A NOT-SO-CHARACTERISTIC EQUATION: THE ART OF LINEAR ALGEBRA

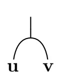
ELISHA PETERSON

Abstract. Can the cross product be generalized? Why are the trace and determinant so important in matrix theory? What do all of the coefficients of the characteristic polynomial represent? This paper describes a technique for 'doodling' equations from linear algebra which offers elegant solutions to all these questions. The doodles, known as trace diagrams, are graphs labeled by matrices that have a correspondence to multilinear functions. This correspondence permits computations in linear algebra to be performed using diagrams. The result is an elegant theory from which standard constructions of linear algebra such as the determinant, the trace, the adjugate matrix, Cramer's rule, and the characteristic polynomial arise naturally. Using the diagrams, it is easy to see how little structure gives rise to these various results, as they all can be traced back to the definition of the determinant and inner product.

1. Introduction

When I was an undergraduate, I remember asking myself: why does the cross product "only work" in three dimensions? And what's so special about the trace and determinant of a matrix? What is the *real* reason behind the connection between the cross product and the determinant? These questions have traditional answers, but I never found them satisfying. Perhaps the reason is that I could not, as a visual learner, "see" what these things really meant.

A few years ago, I was delighted to come across the correspondences

$\mathbf{u} \times \mathbf{v} \leftrightarrow$ 

and $\mathbf{u} \cdot \mathbf{v} \leftrightarrow$ 

in a book by the physicist G.E. Stedman [16]. Moreover, there was a way to perform *rigorous* calculations using the diagrams. In subsequent exploration of the ideas in the book, I found truly satisfying explanations for the above questions.

The main tool in Stedman's book is a *spin network* (also called a *trace diagram*), which is a type of graph labeled by representations of a particular group. Spin networks are similar in appearance to Feynman diagrams, but are useful for different kinds of problems. The earliest work with diagrams of this sort appears to be [8], in which spin networks of rank 2 were used as a tool for investigating quantized angular momenta. The name *spin network* is attributed to Roger Penrose, who used them to construct a discrete model of space-time [12, 13]. In modern terminology, a spin network is a graph whose edges are labeled by representations of a group and whose vertices are labeled by *intertwiners* or maps between tensor powers of representations [9]. Predrag Cvitanovic has shown how to construct the diagrams

Date: January 25, 2008.

for any Lie group, and actually found a novel classification of the classical and exceptional Lie groups using this approach [2, ch. 21].

There are numerous applications of spin networks. They are a standard tool used by physicists for studying various types of particle interactions [2, 16]. Their generalization to quantum groups forms the backbone of skein theory and many 3-manifold invariants [4, chs. I.9/I.16]. In combinatorics, they are closely related to chromatic polynomials of graphs [4, chs. II.6/II.7], and actually the *four-color theorem* can be restated in terms of spin networks [5]. Spin networks also play a role in geometry. They can be used to describe the *character variety* of a surface, which encodes the geometric structures that can be placed on the surface [6, 14, 15]. There are indications that they may also be a powerful tool for the study of matrix invariants [1], [15, Cor. 6.1].

Amidst all these applications, there is a surprising lack of the most basic application of the diagrams: linear algebra. In this paper, spin networks are additionally labeled with matrices and called *trace diagrams* to emphasize this application. Trace diagrams provide simple depictions of traces, determinants, and other linear algebra fare in a way that is mathematically rigorous. The paper concludes with an elegant diagrammatic proof of the Cayley-Hamilton theorem.

The emphasis in this paper is on *illumination*, and in particular on how diagrammatic techniques have the power to both prove and explain. For this reason, several examples are included, and more enlightening proofs are offered. While diagrammatic methods may seem unfamiliar at first, in the end they offer a profound insight into some of the most fundamental structures of linear algebra, such as the determinant, the adjugate matrix, and the characteristic equation. We hope that by the end of the paper the reader is both more comfortable with the appearance of diagrams in mathematics and convinced that these informal "doodles" can actually be quite useful.

2. Symmetry in Linear Algebra and Tensor Products

Perhaps the greatest contribution of diagrams is their ability to capture symmetries inherent in linear algebra. This section reviews some of those symmetries as well as multilinear or *tensor* algebra, the algebraic framework used to interpret diagrams.

The inner product is an example of a *symmetric* (or *commutative*) function, since $\mathbf{u}\cdot\mathbf{v}=\mathbf{v}\cdot\mathbf{u}$. Similarly, the cross product is *antisymmetric* because $\mathbf{u}\times\mathbf{v}=-\mathbf{v}\times\mathbf{u}$. Both functions are *multilinear*, since they are linear in each factor. For example:

$$(\mathbf{u}+\lambda\mathbf{v})\times\mathbf{w}=\mathbf{u}\times\mathbf{w}+\lambda(\mathbf{v}\times\mathbf{w}).$$

Multilinear functions on vector spaces can *always* be considered as functions on tensor products. Informally, a *tensor product* of two complex vector spaces consists of finite sums of vector pairs $(\mathbf{u},\mathbf{v})$ subject to the relation

$$(\lambda\mathbf{u},\mathbf{v})=\lambda(\mathbf{u},\mathbf{v})=(\mathbf{u},\lambda\mathbf{v})$$

for $\lambda\in\mathbb{C}$. The corresponding element is usually written $\mathbf{u}\otimes\mathbf{v}$. If $V=\mathbb{C}^3$, then the domain of both $\cdot$ and $\times$ can be taken to be $V\otimes V$, so that $\cdot: V\otimes V\to\mathbb{C}$ and $\times: V\otimes V\to V$. For a rigorous construction of tensor products, see Appendix B in [3].

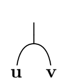
As another example with $V=\mathbb{C}^3$, the determinant function can be written as a function $\mathsf{det}: V\otimes V\otimes V\to\mathbb{C}$, with $\mathbf{u}\otimes\mathbf{v}\otimes\mathbf{w}\mapsto\mathsf{det}[\mathbf{u}\ \ \mathbf{v}\ \ \mathbf{w}]$. The multilinearity of the determinant is usually described as being able to factor constants multiplying a single matrix column outside the determinant. Since switching columns in a matrix introduces a sign on the determinant, the determinant is an *antisymmetric* function.

Remark 2.1. Antisymmetric functions can also be considered as functions on an *exterior (wedge) product* of vector spaces. Indeed, *exterior algebra* provides traditional answers to some of the questions mentioned in the introduction.

3. A Taste of Trace

This section presents *3-vector diagrams* and *trace diagrams* informally as heuristics to aid in calculations. The ideas introduced here are made rigorous in Section 4. Some of the material included here is paralleled in [16, ch. 1].

Consider again the correspondences

$\mathbf{u} \times \mathbf{v} \leftrightarrow$  and $\mathbf{u} \cdot \mathbf{v} \leftrightarrow$ .

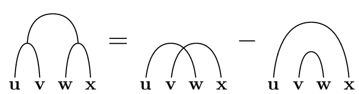
These diagrams are read "bottom to top" so that the inputs $\mathbf{u}$ and $\mathbf{v}$ occur along the bottom, and the output(s) occur along the top. In the case of the cross product, a single output strand implies a *vector* output; in the case of the inner product, the absence of an output strand implies a *scalar* output.

The notation permits an alternate expression of standard vector identities:

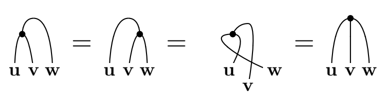
Example 3.1. Draw the identity

$$(\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{w} \times \mathbf{x}) = (\mathbf{u} \cdot \mathbf{w})(\mathbf{v} \cdot \mathbf{x}) - (\mathbf{u} \cdot \mathbf{x})(\mathbf{v} \cdot \mathbf{w}).$$

Solution. Keeping the vector inputs in the same order, the diagram is:

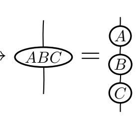
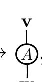


Exercise 3.2. What vector identity does this diagram represent?



The reader is encouraged to guess the meaning of the fourth term (later described in Theorem 4.2).

Now suppose that matrix and vector multiplication could also be encoded diagrammatically, according to the rules

$ABC \leftrightarrow$  and $\mathbf{v}^T A \mathbf{w} \leftrightarrow$ ,

where ABC is a legal matrix product and $\mathbf{v}^T$ represents the transpose of $\mathbf{v}$. Then matrix elements can be represented diagrammatically as well. If the standard row and column bases for $\mathbb{C}^n$ are denoted by $\{\hat{\mathbf{e}}^i\}$ and $\{\hat{\mathbf{e}}_j\}$, respectively, then

$a_{ij} = \hat{\mathbf{e}}^i A \hat{\mathbf{e}}_j \leftrightarrow$ .

Using this notation, trace and determinant diagrams may also be constructed.

Example 3.3. Find a diagrammatic representation of the trace $\mathsf{tr}(A)$.

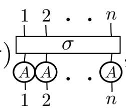
Solution.

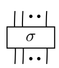
$$\mathsf{tr}(A) \leftrightarrow \sum_{i=1}^{n}$$ 

Example 3.4. Find a diagrammatic representation of the determinant

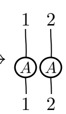
$$\mathsf{det}(A) = \sum_{\sigma \in \Sigma_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}. \tag{1}$$

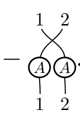
Solution. One approach is to introduce new notation:

$$\mathsf{det}(A) \leftrightarrow \sum_{\sigma \in \Sigma_n} \operatorname{sgn}(\sigma)$$ 

where 

represents a permutation on the n strands, and $\operatorname{sgn}(\sigma)$ is the *sign* or *signature* of the permutation σ. For example, if $n = 2$, then

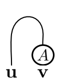
$$\mathsf{det}(A) \leftrightarrow$$ 

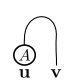
$-$ 

4. Topological Invariance: No Rules, Just Right

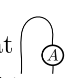
What if there were a set of rules for manipulating the diagrams that was compatible with their interpretations as functions? This section exhibits just such a correspondence, which exists *provided the graphs are given a little extra structure.* Several classical proofs become trivial in this context, as the complexity is funneled from the proof into the construction of a diagram.

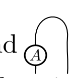
The diagrams are essentially given the structure of a graph whose vertices have degree 1 or n only, and whose edges are labeled by matrices. But there are a few difficulties with ensuring a diagram's function is well-defined. First,

(2) 

$= \mathbf{u} \cdot (A\mathbf{v}) = (A^T\mathbf{u}) \cdot \mathbf{v}$ but 

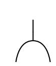
$= (A\mathbf{u}) \cdot \mathbf{v}$.

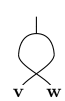
The problem here is that 

and 

are indistinguishable as graphs, but represent different functions. This problem is resolved in Definition 4.1 by requiring the graphs to be *oriented*.

The second problem is

(3) 

$= \mathbf{v} \otimes \mathbf{w}$ but 

$= \mathbf{w} \otimes \mathbf{v} = -\mathbf{v} \otimes \mathbf{w}$.

This time, the problem is that the ordering of edges at a vertex matters, so the diagram *must include a way to order the edges adjacent to a vertex.*

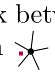
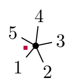
These two extra pieces of structure are incorporated into the formal definition of a trace diagram [14, Defn. 4.2], cf. [2, ch. 6], [15, Defn. 3.2].

Definition 4.1. An *n-trace diagram* is an oriented graph with edges labeled by $n \times n$ matrices whose vertices

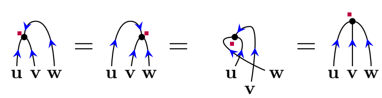
(1) have degree 1 or n only;
(2) are sources or sinks;
(3) are labeled by an ordering of edges incident to the vertex.

If there are no degree 1 vertices, the trace diagram is said to be *closed*. Otherwise, the degree 1 vertices are divided into ordered sets of *inputs* and *outputs*.

The object defined in this way are *combinatorial in nature*, so that diagrams are identical whenever they have the same underlying graph and the same labels.

By convention, trace diagrams are typically drawn in the plane with input vertices at the bottom of the drawing and output vertices at the top. A simple way to maintain the required ordering is to draw a mark between two edges at the vertex, called a *ciliation*. For example, the ciliation on  implies the counterclockwise ordering 

. Given the combinatorial definition, it does not matter how a diagram is drawn, provided the order of the inputs and outputs and the positions of the ciliations do not change. Thus



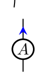
holds automatically since the diagrams are identical as labeled graphs.

The power of the diagrams lies in the fact that they can be identified with certain kinds of functions. The actual function depends upon the way the diagram is drawn in the plane. Since a single trace diagram may be drawn in many different ways, it may correspond to several different functions. However, the following theorem says that these 'different' functions are algebraically equivalent [14, Prop. 4.3].

Theorem 4.2 (Fundamental theorem of trace diagrams)**.** *Let $V = \mathbb{C}^n$. There is a well-defined correspondence between trace diagrams with k inputs and l outputs and functions $V^{\otimes k} \to V^{\otimes l}$. In particular, every decomposition of a trace diagram into the following basic maps gives the same function (or scalar if the diagram is closed):*

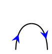
 $: \mathbf{v} \longmapsto \mathbf{v};$

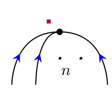
 $: \mathbf{v}^T \longmapsto \mathbf{v}^T;$



$: \mathbf{v} \longmapsto A \cdot \mathbf{v};$

 $: 1 \longmapsto \sum_{i=1}^{n} \hat{\mathbf{e}}^i \otimes \hat{\mathbf{e}}_i;$

 $: \mathbf{v} \otimes \mathbf{w}^T \longmapsto \mathbf{v} \cdot \mathbf{w};$

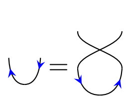


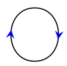
$: \mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_n \longmapsto \mathsf{det}[\mathbf{v}_1 \ \ \cdots \ \ \mathbf{v}_n].$

In addition, the n-vertex with opposite orientation represents the determinant of row vectors, $\mathbf{v}_1^T \otimes \cdots \otimes \mathbf{v}_n^T \longmapsto \mathsf{det}[\mathbf{v}_1 \ \ \cdots \ \ \mathbf{v}_n]$.

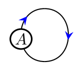
The decomposition assumes the diagram is in 'general position' relative to the vertical direction, and involves "chopping" the diagram up into these particular pieces. Vertical stacking of diagrams corresponds to the composition of functions. Several examples follow.

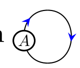
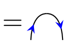
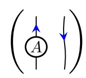
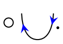
Example 4.3. To compute , switch the order of the outputs:

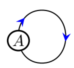
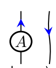
 $: 1$  $\longmapsto \sum_i \hat{\mathbf{e}}^i \otimes \hat{\mathbf{e}}_i$  $\longmapsto \sum_i \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}^i.$

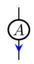
Exercise 4.4. Show that  $: 1 \mapsto \dim(V) = n$. Thus, the basic loop is equivalent to the dimension n of the vector space.

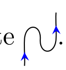
Here is the simplest closed diagram with a matrix, and the reason for the terminology 'trace' diagram:

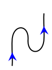
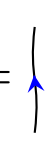
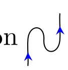
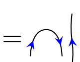
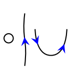
Example 4.5. Show that  $= \mathsf{tr}(A)$.

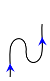
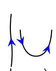
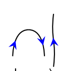
Solution. Decompose the diagram  $=$  $\circ$  $\circ$ . Then

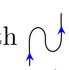
 $: 1$  $\longmapsto \sum_i \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}^i$  $\longmapsto \sum_i (A\hat{\mathbf{e}}_i) \otimes \hat{\mathbf{e}}^i$  $\longmapsto \sum_i (A\hat{\mathbf{e}}_i) \cdot \hat{\mathbf{e}}_i = \sum_i a_{ii} = \mathsf{tr}(A).$

Exercise 4.6. Show that the difficulty encountered in equation (2) is no longer a problem, since  $: \mathbf{v}^T \longmapsto (A\mathbf{v})^T$.

Example 4.7. Compute .

Solution.  $=$  since the decomposition  $=$  $\circ$  gives

 $: \mathbf{v}$  $\longmapsto \sum_i \mathbf{v} \otimes \hat{\mathbf{e}}^i \otimes \hat{\mathbf{e}}_i$  $\longmapsto \sum_i (\mathbf{v} \cdot \hat{\mathbf{e}}_i)\hat{\mathbf{e}}_i = \sum_i \mathbf{v}_i \hat{\mathbf{e}}_i = \mathbf{v}.$

This last result was expected since both  and  are equivalent forms of the same trace diagram. Moreover, this fact essentially proves Theorem 4.2:

Sketch proof of Theorem 4.2. Consider two drawings that represent the same trace diagram and can be decomposed into the components in the statement of the theorem. Then the drawing must look 'locally' like one of these component maps: all matrices must occur on upward oriented strands, and all nodes must occur at local maxima. The only possible difference between the two diagrams is therefore the number of "kinks" between strands between the nodes, the matrix labels, and the inputs and outputs. But the calculation in Example 4.7 shows that adding or removing a kink does not change the underlying function. Therefore, any two drawings of the same trace diagram must represent the same function. □

Remark 4.8. The correspondence between diagrams and functions established by Theorem 4.2 asserts the existence of a *functor* between the *category* of trace diagrams and the *category* of multilinear functions. This functor is well-known in the literature on spin networks. This fact is commonly proven by demonstrating that the algebra of diagrams modulo specific relations is isomorphic to the algebra of functions modulo their relations [15, Thm. 3.7].

Remark 4.9. In particular cases, the diagrams can be simplified somewhat.

If n is odd, then the determinant is *cyclically invariant*:

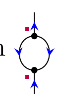
$$\mathsf{det}[\mathbf{v}_1 \ \ \mathbf{v}_2 \ \ \cdots \ \ \mathbf{v}_n] = \mathsf{det}[\mathbf{v}_2 \ \ \cdots \ \ \mathbf{v}_n \ \ \mathbf{v}_1],$$

so the ciliation is unnecessary; the cyclic orientation implied by the drawing of a diagram in the plane is sufficient.

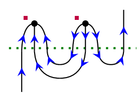
For $n = 2$, the orientation is unnecessary, and frequently the cap $\frown$ is defined to be $\mathbf{v} \otimes \mathbf{w} \mapsto i \cdot \mathsf{det}[\mathbf{v} \ \ \mathbf{w}]$ rather than the inner product. With the factor of i, there is no need to keep track of ciliations, and the diagrams are simply collections of arcs and loops labeled by matrices [4, p. 444].

5. Simplifying Diagrammatic Calculations

Theorem 4.2 says that a function's diagram may be computed by transforming the diagram into some sort of "standard form." In practice, it is better to have a working knowledge of the functions of the few basic diagrams described in this section. This is abundantly clear in the following computation:

Example 5.1. Show that the diagram  is the function $\mathbf{v} \mapsto -2\mathbf{v}$.

Solution. Arrange the diagram so it decomposes into the basic component maps. Evaluate the map in two steps (divided by the dotted line in the figure):



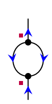
$$: \mathbf{v} \longmapsto \sum_{i,j,k} \mathbf{v} \otimes \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}^j \otimes \hat{\mathbf{e}}^i \otimes \hat{\mathbf{e}}^k \otimes \hat{\mathbf{e}}_k$$

$$\longmapsto \sum_{i\neq j, j\neq k, i\neq k} \mathsf{det}[\mathbf{v} \ \ \hat{\mathbf{e}}_i \ \ \hat{\mathbf{e}}_j]\mathsf{det}[\hat{\mathbf{e}}_j \ \ \hat{\mathbf{e}}_i \ \ \hat{\mathbf{e}}_k]^T \hat{\mathbf{e}}_k.$$

But

$$\mathsf{det}[\mathbf{v} \ \ \hat{\mathbf{e}}_i \ \ \hat{\mathbf{e}}_j]\mathsf{det}[\hat{\mathbf{e}}_j \ \ \hat{\mathbf{e}}_i \ \ \hat{\mathbf{e}}_k]^T = \mathsf{det}\begin{bmatrix} \mathbf{v}\cdot\hat{\mathbf{e}}_j & \mathbf{v}\cdot\hat{\mathbf{e}}_i & \mathbf{v}\cdot\hat{\mathbf{e}}_k \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = -\mathbf{v}\cdot\hat{\mathbf{e}}_k.$$

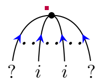
Therefore,



$$: \mathbf{v} \mapsto \sum_{i\neq j, j\neq k, i\neq k} (-\mathbf{v}\cdot\hat{\mathbf{e}}_k)\hat{\mathbf{e}}_k = -2\sum_k (\mathbf{v}\cdot\hat{\mathbf{e}}_k)\hat{\mathbf{e}}_k = -2\mathbf{v}.$$

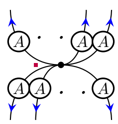
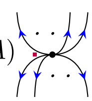
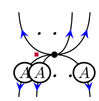
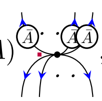
Fortunately, there is an easier approach. The remainder of this section gives several results which make such technical calculations unnecessary. Of course, the proofs of these results are themselves quite technical, and several of the proofs are deferred to the Appendix.

Proposition 5.2. *The n-vertices of trace diagrams are antisymmetric, meaning a sign is introduced if either (i) two inputs at a node are switched or (ii) a ciliation* *is moved across an edge. Alternately, if any collection of inputs at an n-vertex are linearly dependent, the diagram evaluates to zero. In particular,*

 $= 0.$

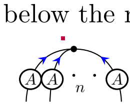
Proof. These facts are restatements of standard facts regarding the determinant: switching two columns introduces a sign, while the determinant of a matrix with linearly dependent rows or columns is zero. $\square$

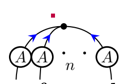
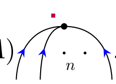
Proposition 5.3 (Matrix invariance)**.** *Matrices may be 'cancelled' at any node with the addition of a determinant factor. In particular,*

 $= \mathsf{det}(A)$  *and*  $= \mathsf{det}(A)$ ,

where $\bar{A} = A^{-1}$.

Proof. By Theorem 4.2, all diagrams with nodes may be computed by first drawing the nodes as local maxima. Therefore, it suffices to check the case where all edges are below the node.

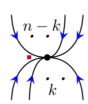
 $: \mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_n \mapsto \mathsf{det}[A\mathbf{v}_1 \ \ \cdots \ \ A\mathbf{v}_n] = \mathsf{det}(A)\mathsf{det}[\mathbf{v}_1 \ \ \cdots \ \ \mathbf{v}_n].$

This calculation shows that  $= \mathsf{det}(A)$ .

The second relation follows from the first. $\square$

The next two propositions, whose proofs are found in the Appendix, describe the functions at a diagram's nodes:

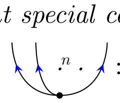
Proposition 5.4. *The following diagram is a "complemental antisymmetrizer":*

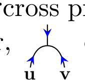
(4)  $: \hat{\mathbf{e}}_{a_1} \otimes \cdots \otimes \hat{\mathbf{e}}_{a_k} \longmapsto \sum_{\sigma \in \Sigma_{\tilde{A}}} \mathrm{sgn}(\mathsf{Id}_A|\sigma)\hat{\mathbf{e}}^{\sigma(n)} \otimes \cdots \otimes \hat{\mathbf{e}}^{\sigma(k+1)},$

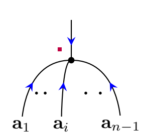
where the sum is over permutations on $\tilde{A} = N \setminus \{a_1, \ldots, a_k\} = \{\tilde{a}_1, \ldots, \tilde{a}_{n-k}\}$ *and* $\mathrm{sgn}(\mathsf{Id}_A|\sigma)$ *is the* sign *or* signature *of the permutation*

$$\begin{pmatrix} a_1 & \cdots & a_k & \sigma(\tilde{a}_1) & \cdots & \sigma(\tilde{a}_{n-k}) \\ 1 & \cdots & k & k+1 & \cdots & n \end{pmatrix}.$$

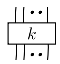
One important special case is the codeterminant *map*

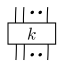
(5)  $: 1 \mapsto \sum_{\sigma \in \Sigma_n} (-1)^{\lfloor \frac{n}{2} \rfloor} \mathrm{sgn}(\sigma)\hat{\mathbf{e}}_{\sigma(1)} \otimes \cdots \otimes \hat{\mathbf{e}}_{\sigma(n)}.$

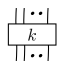
Remark 5.5. Proposition 5.4 shows that there is a way to define a "cross product" in *any* dimension, although $n-1$ inputs are required. In particular,  extends to the family of diagrams

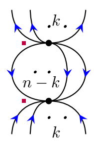
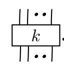
 .

This is the analog of the *wedge product* alluded to in Remark 2.1.

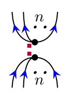
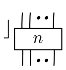
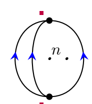
Proposition 5.6. *Let*  *denote the antisymmetrizer on k vertices defined by*

(6)  $: \mathbf{a}_1 \otimes \cdots \otimes \mathbf{a}_k \longmapsto \sum_{\sigma \in \Sigma_k} \operatorname{sgn}(\sigma) \mathbf{a}_{\sigma(1)} \otimes \cdots \otimes \mathbf{a}_{\sigma(k)}.$

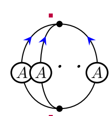
If $k > n$, then  $= 0$. *Otherwise, for $0 \leq k \leq n$,*

 $= (-1)^{\lfloor \frac{n}{2} \rfloor}(n-k)!$ .

In the special cases $k = n$ and $k = 0$:

(7)  $= (-1)^{\lfloor \frac{n}{2} \rfloor}$   $= (-1)^{\lfloor \frac{n}{2} \rfloor} n!$

Corollary 5.7. *Combining Propositions 5.3 and 5.6, one can find a closed diagram representing the determinant:*

(8)  $= (-1)^{\lfloor \frac{n}{2} \rfloor} n! \mathsf{det}(A).$

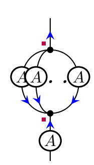
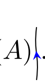
Remark 5.8. These propositions offer one explanation of the importance of the trace and determinant: the trace is the simplest closed diagram with a matrix (Example 4.5), while the determinant is the "simplest" closed diagram with two nodes and matrices on all the edges (Proposition 5.7). This begs the question: what if some of the matrices in (8) are removed? The answer is provided in the proof of Proposition 7.3; they are essentially the coefficients of the characteristic polynomial!

6. The Elegant Adjugate and Clandestine Cramer

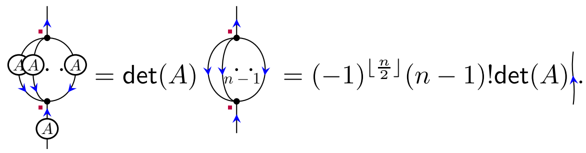
We now turn to the main results of this paper, the diagrammatic proofs of the adjugate formula, Cramer's rule, and the Cayley-Hamilton theorem. Each diagrammatic result is easy to prove using the results of the previous section, but it is not always so easy to show that the diagrammatic relation is equivalent to its standard definition in linear algebra. The reader unfamiliar with the linear algebra concepts may wish to consult a linear algebra text such as [7].

The diagrammatic version of the adjugate formula $\mathsf{adj}(A) \cdot A = \mathsf{det}(A) I$ follows.

Proposition 6.1 (Diagrammatic adjugate formula)**.**

(9)  $= (-1)^{\lfloor \frac{n}{2} \rfloor}(n-1)! \mathsf{det}(A)$ .

Proof. Use Propositions 5.3 and 5.6:



□

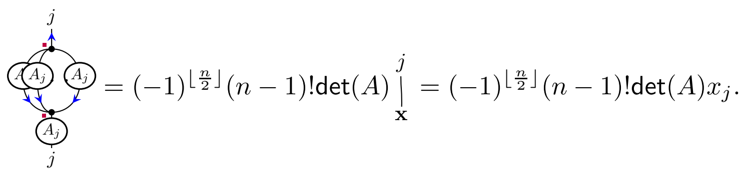
Cramer's rule is actually 'hiding' in this diagram. Recall that given a matrix $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_n]$, *Cramer's rule* states that for the matrix equation $A\mathbf{x} = \mathbf{b}$, the elements of the solution $\mathbf{x}$ are given by

$$x_j = \frac{\mathsf{det}(A_j)}{\mathsf{det}(A)} = \frac{\mathsf{det}[\mathbf{a}_1 \ \cdots \ \mathbf{a}_{j-1} \ \mathbf{b} \ \mathbf{a}_{j+1} \ \cdots \ \mathbf{a}_n]}{\mathsf{det}(A)}.$$

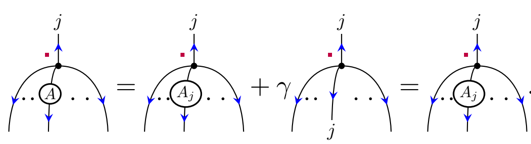
Here A_j is the matrix obtained from A by replacing the jth column with the vector $\mathbf{b}$.

Proposition 6.2 (Diagrammatic Cramer's rule)**.** *Let $A\mathbf{x} = \mathbf{b}$, and suppose that the columns of A_j are identical to the columns of A, except that the jth column of A_j is $\mathbf{b}$. Then*

(10)

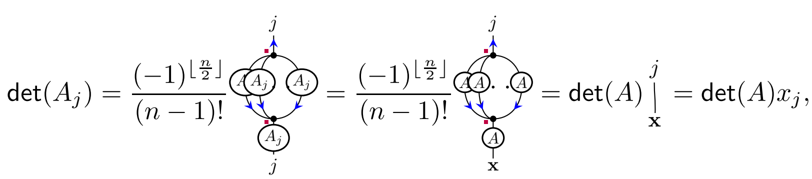


Proof. Given an arbitrary vector $\mathbf{v}$, $\mathbf{v}^T A = \mathbf{v}^T A_j + \gamma \hat{\mathbf{e}}^j$ for some γ. It follows from Proposition 5.2 that



Given this and the fact that $A_j\hat{\mathbf{e}}_j = \mathbf{b} = A\mathbf{x}$,

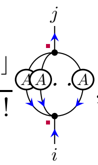
(11)



where the first and third steps follow from Propositions 5.3 and 5.6. This proves (10), as well as Cramer's rule. □

It should be clear from (10) that Proposition 6.2 implies Cramer's rule, but it may not be clear where the adjugate matrix shows up in Proposition 6.1. In traditional texts, the adjugate matrix is constructed by (i) building a matrix of cofactors, (ii) applying sign changes along a checkerboard pattern, and (iii) transposing the result. In contrast, the diagram of $\mathsf{adj}(A)$ is quite simple:

Proposition 6.3. *The matrix elements of* $\mathsf{adj}(A)$ *may be expressed as*

$$(\mathsf{adj}(A))_{ji} = \frac{(-1)^{\lfloor \frac{n}{2} \rfloor}}{(n-1)!}$$ ,

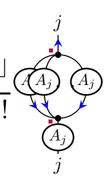
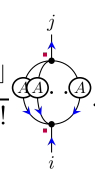
Proof. The matrix element of the diagram must somehow encode all the traditional steps required for computing the adjugate. But how? First, 'crossing out' occurs when the basis elements $\hat{\mathbf{e}}_i$ and $\hat{\mathbf{e}}_j$ are placed adjacent to the nodes, sign changes are encoded in the orientation of the node, and the transpose comes into play because the matrices in the diagram are along downward-oriented strands.

Formally, note that the signed (i,j)-cofactor may be expressed as the determinant of the matrix

$$A_j = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_{j-1} \ \hat{\mathbf{e}}_i \ \mathbf{a}_{j+1} \ \cdots \ \mathbf{a}_n] = \begin{bmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & 1 & \cdots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{bmatrix},$$

in which the jth column of A is replaced by $\hat{\mathbf{e}}_i$. By definition, $(\mathsf{adj}(A))_{ji} = \mathsf{det}(A_j)$.

The following equation, which is remarkably similar to (11), shows how to find the adjugate matrix:

(12) $$(\mathsf{adj}(A))_{ji} = \mathsf{det}(A_j) = \frac{(-1)^{\lfloor \frac{n}{2} \rfloor}}{(n-1)!}$$  $$= \frac{(-1)^{\lfloor \frac{n}{2} \rfloor}}{(n-1)!}$$ .

This establishes the result. ∎

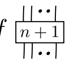
7. The Not-So-Characteristic Equation

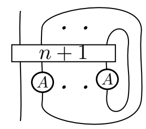
Recall that for an $n \times n$ matrix A, $p(\lambda) = \mathsf{det}(A - \lambda I)$ is a degree n polynomial in λ called the *characteristic polynomial* of A. By the fundamental theorem of algebra, it must have n real or complex roots, counted with multiplicity, which are the *eigenvalues* of A. The *Cayley-Hamilton theorem* says that a matrix satisfies its own characteristic equation $p(A) = 0$. For example, for any 2×2 matrix, the characteristic polynomial is

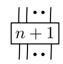
$$\mathsf{det}(A - \lambda I) = \lambda^2 - (\mathsf{tr}A)\lambda + \mathsf{det}(A) = 0.$$

Consequently, $A^2 - (\mathsf{tr}A)A + \mathsf{det}(A)I = 0$.

The final result is the "not-so-characteristic" equation, a diagrammatic statement equivalent to the Cayley-Hamilton theorem whose proof is perhaps the simplest in this entire paper.

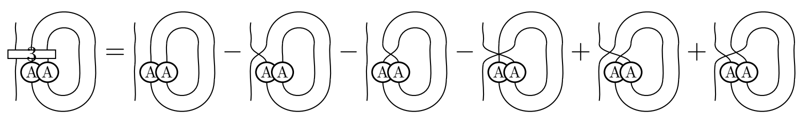
Theorem 7.1 (Diagrammatic Cayley-Hamilton theorem)**.** *If*  *represents the antisymmetrizer on* $n+1$ *vectors* (6)*, then*

(13)  $= 0.$

Proof. Apply Proposition 5.6, which states that  $= 0$. ◻

The reader may not yet be convinced that the diagram is in fact the characteristic equation. How does this relate to the formula $\mathsf{det}(A - \lambda I) = 0$? Consider the following example for 2×2 matrices:

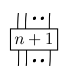
Example 7.2.



$$= \mathsf{tr}(A)^2 I - \mathsf{tr}(A)A - \mathsf{tr}(A^2)I - \mathsf{tr}(A)A + A^2 + A^2$$
$$= 2(A^2 - \mathsf{tr}(A)A + \mathsf{det}(A)I) = 0,$$

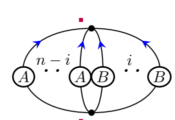
using the fact that $\mathsf{det}(A) = \frac{1}{2}\left(\mathsf{tr}(A)^2 - \mathsf{tr}(A^2)\right)$.

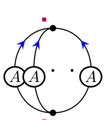
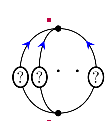
As is usually the case in diagrammatic statements, the "hard part" is demonstrating the equivalence to the traditional construction. In the remainder of this section, we show that the coefficients of A^i in the expansion of (13) are precisely $n!$ times the coefficients of the characteristic polynomial.

Proposition 7.3. *When the antisymmetrizer*  *in* (13) *is expanded, the coefficients of* A^i *are equal to* $n!$ *times the coefficients of* λ^i *in the characteristic polynomial* $\mathsf{det}(A - \lambda I) = 0$.

First, how can the coefficients of the characteristic polynomial be described diagrammatically?

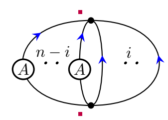
Lemma 7.4. *Given* $A, B \in M_{n\times n}$*, the determinant sum* $\mathsf{det}(A + B)$ *is expressed diagrammatically as*

$$\mathsf{det}(A+B) = (-1)^{\lfloor \frac{n}{2} \rfloor} \sum_{i=0}^{n} \frac{1}{i!(n-i)!}$$  .

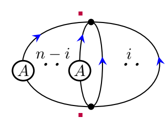
Proof. Replace A in $\mathsf{det}(A) = \frac{(-1)^{\lfloor \frac{n}{2} \rfloor}}{n!}$  with $A + B$. The result consists of 2^n diagrams, each of which has the form  where ? is either A or B. But each of the strands of these diagrams can be rearranged, by Proposition 5.2:

swapping adjacent strands introduces a sign of $(-1)\cdot(-1) = +1$. For this reason, diagrams may be grouped by the number of A's and B's occurring in each. The result follows since there are precisely $\binom{n}{i} = \frac{n!}{i!(n-i)!}$ summands that have i B strands. ∎

Corollary 7.5. *In terms of diagrams, the characteristic polynomial is*

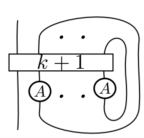
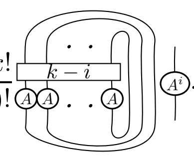
(14) $$\mathsf{det}(A - \lambda I) = \sum_{i=0}^{n} \left(\frac{(-1)^{i+\lfloor \frac{n}{2} \rfloor}}{i!(n-i)!} \right.$$  $$\left. \vphantom{\frac{1}{1}} \right) \lambda^i = \sum_{i=0}^{n} c_i \lambda^i,$$

where

(15) $$c_i = \frac{(-1)^{i+\lfloor \frac{n}{2} \rfloor}}{i!(n-i)!}$$  .

This means that the coefficients of the characteristic polynomial are, up to a constant factor, the $n+1$ "simplest" diagrams with two nodes. Are these also the coefficients of A^i in (13)? The next lemma provides the combinatorial decomposition of the antisymmetrizer that is required to demonstrate this fact.

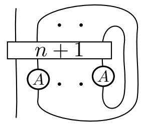
Lemma 7.6. *For any k with $0 \le k \le n$,*

(16)  $$= \sum_{i=0}^{k} \frac{(-1)^i k!}{(k-i)!}$$  .

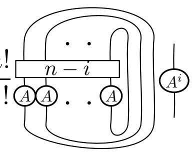
Proof. Choose a summand corresponding to a permutation $\sigma \in \Sigma_{k+1}$. Write the permutation as $\sigma = \tau\nu$ where τ is the cycle containing the first element and ν contains the remaining cycles. Then $|\tau| = i+1$ in the summand, since the left strand passes through i A strands, and the summand contributes to the A^i term.

There are $\frac{k!}{(k-i)!}$ ways to select τ such that $|\tau| = i+1$. In each case, the remainder of the diagram is closed and all choices of ν can be consolidated into a single $\boxed{k-i}$ term. The sign of a summand is given by $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\tau)\mathrm{sgn}(\nu) = (-1)^i\mathrm{sgn}(\nu)$, and $\mathrm{sgn}(\nu)$ is also incorporated within the $\boxed{k-i}$. Hence, the coefficient of A^i is $(-1)^i \frac{k!}{(k-i)!}$ times the closed diagram shown. ∎

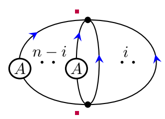
Proof of Proposition 7.3. Letting $k = n$ in Lemma 7.6 and applying Proposition 5.6 shows that



$$= \sum_{i=0}^{n} \frac{(-1)^i n!}{(n-i)!}$$



$$= \sum_{i=0}^{n} \frac{(-1)^i (-1)^{\lfloor \frac{n}{2} \rfloor} n!}{(n-i)!i!}$$



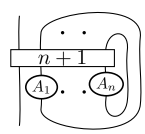
$$A^i$$

$$= n! \sum_{i=0}^{n} c_i A^i. \qquad \square$$

8. Closing Thoughts

The results contained in this paper are merely a beginning, and an invitation to further exploration. It is an easy exercise to flip through a linear algebra textbook and find more results than can be expressed diagrammatically. For example, trace diagrams are remarkably good at capturing the concept of a matrix minor. Steven Morse has shown that both the *Jacobi determinant theorem* and Charles Dodgson's *condensation method* for calculating determinants have simple diagrammatic proofs [10]. Conversely, every diagrammatic relation has a linear algebra interpretation, so they are quite good at "generating" new formulas.

Trace diagrams are also a powerful technique for generating trace identities [1], [15, Cor. 6.1]. Indeed, for n matrices $\{A_1, \ldots, A_n\}$ it is an immediate consequence of Proposition 5.6 that



$$= 0.$$

This equation, which generalizes the characteristic equation, is sometimes called a *polarization* of the characteristic polynomial, and is a fundamental tool used in the study of matrix invariants [11, ch. 6]. It seems plausible that some problems arising in invariant theory that are intractable with classical techniques might yield to the diagrammatic method.

Finally, there are strong connections between trace diagrams and coloring theory. In particular, n-trace diagrams without matrices can be evaluated by simply counting the number of possible coloring of edges by n colors so that no two edges at the same vertex have the same color. Several important open questions in graph theory can be restated in the language of trace diagrams [14, ch. 8].

Perhaps the most promising feature of trace diagrams, however, is their ability to bring together these diverse areas. For example, 3-trace diagrams are a natural setting for working with the cross product, with trace identities and the invariant theory of 3×3 matrices, and with 3-edge colorable graphs. Several connections between these distinct areas have already been discovered using standard techniques. It seems likely to this author, however, that several more connections may be revealed by trace diagrams in the future.

Appendix A. Proofs of Node Identities

The proofs that follow require some facts regarding permutations and some new notation. Given a permutation α on the set $N = \{1, 2, \ldots, n\}$ taking $i \mapsto \alpha(i)$, denote by $\mathrm{sgn}(\alpha) = \mathsf{det}\,[\hat{\mathbf{e}}_{\alpha_1} \ \cdots \ \hat{\mathbf{e}}_{\alpha_m}]$ the *sign* or *signature* of the permutation.

If $A = (a_1, \ldots, a_k)$ represents an *ordered* subset of N (no repetitions allowed), let $\tilde{A}$ represent a particular ordering $(\tilde{a}_1, \ldots, \tilde{a}_{n-k})$ of the elements in $N \setminus A$. Given permutations $\tau \in \Sigma_A$ of the elements of A and $\sigma \in \Sigma_{\tilde{A}}$ of the elements not in A, let $(\tau|\sigma)$ represent the permutation on N given by

$$(\tau|\sigma) = \begin{pmatrix} \tau(a_1) & \cdots & \tau(a_k) & \sigma(\tilde{a}_1) & \cdots & \sigma(\tilde{a}_{n-k}) \\ 1 & \cdots & k & k+1 & \cdots & n \end{pmatrix}.$$

Note that

$$\mathrm{sgn}(\tau|\sigma) = \mathrm{sgn}(\mathsf{Id}_A|\mathsf{Id}_{\tilde{A}}) \cdot \mathrm{sgn}(\tau) \cdot \mathrm{sgn}(\sigma),$$

where Id_A represents the identity permutation on A.

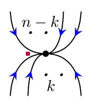
Let $\overleftarrow{\alpha}$ denote the permutation reversing the order of elements in the permutation α, hence taking $i \mapsto \alpha(n+1-i)$. Then $\mathrm{sgn}(\overleftarrow{\alpha}) = (-1)^{\lfloor \frac{n}{2} \rfloor}\mathrm{sgn}(\alpha)$, since $\lfloor \frac{n}{2} \rfloor$ transpositions are required to change $\alpha \leftrightarrow \overleftarrow{\alpha}$.

The following lemma is used in the proof of Proposition 5.6.

Lemma A.1.

$$\begin{aligned}\mathrm{sgn}(\mathsf{Id}_A|\sigma) \cdot \mathrm{sgn}(\overleftarrow{\sigma}|\overleftarrow{\tau}) &= \mathrm{sgn}(\mathsf{Id}_A|\mathsf{Id}_{\tilde{A}})\mathrm{sgn}(\sigma) \cdot (-1)^{\lfloor \frac{n}{2} \rfloor}\mathrm{sgn}(\mathsf{Id}_A|\mathsf{Id}_{\tilde{A}})\mathrm{sgn}(\tau)\mathrm{sgn}(\sigma) \\ &= (-1)^{\lfloor \frac{n}{2} \rfloor}\mathrm{sgn}(\mathsf{Id}_A|\mathsf{Id}_{\tilde{A}})^2\mathrm{sgn}(\sigma)^2\mathrm{sgn}(\tau) \\ &= (-1)^{\lfloor \frac{n}{2} \rfloor}\mathrm{sgn}(\tau).\end{aligned}$$

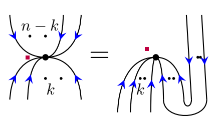
Proposition A.2 (Proposition 5.4)**.**



$$: \hat{\mathbf{e}}_{a_1} \otimes \cdots \otimes \hat{\mathbf{e}}_{a_k} \longmapsto \sum_{\sigma \in \Sigma_{\tilde{A}}} \mathrm{sgn}(\mathsf{Id}_A|\sigma)\hat{\mathbf{e}}^{\sigma(n)} \otimes \cdots \otimes \hat{\mathbf{e}}^{\sigma(k+1)},$$

where the sum is over permutations on the complement of $\{a_1, \ldots, a_k\}$.

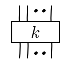
Proof. The function is computed by redrawing the diagram

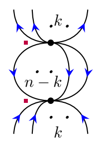
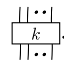


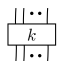
This diagram can be decomposed into two pieces to see how the input is transformed:

$$\begin{aligned}\hat{\mathbf{e}}_{a_1} \otimes \cdots \otimes \hat{\mathbf{e}}_{a_k} &\longmapsto \sum_{i_n, \ldots, i_{k+1}=1}^{n} \hat{\mathbf{e}}_{a_1} \otimes \cdots \otimes \hat{\mathbf{e}}_{a_k} \otimes \left(\hat{\mathbf{e}}_{i_{k+1}} \otimes \cdots \otimes \hat{\mathbf{e}}_{i_n}\right) \otimes \left(\hat{\mathbf{e}}^{i_n} \otimes \cdots \otimes \hat{\mathbf{e}}^{i_{k+1}}\right) \\ &\longmapsto \sum_{\sigma \in \Sigma_{\tilde{A}}} \mathsf{det}[\hat{\mathbf{e}}_{a_1} \ \cdots \ \hat{\mathbf{e}}_{a_k} \ \hat{\mathbf{e}}_{\sigma(k+1)} \ \cdots \ \hat{\mathbf{e}}_{\sigma(n)}]\hat{\mathbf{e}}^{\sigma(n)} \otimes \cdots \otimes \hat{\mathbf{e}}^{\sigma(k+1)} \\ &= \sum_{\sigma \in \Sigma_{\tilde{A}}} \mathrm{sgn}(\mathsf{Id}_A|\sigma)\hat{\mathbf{e}}^{\sigma(n)} \otimes \cdots \otimes \hat{\mathbf{e}}^{\sigma(k+1)}.\end{aligned}$$

In the second step, the summation may be restricted to permutations on $\tilde{A}$ since the determinant of any matrix with repeated columns is zero. $\square$

Proposition A.3 (Proposition 5.6)**.** *If $k > n$, then*  $= 0$. *Otherwise, for $0 \le k \le n$,*

 $= (-1)^{\lfloor \frac{n}{2} \rfloor}(n-k)!$ .

Proof. The fact that  $= 0$ for $k > n$ follows from Proposition 5.2.

For $k \le n$, the function is computed by applying Proposition 5.4 twice. The image of $\hat{\mathbf{e}}_{a_1} \otimes \cdots \otimes \hat{\mathbf{e}}_{a_k}$ is (using Lemma A.1):

$$\sum_{\tau \in \Sigma_A} \sum_{\sigma \in \Sigma_{\tilde{A}}} \mathrm{sgn}(\mathsf{Id}_A | \sigma) \mathrm{sgn}(\overleftarrow{\sigma} | \overleftarrow{\tau}) \hat{\mathbf{e}}_{\tau(a_1)} \otimes \cdots \otimes \hat{\mathbf{e}}_{\tau(a_k)}$$
$$= \sum_{\tau \in \Sigma_A} \sum_{\sigma \in \Sigma_{\tilde{A}}} (-1)^{\lfloor \frac{n}{2} \rfloor} \mathrm{sgn}(\tau) \hat{\mathbf{e}}_{\tau(a_1)} \otimes \cdots \otimes \hat{\mathbf{e}}_{\tau(a_k)}$$
$$= (-1)^{\lfloor \frac{n}{2} \rfloor}(n-k)! \sum_{\tau \in \Sigma_A} \mathrm{sgn}(\tau) \hat{\mathbf{e}}_{\tau(a_1)} \otimes \cdots \otimes \hat{\mathbf{e}}_{\tau(a_k)}. \qquad \square$$

Acknowledgments

I am grateful to Sean Lawton, Jenni Peterson, and Brian Winkel for comments on an early draft. Discussions with Steven Morse regarding this material were extremely helpful in developing the point of view taken here, and in particular for developing the ideas and notation used in the proofs of Propositions 5.4 and 5.6.

References

1. Doug Bullock, *Rings of $SL(2,C)$-characters and the Kauffman bracket skein module*, Comment. Math. Helv. **72** (1997), 521–542.
2. P. Cvitanovic, *Group theory*, Princeton University Press, 2008.
3. W. Fulton and J. Harris, *Representation theory: a first course*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, 1991.
4. Louis Kauffman, *Knots and physics*, World Scientific, 1991.
5. Louis Kauffman and Robin Thomas, *Temperley-Lieb algebras and the four-color theorem*, Combinatorica **23** (2003), 653–667.
6. Sean Lawton and Elisha Peterson, *Handbook of Teichmüller theory volume II*, ch. Spin networks and $SL(2,C)$-character varieties, unknown, 2008.
7. David C. Lay, *Linear algebra and its applications*, second ed., Addison Wesley, 1997.
8. I. B. Levinson, *Sum of Wigner coefficients and their graphical representation*, Proceed. Physical-Technical Inst. Acad. Sci. Lithuanian SSR **2** (1956), 17–30.
9. Seth A. Major, *A spin network primer*, Am. J. Phys. **67** (1999), 972.
10. Steven Morse, *Diagramming linear algebra: Some results and trivial proofs*, Undergraduate thesis, United States Military Academy, 2008.
11. Mara D. Neusel, *Invariant theory*, American Mathematical Society, 2007.
12. Roger Penrose, *Combinatorial mathematics and its applications*, ch. Applications of negative dimensional tensors, Academic Press, 1971.
13. ———, *Quantum theory and beyond*, ch. Angular Momentum: an Approach to Combinatorial Space-Time, pp. 151–180, Cambridge University Press, 1971.
14. Elisha Peterson, *Trace diagrams, representations, and low-dimensional topology*, Ph.D. thesis, University of Maryland, College Park, 2006.
15. Adam Sikora, *$SL(n)$-character varieties as spaces of graphs*, Trans. Amer. Math. Soc. **353** (2001), 2773–2804.
16. G. E. Stedman, *Diagram techniques in group theory*, Cambridge University Press, 1990.

Department of Mathematical Sciences, United States Military Academy, West Point, NY 10996

E-mail: **elisha.peterson@usma.edu**