS AND PIECEWISE LINEAR REPRESENTATION THEORY – SYMMETRIC GROUPS

JOEL GIBSON, DANIEL TUBBENHAUER AND GEORDIE WILLIAMSON

Contents

We assume familiarity with the main paper's notation [**GTW23b**].

1. Example: symmetric groups

We now consider the ***symmetric group*** S_n on $\{1,...,n\}$. Let $m_{ij}=2$ for $|i-j|\neq 1$, $m_{ij}=3$ otherwise. The ***Coxeter presentation*** of the symmetric group, that we will use, is $S_n\cong\langle s_1,...,s_{n-1}|(s_is_j)^{m_{ij}}=1\rangle$ with s_i corresponding to the simple transposition that swaps i and $i+1$. The ***regular*** S_n***-representation*** is the $n!$ dimensional $\mathbb{R}$-vector space $R=\mathbb{R}\{x_\sigma|\sigma\in S_n\}$ with left and right actions

$$s_i \centerdot x_\sigma = x_{s_i\sigma},\quad x_\sigma \centerdot s_i = x_{\sigma s_i}. \tag{1.1}$$

We use the ordered basis of R given by $\{x_{\sigma_1}<...<x_{\sigma_{n!}}\}$ where $x_\sigma\leq x_\tau$ if and only if σ is lexicographically smaller than τ in one-line notation. We use the same order on S_n itself.

Example 1.2. We have $S_3=\{1<s_2<s_1<s_1s_2<s_2s_1<s_2s_1s_2=s_1s_2s_1\}$. Let us write $\mathtt{M}_i=\mathtt{M}_{s_i}$ for the left action matrix of s_i, and similar for the right action matrix. Then the left and right actions of (1.1) in matrices are:

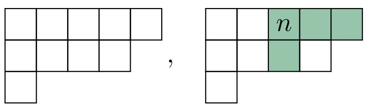
$$\mathtt{M}_1=\begin{pmatrix}0&0&1&0&0&0\\0&0&0&0&1&0\\1&0&0&0&0&0\\0&0&0&0&0&1\\0&1&0&0&0&0\\0&0&0&1&0&0\end{pmatrix},\quad \mathtt{M}_2=\begin{pmatrix}0&1&0&0&0&0\\1&0&0&0&0&0\\0&0&0&1&0&0\\0&0&1&0&0&0\\0&0&0&0&0&1\\0&0&0&0&1&0\end{pmatrix},\quad \mathtt{M}_1^r=\begin{pmatrix}0&0&1&0&0&0\\0&0&0&1&0&0\\1&0&0&0&0&0\\0&1&0&0&0&0\\0&0&0&0&0&1\\0&0&0&0&1&0\end{pmatrix},\quad \mathtt{M}_2^r=\begin{pmatrix}0&1&0&0&0&0\\1&0&0&0&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\\0&0&1&0&0&0\\0&0&0&1&0&0\end{pmatrix}.$$

Note that $S_3\cong D_{2\cdot 3}$, and in fact, the above S_3-representation is equivalent to the $D_{2\cdot 3}$-representation from [**GTW23b**, ???], but with different generators and on a differently ordered basis. ◇

Of course, (1.1) also makes sense for general elements and not just for the s_i.

1A. **Simple representations, projections and inclusions.** Let $P(n)$ denote the set of ***partitions of*** n. We illustrate elements of $P(n)$ using Young diagrams in English convention. For $\lambda\in P(n)$ let $N(\lambda)$ denote the set of nodes of λ and let $h_\lambda(n)$ denote the hook length of $n\in N(\lambda)$, e.g.

$\lambda=(5,4,1)\leftrightsquigarrow$



$\Rightarrow h_\lambda(n)=4.$

For $\lambda\in P(n)$ we let $\dim\lambda=n!/\prod_{n\in N(\lambda)}h_\lambda(n)$ denote the value of the ***hook length formula*** on λ.

Recall from e.g. [**FH91**, Section I.4] that there is a rational S_n-representation, the ***Specht module***, associated to $\lambda\in P(n)$ of dimension $\dim\lambda$. We denote by L_λ its scalar extension to $\mathbb{R}$.

Remark 1A.1. [**FH91**, Section I.4] use $\mathbb{C}$ as the ground field, but they point out that everything works over $\mathbb{Q}$ as well. In fact, Specht modules can be defined over $\mathbb{Z}$ but that will not play any role for us.

Lemma 1A.2. *As real* S_n*-representations we have*

$$R\cong\bigoplus_{\lambda\in P(n)}L_\lambda^{\oplus\dim\lambda}.$$

All of the appearing real S_n*-representations are simple and pairwise nonisomorphic. All simple real* S_n*-representations appear in this way.*

Proof. As for the dihedral group there is no significant difference between simple real and complex S_n-representations, so this lemma is classical. See for example [**FH91**, Section I.4] for a concise discussion. ☐

For $\lambda \in P(n)$ let $ST(\lambda)$ denote the set of all ***standard tableaux*** of shape λ, that is, fillings of the nodes of λ with the entries from $\{1, ..., n\}$ such that the entries strictly increase when reading along rows and columns. For example,

$$\lambda = \begin{array}{|c|c|}\hline & \\ \hline & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \Rightarrow ST(\lambda) = \left\{ \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \right\}.$$

Notation 1A.3. Recall that $\#ST(\lambda) = \dim \lambda$. And we will index the simple real S_n-representations in Lemma 1A.2 by $T \in ST(\lambda)$ and write L_T for them. Their projectors, constructed below, are denoted p_T.

The central projectors are the so-called ***Young symmetrizers*** and can be found in many textbooks on S_n-representations, see for example [**FH91**, Section I.4]. In general, ***Young's seminormal basis*** gives the Artin–Wedderburn decomposition. We will briefly recall what we need from this theory.

For $T \in ST(\lambda)$ let $r(T) \in \mathbb{R}[S_n]$ denote the formal sum of all permutations which stabilize the rows of T, and dually, let $c(T) \in \mathbb{R}[S_n]$ be the signed formal sum of permutations that stabilize the columns of T. As one final piece of notation let $T(n-1)$ denote the standard tableaux with $n-1$ nodes obtained from T by removing the node with entry n.

Lemma 1A.4. *Let $\lambda \in P(n)$ and $T \in ST(\lambda)$, and set $\kappa_\lambda = \prod_{n \in N(\lambda)} h_\lambda(n) = n!/\dim \lambda$. We have the following inductive formulas. First, $p(\emptyset) = 1$, and otherwise*

$$p_T = \frac{1}{\kappa_\lambda} p_{T(n-1)}\, r(T)c(T)\, p_{T(n-1)} \in \mathbb{R}[S_n].$$

Proof. Since the theories over $\mathbb{R}$ and $\mathbb{C}$ coincide, explicit forms for these projectors are well-known, see for example [**JK81**, Section 3.2]. □

The projectors to L_T, obtained by replacing $\sigma \in S_n$ by its right action matrix as in (1.1), are denoted by the same symbol. We will write $\vec{p}_T \in \mathbb{R}^{n!}$ for the coefficient vector that one gets when one writes p_T in the basis of $\mathbb{R}[S_n]$ fixed above.

Notation 1A.5. We want to give matrices so we fix an order on standard tableaux as follows: Let $\lambda, \lambda' \in P(n)$ be the partitions underlying T and T', respectively. Then $T \leq T'$ if and only if (λ is smaller or equal to λ' lexicographically, and if $\lambda = \lambda'$ then additionally $col(T) \leq col(T')$ lexicographically) where $col()$ is the sequence of the entries in column reading.

Example 1A.6. For $n = 3$ we have

$$T(1) = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline\end{array}, \quad \begin{matrix} T(2) = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \\ T(3) = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}, \end{matrix} \quad T(4) = \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline\end{array},$$

$$\vec{p}_{T(1)} = 1/6(1,1,1,1,1,1), \quad \begin{matrix} \vec{p}_{T(2)} = 1/6(2,-1,2,-1,-1,-1), \\ \vec{p}_{T(3)} = 1/6(2,1,-2,-1,-1,1), \end{matrix} \quad \vec{p}_{T(4)} = 1/6(1,-1,-1,1,1,-1).$$

In terms of matrices we get

$$p_{T(1)} = \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix},$$

$$p_{T(2)} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} \\ \frac{1}{3} & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} & \frac{1}{3} \\ -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} & \frac{1}{3} \end{pmatrix},$$

$$p_{T(3)} = \begin{pmatrix} \frac{1}{3} & \frac{1}{6} & -\frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{3} & -\frac{1}{6} \\ -\frac{1}{3} & -\frac{1}{6} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & -\frac{1}{3} & \frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{3} & \frac{1}{6} & \frac{1}{3} \end{pmatrix},$$

$$p_{T(4)} = \begin{pmatrix} \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \end{pmatrix},$$

as the projectors. ◇

As for the dihedral group, the relevant change-of-basis matrix $B \in \mathrm{End}_{\mathbb{R}}(R)$ is defined by choosing $\dim L_\lambda = \dim \lambda$ linearly independent rows of the projectors, collect them into a $n!$-by-$n!$ matrix $\mathtt{Q}^{-1}$. With the same notation as for dihedral groups we get:

Lemma 1A.7. *We have the following.*

(a) *The matrix $\mathtt{Q}^{-1}$ is invertible, with inverse denoted by $\mathtt{Q}$.*

(b) *The matrix $\mathtt{Q}$ is S_n-equivariant and satisfies $\mathtt{N}_\sigma = \mathtt{Q}^{-1}\mathtt{M}_\sigma\mathtt{Q}$.*

Proof. Using Lemma 1A.4, the arguments from [**GTW23b**, ???] can be copied. □

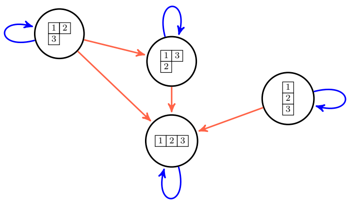
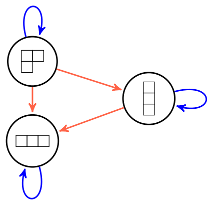
Example 1A.8. Again back to $n = 3$. We get

$$\mathtt{Q} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 0 & -1 & 0 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 0 & 1 & 0 & -1 & 1 \\ 1 & 0 & 1 & 0 & 1 & -1 \end{pmatrix}, \quad \mathtt{Q}^{-1} = \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{3} & -\frac{1}{6} & \frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & -\frac{1}{3} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & -\frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & -\frac{1}{6} \end{pmatrix},$$

where we have chosen the first and the last row in the projectors to two dimensional summands. $\diamond$

1B. **ReLU and symmetric groups.** We do not know a general statement for the symmetric group, and we can only discuss $n \in \{3, 4, 5\}$. We will also see a difference between the interaction graphs and the isotypic interaction graphs.

Example 1B.1. Let $T(i)$ be as in Example 1A.6. Writing $T(i)$ instead of $L_{T(i)}$ and similarly for the isotypic components, for S_3 we get

$\Gamma_{\mathsf{ReLU}} =$  , $i\Gamma_{\mathsf{ReLU}} =$  .

Note the asymmetry between the isomorphic S_3-representations $L_{T(2)}$ and $L_{T(3)}$.

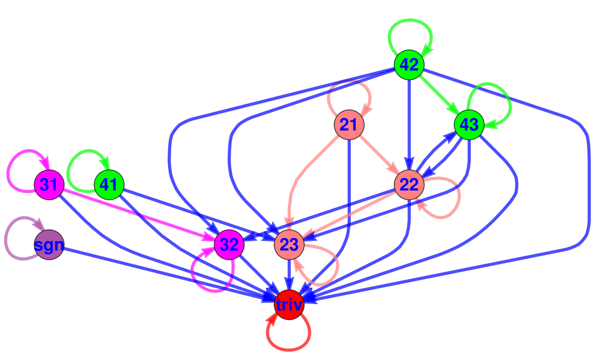
Moreover, the isotypic graph is not obtained from Γ_{ReLU} by collapsing vertices along isomorphism classes, but the isotypic graph is the same as for the dihedral group under the isomorphism $S_3 \cong D_{2\cdot 3}$. $\diamond$

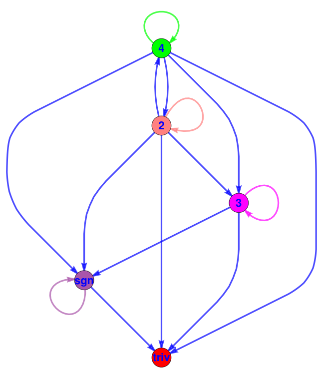
Remark 1B.2. Example 1B.1 shows that the interaction graphs crucially depend on the choice of projectors: The projectors to isotypic components are canonical and indeed the isotypic intersection graphs for the dihedral group and the symmetric group are the same. The projectors to the simple summands are not canonical and the different choices for the dihedral group and the symmetric group changed the intersection graphs.

Example 1B.3. For S_4 the pattern of the graph Γ_{ReLU} is already quite sophisticated. Let us therefore also display $i\Gamma_{\mathsf{ReLU}}$. Before we display the graphs, let us fix the notation:

$$1 \leftrightsquigarrow \ytableaushort{1234}, \quad 21 \leftrightsquigarrow \ytableaushort{134,2}, \quad 22 \leftrightsquigarrow \ytableaushort{124,3}, \quad 23 \leftrightsquigarrow \ytableaushort{123,4}, \quad 31 \leftrightsquigarrow \ytableaushort{13,24}, \quad 32 \leftrightsquigarrow \ytableaushort{12,34},$$
$$41 \leftrightsquigarrow \ytableaushort{14,2,3}, \quad 42 \leftrightsquigarrow \ytableaushort{13,2,4}, \quad 43 \leftrightsquigarrow \ytableaushort{12,3,4}, \quad 5 \leftrightsquigarrow \ytableaushort{1,2,3,4}.$$

Only keeping the first entry gives the isotypic components. In this notation one gets (the colors are a visual aid only and indicate isotypic components):

$\Gamma_{\mathsf{ReLU}} =$ 

, $i\Gamma_{\mathsf{ReLU}} =$ 

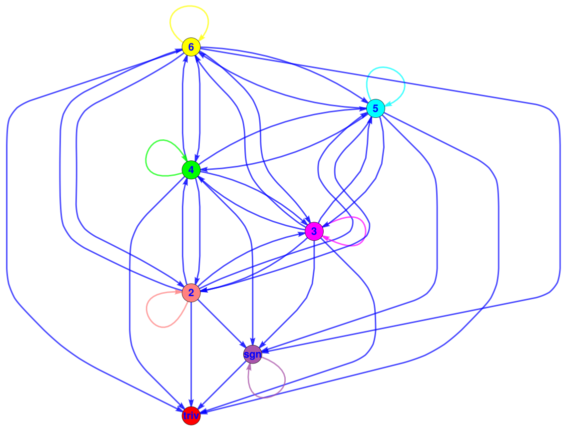
.

These graphs can be produced with the Mathematica code available on [GTW23a]. $\diamond$

Example 1B.4. Let us fix the following notation for the seven partitions of five:

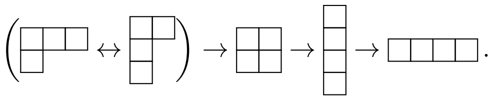
$$1 \leftrightsquigarrow \ydiagram{5}, \quad 2 \leftrightsquigarrow \ydiagram{4,1}, \quad 3 \leftrightsquigarrow \ydiagram{3,2}, \quad 4 \leftrightsquigarrow \ydiagram{3,1,1}, \quad 5 \leftrightsquigarrow \ydiagram{2,2,1}, \quad 6 \leftrightsquigarrow \ydiagram{2,1,1,1}, \quad 7 \leftrightsquigarrow \ydiagram{1,1,1,1,1}.$$

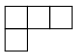
The standard tableaux are ordered such that column reading is minimized, see Example 1B.3 for the $n = 4$ case. For S_5 the graph, using the same color code as in Example 1B.3, $i\Gamma_{\mathsf{ReLU}}$ is

$i\Gamma_{\mathsf{ReLU}} =$  .

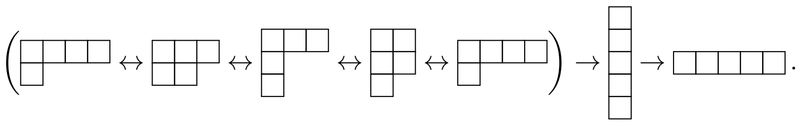
The graph Γ_{ReLU} itself is large, so we do not display it here. It can be produced using the Mathematica code available on [**GTW23a**]. ◇

Remark 1B.5. The order on partitions that comes from $i\Gamma_{\mathsf{ReLU}}$ in Example 1B.3 is



Thus, the standard S_4-representation corresponding to  and its conjugate are the easiest ones with respect to [**GTW23b**, ???].

For S_5 we get



from Example 1B.4.

Theorem 1B.6. *We have the following regarding Γ_{ReLU} and $i\Gamma_{\mathsf{ReLU}}$.*

(a) *All L have an edge to the trivial S_n-representation L_1.*

(b) *The interaction graphs Γ_{ReLU} for S_3 and S_4, and $i\Gamma_{\mathsf{ReLU}}$ for S_5 are as in Example 1B.1, Example 1B.3 and Example 1B.4. The remaining ones can be displayed using the code on available on [**GTW23a**].*

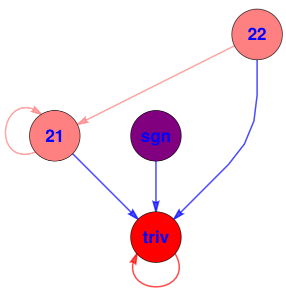
The proof of this theorem is postpone to the end of Section 1D.

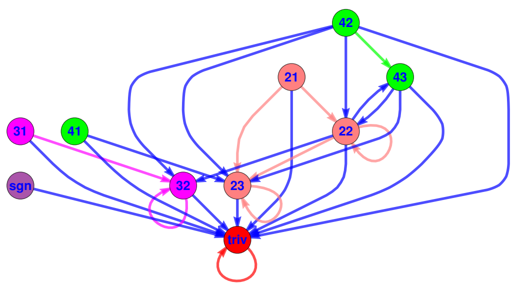
1C. **Absolute value and symmetric groups.** As for ReLU, we only have restricted knowledge of the pattern:

Theorem 1C.1. *We have the following regarding Γ_{Abs} and $i\Gamma_{\mathsf{Abs}}$.*

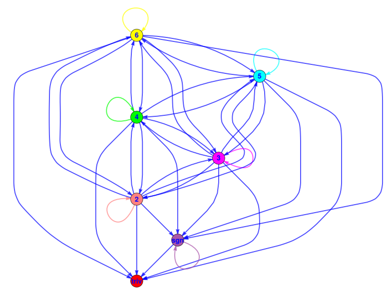
(a) *All L have an edge to the trivial S_n-representation L_1.*

(b) *The interaction graphs Γ_{Abs} for S_3 and S_4, and $i\Gamma_{\mathsf{Abs}}$ for S_5 are*

$n = 3$: 

, $n = 4$:  ,

$n = 5$:



*The remaining ones can be displayed using the code on available on [**GTW23a**].*

As for ReLU, the proof will follow below.

1D. **The piecewise linear maps for the symmetric group.** Let us first discuss $n = 4$ and the piecewise linear maps to the trivial S_4-representation. With the notation in Example 1B.3, we will explicitly give

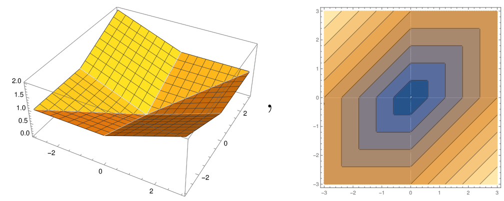
$$\mathsf{ReLU}^1_1, \quad \mathsf{ReLU}^1_{21}, \quad \mathsf{ReLU}^1_{31}, \quad \mathsf{ReLU}^1_4,$$

and the remaining maps are similar.

First, ReLU^1_1 is just ReLU and ReLU^1_5 is Abs, as before.

The piecewise linear map ReLU^1_{31} can be illustrated via level sets and we get

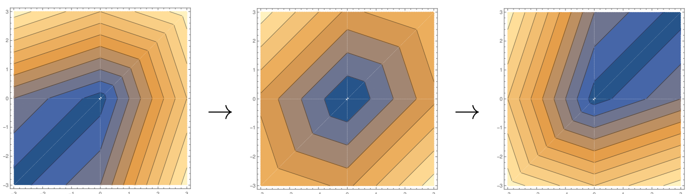
ReLU^1_{31} :



Surprisingly, this map is the same as ReLU^1_{21} for S_3.

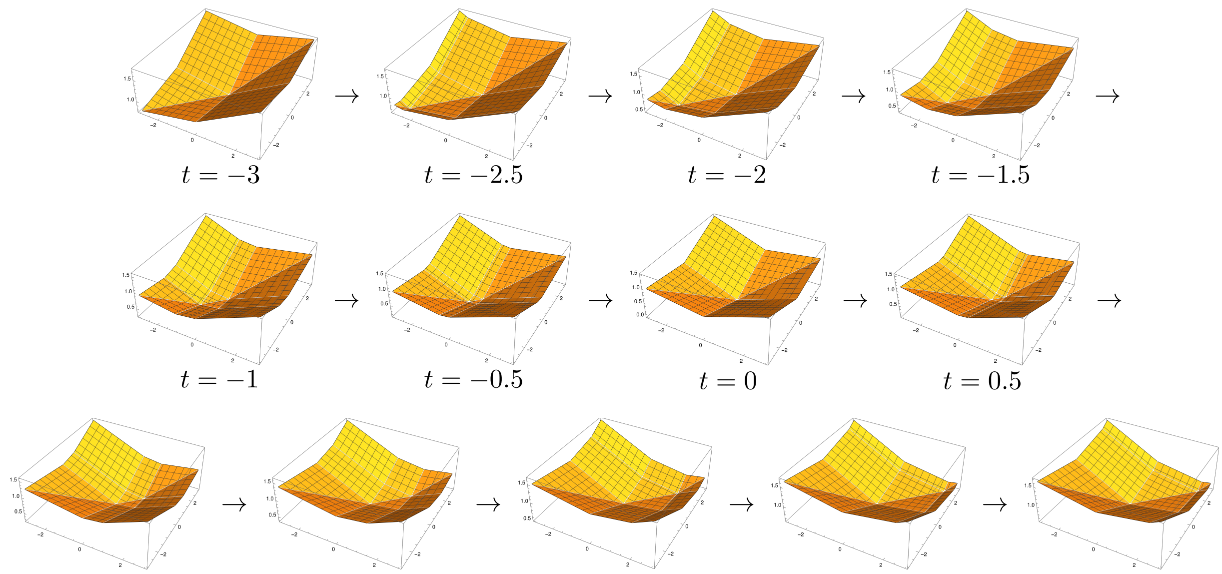
Finally, ReLU^1_{21} is a map $\mathbb{R}^3 \to \mathbb{R}$, which we illustrate as a movie thinking of the final entry as time. That is, we think of ReLU^1_{21} as $f(x, y, t)$, display $f(x, y, t)$ for fixed t and then vary t. What one gets is a map

ReLU^1_{21} :



$t = -3 \quad \to \quad t = 0 \quad \to \quad t = 3$,

with level sets for $t = -3$, $t = 0$ and $t = 3$. The movie reads as



$t = -3 \to t = -2.5 \to t = -2 \to t = -1.5 \to$

$t = -1 \to t = -0.5 \to t = 0 \to t = 0.5 \to$

$t = 1 \to t = 1.5 \to t = 2 \to t = 2.5 \to t = 3$.

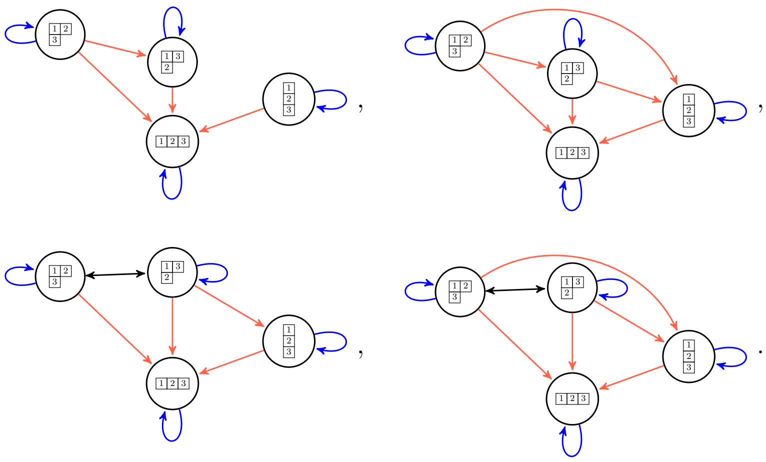
An honest movie of this map is available at [**GTW23a**].

Proof of Theorem 1B.6 and Theorem 1C.1. (a) follows from [**GTW23b**, ???].

Part (b) is a computer calculation, see the discussion above for $n = 3$ and $n = 4$, and partially for $n = 5$. The code is available at [**GTW23a**]. □

1E. **Minimal interaction graphs for the symmetric group with six elements.** Recall that interaction graphs in general depend on the choice of idempotents, cf. [**GTW23b**, ???]. In particular, the interaction graph for S_n depends on this choice whenever $n \geq 3$.

Example 1E.1. Here are a few example of how Γ_{ReLU} could be:



The northwest graph in Theorem 1E.2 is the one for the Young idempotents and the southeast graph is the one for the dihedral idempotents under the identification $S_3 \cong D_{2\cdot 3}$. ◇

For $n = 3$ we can classify the interaction graphs:

Theorem 1E.2. *The interaction graph for the Young idempotents is minimal. The case of Γ_{Abs} is verbatim.*

Proof. This boils down to a minimization problem after observing that the only choice involved in this case is the choice of a 4-4 idempotent matrix. The code verifying this is available at [**GTW23a**]. □

1F. **The polynomial representation of the symmetric groups.** Assume $n \in \mathbb{Z}_{\geq 3}$. There is another S_n-representation which is much smaller than R: the ***polynomial representation*** $\mathsf{Fun}(n, \mathbb{R}) = \mathsf{Fun}(\{1, ..., n\}, \mathbb{R})$. For it we have a complete answer regarding ReLU and Abs as we will see in this section.

Lemma 1F.1. *Let (n) and $(n-1, 1)$ denote the respective partitions of n.*

(a) *As real S_n-representations we have*

$$\mathsf{Fun}(n, \mathbb{R}) \cong L_{(n)} \oplus L_{(n-1,1)}.$$

(b) *The S_n-equivariant change-of-basis matrix is given by $\mathtt{Q} = \big((1, ..., 1), e_1 - e_2, ..., e_1 - e_n\big)$.*

(c) *$\mathtt{Q}$ is invertible with inverse given by*

$$\mathtt{Q}^{-1} = \frac{1}{n} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & -(n-1) & 1 & 1 & 1 \\ 1 & 1 & \ddots & \ddots & 1 \\ 1 & 1 & \ddots & \ddots & 1 \\ 1 & 1 & 1 & 1 & -(n-1) \end{pmatrix}.$$

Proof. Easy and omitted. □

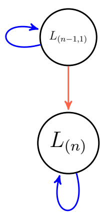
Theorem 1F.2. *Consider the interaction graph Γ_{ReLU}. Every vertex has a loop. Moreover, there is a non-loop edge from $L_{(n-1,1)}$ to $L_{(n)}$.*

The interaction graph Γ_{Abs} is exactly the same, and similarly for the isotypic interaction graphs.

Proof. This can be directly read-off from $\mathtt{Q}$ and its inverse $\mathtt{Q}^{-1}$ as in Lemma 1F.1. □

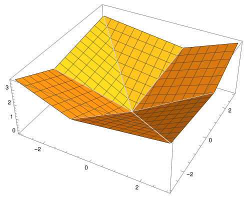
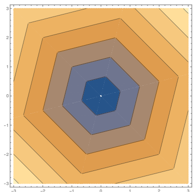
Example 1F.3. Independent of n we have

$$\Gamma_{\mathsf{ReLU}} = \Gamma_{\mathsf{Abs}} =$$

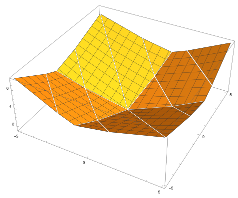
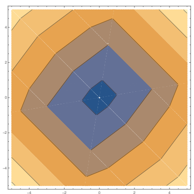


for both, ReLU and Abs. ◇

The explicit description of $\mathsf{ReLU}^{(n)}_{(n-1,1)}\colon L_{(n-1,1)} \to L_{(n)}$ is not difficult. Identifying $L_{(n-1,1)} \cong \mathbb{R}^{n-1}$ and $L_{(n)} \cong \mathbb{R}$ shows that $\mathsf{ReLU}^{(n)}_{(n-1,1)}$ can be seen as a map $\mathbb{R}^{n-1} \to \mathbb{R}$. Fixing all but two coordinates to be zero gives a map $\mathbb{R}^2 \to \mathbb{R}$ which can be illustrated by now familiar pictures:

 ,  .

This description is independent of n. Varying the coordinates set to zero to other values creates new hyperplanes, e.g.:

 ,  .

An animation of this can be found on [**GTW23a**].

References

[FH91] W. Fulton and J. Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991. A first course, Readings in Mathematics. doi:10.1007/978-1-4612-0979-9.

[GTW23a] J. Gibson, D. Tubbenhauer, and G. Williamson. Code related to the paper "Equivariant neural networks and piecewise linear representation theory". 2023. https://github.com/dtubbenhauer/pl-reps and https://www.dtubbenhauer.com/pl-reps/site/index.html.

[GTW23b] J. Gibson, D. Tubbenhauer, and G. Williamson. Equivariant neural networks and piecewise linear representation theory. 2023. URL: https://arxiv.org/abs/0706.0351.

[JK81] G. James and A. Kerber. *The representation theory of the symmetric group*, volume 16 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Co., Reading, Mass., 1981. With a foreword by P. M. Cohn, With an introduction by Gilbert de B. Robinson.

J.G.: The University of Sydney, School of Mathematics and Statistics F07, Office Carslaw 827, NSW 2006, Australia, www.jgibson.id.au
Email address: joel@jgibson.id.au

D.T.: The University of Sydney, School of Mathematics and Statistics F07, Office Carslaw 827, NSW 2006, Australia, www.dtubbenhauer.com, https://orcid.org/0000-0001-7265-5047
Email address: daniel.tubbenhauer@sydney.edu.au

G.W.: The University of Sydney, School of Mathematics and Statistics F07, Office L4.41 Quadrangle A14, NSW 2006, Australia, www.maths.usyd.edu.au/u/geordie/, https://orcid.org/0000-0003-3672-5284
Email address: g.williamson@sydney.edu.au