LIE ALGEBRAS AND THE FOUR COLOR THEOREM

DROR BAR-NATAN

Appeared in Combinatorica **17-1** *(1997) 43–52.*

ABSTRACT. We present a statement about Lie algebras that is equivalent to the Four Color Theorem.

CONTENTS

1. INTRODUCTION

Let us start by recalling a well-known construction that associates to any finite dimensional metrized Lie algebra L a numerical-valued functional W_L defined on the set of all oriented trivalent graphs G (that is, trivalent graphs in which every vertex is endowed with a cyclic ordering of the edges emanating from it). This construction underlies the gauge-group dependence of gauge theories in general and of the Chern-Simons topological field theory in particular (see e.g. [B-N1, AS1, AS2]) and plays a prominent role in the theory of finite type (Vassiliev) invariants of knots ([B-N2, B-N3, B-N4]) and most likely also in the theory of finite type invariants of 3-manifolds ([O, GO, R]).

Fix a finite dimensional metrized Lie algebra L (that is, a finite dimensional Lie algebra with an *ad*-invariant symmetric non-degenerate bilinear form), choose some basis $\{L_a\}_{a=1}^{\dim L}$ of L, let $t_{ab} = \langle L_a, L_b\rangle$ be the metric tensor, let t^{ab} be the inverse matrix of t_{ab}, and let f_{abc} be the structure constants of L relative to $\{L_a\}$:

$$f_{abc} = \langle L_a, [L_b, L_c]\rangle.$$

Let G be some oriented trivalent graph. To define W_L, label all half-edges of G by symbols from the list $a, b, c, \ldots, a_1, b_1, \ldots$, and sum over $a, b, \ldots, a_1, \ldots \in \{1, \ldots, \dim L\}$ the product over the vertices of G of the structure constants "seen" around each vertex times the product over the edges of the $t^{\cdot\cdot}$'s seen on each edge. This definition is much better explained by an example, as in figure 1.

Date: This edition: October 22, 1999; First edition: August 18, 1995.
This paper is available electronically at `http://www.ma.huji.ac.il/~drorbn` and at `http://xxx.lanl.gov/abs/q-alg/9606016`.

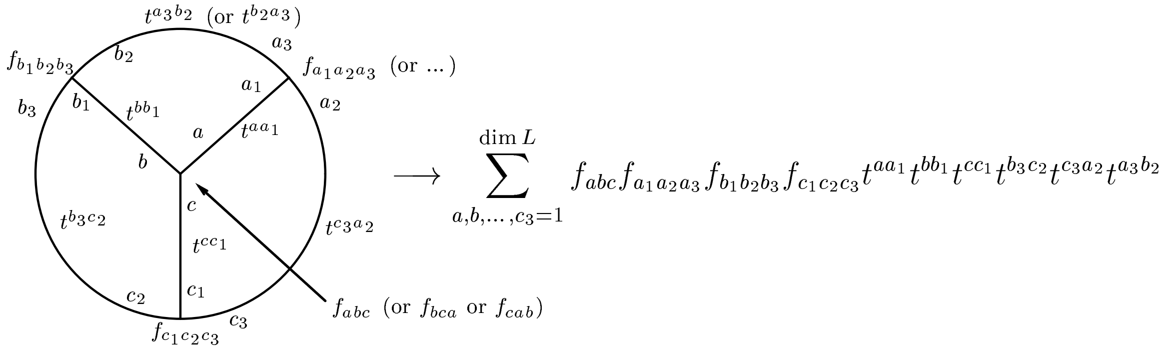


Figure 1. An example illustrating the construction of $W_L(G)$. Notice that when G is drawn in the plane, we assume counterclockwise orientation for all vertices (unless noted otherwise), and that the cyclic symmetry $f_{abc} = f_{bca} = f_{cab}$ of the structure constants and the symmetry $t^{ab} = t^{ba}$ of the inverse metric ensures that $W_L(G)$ is well defined.

By introducing an explicit change-of-basis matrix as in [B-N2] or by re-interpreting $W_L(G)$ in terms of abstract tensor calculus as in [B-N3], one can verify that $W_L(G)$ does not depend on the choice of the basis $\{L_a\}$. Typically one chooses a "nice" orthonormal (or almost orthonormal) basis $\{L_a\}$, so that most of the constants t^{ab} and f_{abc} vanish, thus greatly reducing the number of summands in the definition of $W_L(G)$.

Unless otherwise stated, whenever dealing with a Lie algebra of matrices, we will take the metric to be the matrix trace in the defining representation: $\langle L_a, L_b\rangle = \operatorname{tr}(L_aL_b)$.

Lemma-Definition 1.1. (proof in section 2) If a connected G has v vertices, then $W_{sl(N)}(G)$ is a polynomial in N of degree at most $\frac{v}{2}+2$ in N. Thus we can set $W^{\text{top}}_{sl(N)}(G)$ to be the coefficient of $N^{\frac{v}{2}+2}$ in $W_{sl(N)}(G)$.

The following statement sounds rather reasonable; it just says that if G is "$sl(2)$-trivial", then it is at least "$sl(N)$-degenerate". For us who grew up thinking that all that there is to learn about $sl(N)$ is already in $sl(2)$, this is not a big surprise:

Statement 1. *For a connected oriented trivalent graph G, $W_{sl(2)}(G) = 0$ implies $W^{top}_{sl(N)}(G) = 0$.*

Lie-theoretically, there is much to say about $sl(2)$ and $sl(N)$. There are representations of $sl(2)$ into $sl(N)$, there is an "almost decomposition" of $sl(N)$ into a product of $sl(2)$'s[1], and there are many other similarities. A-priori, the above statement sounds within reach. The purpose of this note is to explain why statement 1 is equivalent to the Four Color Theorem[2].

This equivalence follows from the following two propositions, proven in sections 2 and 3, respectively:

Proposition 1.2. *Let G be a connected oriented trivalent graph. If G is 2-connected, $|W^{top}_{sl(N)}(G)|$ is equal to the number of embeddings of G in an oriented sphere. Otherwise, $W^{top}_{sl(N)}(G) = 0$.*

[1] See [B-NG] for a similar context in which the different $sl(2)$'s "decouple".

[2] The Four Color Theorem was conjectured by Francis Guthrie in 1852 and proven by K. I. Appel and W. Haken [AH] in 1976. See also [SK].

Proposition 1.3. *(Penrose* [P]. *See also* [Ka1, Ka2, KS].*) If G is planar with v vertices and G^c is the map defined by its complement, then $|W_{sl(2)}(G)|$ is $2^{\frac{v}{2}-2}$ times the total number of colorings of G^c with four colors so that adjacent states are colored with different colors.*

Indeed, statement 1 is clearly equivalent to

$$|W^{\mathrm{top}}_{sl(N)}(G)| \neq 0 \quad \Rightarrow \quad |W_{sl(2)}(G)| \neq 0,$$

which by propositions 1.2 and 1.3 is the same as saying

$$\begin{pmatrix} G \text{ has a planar embedding} \\ \text{with } G^c \text{ a map} \end{pmatrix} \quad \Rightarrow \quad (G^c \text{ has a 4-coloring}).$$

Notice that if G is connected, G^c is a map (does not have states that border themselves) iff G is 2-connected.

Remark 1.4. We've chosen the formulation of statement 1 that we felt was the most appealing. With no change to the end result, one can replace $sl(N) = A_{N-1}$ by B_N, C_N, D_N, or $gl(N)$ and $sl(2)$ by $so(3)$ in the formulation of statement 1. In fact, in the proofs we actually work with $gl(N)$ and $so(3)$ rather than with $sl(N)$ and $sl(2)$.

1.1. **Acknowledgement.** I wish to thank G. Brinkmann, D. Goldberg, J. Goldman, M. Hutchings, L. Kauffman, D. P. Thurston and M. Wunderlich for many helpful conversations.

2. Understanding $W_{sl(N)}$

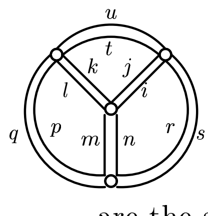
As Lie algebras, $gl(N)$ is just $sl(N)$ plus an Abelian factor. As Abelian Lie algebras have vanishing structure constants, $W_{sl(N)}(G) = W_{gl(N)}(G)$ for any oriented trivalent graph G. So let us concentrate on computing $W_{gl(N)}(G)$ for such G. For the basis of $gl(N)$, we pick the matrices $\{L_a\}_{a=1}^{N^2} = \{L_{ij}\}_{i,j=1}^{N}$, where L_{ij} is the matrix with 1 in the ij entry and 0 everywhere else. As the basis is indexed by a double index rather than by a single index, it is convenient to label every half-edge of G by two symbols from the list $i, j, \ldots, i_1, \ldots$ and double all the edges:

(1)

The metric $t_{ab} = t_{(ij)(kl)}$ of $gl(N)$ is given by $t_{(ij)(kl)} = \operatorname{tr} L_{ij}L_{kl} = \delta_{jk}\delta_{il}$, and its inverse is given by the same formula:

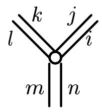
$$t^{(ij)(kl)} = \delta_{jk}\delta_{li}.$$

This formula means that in the summation defining $W_{gl(N)}(G)$ we can assume the equalities $j = k$ and $l = i$ along each edge as in (1). In other words, it is enough to label every doubled edge with just one pair of indices, getting an overall picture like

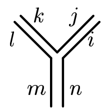


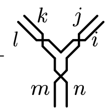
$$\longrightarrow \sum_{i,j,\ldots,u=1}^{N} f_{(ij)(kl)(mn)} f_{(ji)(rs)(ut)} f_{(lk)(tu)(qp)} f_{(nm)(pq)(sr)},$$

where $f_{(ij)(kl)(mn)}$ are the structure constants in our basis:



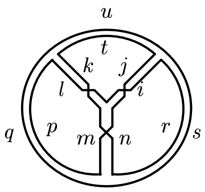
$$= f_{(ij)(kl)(mn)} = \langle L_{ij}, [L_{kl}, L_{mn}] \rangle = \operatorname{tr}(L_{ij}L_{kl}L_{mn}) - \operatorname{tr}(L_{mn}L_{kl}L_{ij})$$

$$= \delta_{jk}\delta_{lm}\delta_{ni} - \delta_{nk}\delta_{li}\delta_{jm} =$$ 

$-$ 

. (2)

In the last equation, indices connected by a line can be assumed to be equal in the summation defining $W_{gl(N)}(G)$. Once the edges and vertices of G are "thickened" as in (1) and (2), the summation over $i, j, \dots$ becomes the counting of the number of solutions of the equalities determined by the connected components of the thickened G. This number is simply N raised to the number of connected components:



$$\longrightarrow \sum_{i,j,\dots,u=1}^{N} \delta_{il}\delta_{lp}\cdots$$

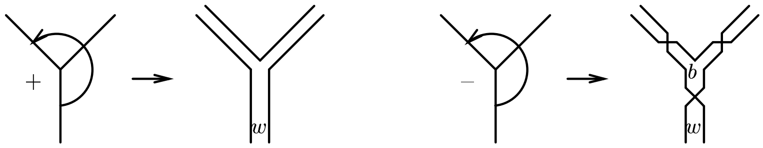
$$= \#\left\{1 \le i, j, \dots \le N : \begin{array}{c} i = l = p = m = j = t = k = n = r \\ u = q = s \end{array}\right\} = N^2.$$

Summarizing, we find the formula[3]

$$W_{gl(N)}(G) = \sum_{\text{markings } M \text{ of } G} \operatorname{sign}(M) N^{b(T_M)}, \tag{3}$$

where:

- A markings M of G is a marking of each vertex of G by a sign in $\{+, -\}$, and $\operatorname{sign}(M)$ is the product of these signs.
- The thickening T_M corresponding to a marking M is the oriented surface with boundary obtained from G as follows:
  - Replace the vertices marked by a "+" with "joints" and the vertices marked by a "−" with "twisted joints" as in (2).
  - Orient these surface pieces using the following black(b)/white(w) convention for the thickening of vertices:

    .

    In other words, "+"-vertices are embedded in the thickening of G so that they are seen as oriented counterclockwise from the white side of the thickening, while "-"-vertices are seen as oriented clockwise from the white side of the thickening.
  - Finally, connect the joints together along the edges of G by bands, in the only way consistent with the orientations of the joints.
- $b(T_M)$ is the number of boundary components of T_M.

If M is a marking of G, let S_M be the closed oriented surface obtained by gluing a disk into each boundary component of the thickening T_M. With χ denoting Euler characteristic and g denoting genus, we have

$$2 - 2g(S_M) = \chi(S_M) = \chi(T_M) + b(T_M) = \chi(G) + b(T_M).$$

[3]Compare with [B-N3, equation (36)]; for similar formulas in the cases of $so(N)$ and $sp(N)$, see [B-N3, equation (33)] and [B-N3, exercise 6.37].

Remembering that G is trivalent and thus $\chi(G) = -\frac{v}{2}$, we get

$$b(T_M) = -\chi(G) + 2 - 2g(S_M) = \frac{v}{2} + 2 - 2g(S_M).$$

Thus $b(T_M)$ is maximal when $g(S_M) = 0$ and in that case $b(T_M) = \frac{v}{2} + 2$. With (3), this proves lemma 1.1. Furthermore, calling a marking M *spherical* when S_M is a sphere, we get the formula

$$W^{\text{top}}_{sl(N)}(G) = W^{\text{top}}_{gl(N)}(G) = \sum_{\text{spherical markings } M \text{ of } G} \operatorname{sign}(M).$$

Proof of proposition 1.2. Let G be 2-connected, and consider the map

$$\Theta : \{\text{spherical markings } M \text{ of } G\} \longrightarrow \{\text{embeddings of } G \text{ in an oriented sphere}\}$$

defined by mapping M to the natural embedding of G in S_M. It is clear that Θ is a bijection. Indeed, if an embedding of G in an oriented sphere S^2 is given, one can reconstruct M by marking the vertices that are oriented counterclockwise within S_M (as seen from the white side of S_M) by a "+" sign, and marking all other vertices by a "−".

To conclude the proof, it is enough to show that for spherical markings, $\operatorname{sign}(M)$ is independent of the marking. Clearly, $\operatorname{sign}(M)$ is also equal to $(-1)^{v_-}$, where v_- is the number of vertices of G that are embedded clockwise in S_M by $\Theta(M)$ (with S_M viewed from its white side). By a theorem of H. Whitney [W1, W2][4], one can get from any spherical embedding of a trivalent 2-connected graph G to any other such embedding by a sequence of flips as in figure 2. Such flips do not change the parity of v_-, since the number of vertices that is flipped is even. □

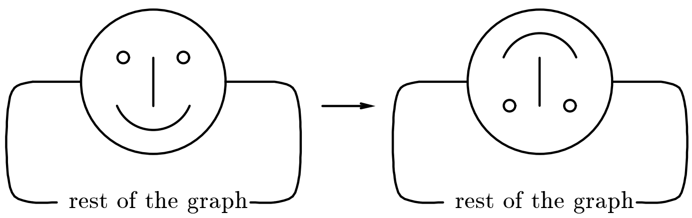


Figure 2. A flip takes a part of a graph that connects to the rest via only two edges, and flips it over.

3. Understanding $W_{sl(2)}$

Proposition 1.3 is due to Penrose [P] (see also [Ka1, Ka2, KS]). For completeness, we reproduce its proof in this section.

As Lie algebras, $sl(2)$ is isomorphic to $so(3)$, so let us work with $so(3)$ instead. The standard basis of $so(3)$ is given by the matrices

$$L_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad L_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad L_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

[4] Check [L, corollary 6] for a version closer to what we need, and remember that our graph is 2-connected and hence some of the moves in [L] are irrelevant for us.

Let us pick the scalar product $\langle\cdot,\cdot\rangle$ on $so(3)$ to be the one that makes this basis orthonormal, and let us denote the corresponding functional on graphs by $W_{\widetilde{so(3)}}$, with the "$\sim$" on top of the $so(3)$ to remind us that we are not using the standard matrix-trace scalar product.

One can easily verify that $\frac{1}{2}\langle\cdot,\cdot\rangle$ is the scalar product induced on $so(3)$ from matrix-trace in $sl(2)$. Thus, remembering that in the construction of W_L vertices scale with the scalar product and edges scale with its inverse, we find that

$$W_{sl(2)}(G) = \left(\frac{1}{2}\right)^{v-e} W_{\widetilde{so(3)}}(G) = 2^{\frac{v}{2}} W_{\widetilde{so(3)}}(G), \tag{4}$$

where G is an oriented trivalent graph with v vertices and e edges. Let us fix such a G once and for all, and let us assume that it is planar and that all the vertices of G are oriented counterclockwise in the plane. Flipping the orientation of any given vertex just reverses the sign of $W_{\widetilde{so(3)}}(G)$, and so the latter assumption does not limit the generality of our arguments.

Proof of proposition 1.3. With (4) in mind, proposition 1.3 clearly follows from the following two lemmas. □

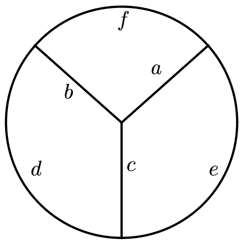
Lemma 3.1. *(Penrose* [P]. *See also* [Ka1, Ka2, KS].*) For a planar G as above, $|W_{\widetilde{so(3)}}(G)|$ is the number of colorings of the edges of G with three colors $\{1,2,3\}$, so that the edges emanating from any single vertex are of different colors.*

Lemma 3.2. *(Tait's theorem* [T]*) Edge-3-colorings as in the previous lemma are in a bijective correspondence with 4-colorings of the map G^c that fix the color of the "state at infinity".*

Proof of lemma 3.1. In the basis $\{L_a\}$, the structure constants of $so(3)$ are given by

$$f_{abc} = \epsilon_{abc} = \begin{cases} \mathrm{sign}(abc) & \text{if } abc \text{ is a permutation,} \\ 0 & \text{otherwise.} \end{cases}$$

Remembering also that $\{L_a\}$ is orthonormal by choice, the computation of $W_{\widetilde{so(3)}}(G)$ is given (on a simple example) by:



$$\longrightarrow \sum_{a,b,c,d,e,f=1}^{3} \epsilon_{abc}\epsilon_{aef}\epsilon_{bfd}\epsilon_{cde}.$$

The ϵ symbols force the indices coming into each vertex to be different, and hence clearly

$$W_{\widetilde{so(3)}}(G) = \sum_{\text{edge-3-colorings of } G} \prod \left(\text{a sign per vertex}\right), \tag{5}$$

where the sign at each vertex is the sign of the permutation of $\{1,2,3\}$ induced by an edge-3-coloring, as read counterclockwise around the vertex.

The only thing left to show is that the product of signs in (5) is independent of the edge-3-coloring. A clever way to do that, discovered by L. H. Kauffman, is to replace every edge colored by a "3" by a pair of edges colored "1" and "2" (in symbols, $1+2=3$). This defines two families of circles in the plane, labeled by "1" and by "2" (see figure 3). By lumping together the signs on each end of a "3" edge and taking the product over all of those edges, one sees that the overall sign depends only on the parity of the number of "3" edges (always $\frac{e}{3}$), and the $\mathbf{Z}/2\mathbf{Z}$ intersection number of the "1" family of circles with the "2" family of circles. By the Jordan curve theorem, the latter is always 0. $\square$

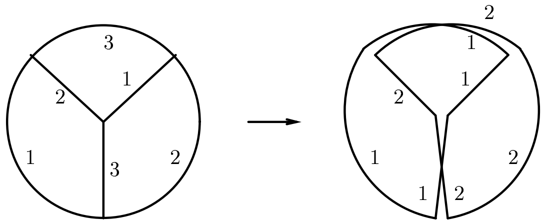


Figure 3. The two families of circles obtained by splitting every "3" edge.

Proof of lemma 3.2. This is a well known result (see e.g. [BM, theorem 9.12]), so let us only sketch the proof. Consider the group $H = \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z}$. Given any 4-coloring of G^c by elements of H, one may associate to it an edge-3-coloring of G by the non-zero elements of H, by coloring every edge by the difference of the colors in the two faces adjacent to it. One then verifies that this edge 3-coloring is well defined and that we get a bijection between the set of 4-colorings of G^c that color the state at infinity with 0 and the set of edge-3-coloring of G. $\square$

References

[AH] K. I. Appel and W. Haken, *Every planar map is four colorable*, Contemp. Math. **98**, Amer. Math. Soc., Providence 1989.

[AS1] S. Axelrod and I.M. Singer, *Chern-Simons perturbation theory*, Proc. XXth DGM Conference (New York, 1991) (S. Catto and A. Rocha, eds.), World Scientific, 1992, 3–45.

[AS2] ———, *Chern-Simons perturbation theory II*, Jour. Diff. Geom., **39** (1994) 173–213.

[B-N1] D. Bar-Natan, *Perturbative aspects of the Chern-Simons topological quantum field theory*, Ph.D. thesis, Princeton Univ., June 1991, Dep. of Mathematics.

[B-N2] ———, *Weights of Feynman diagrams and the Vassiliev knot invariants*, February 1991, preprint.

[B-N3] ———, *On the Vassiliev knot invariants*, Topology **34** (1995) 423–472.

[B-N4] ———, *Bibliography of Vassiliev Invariants*, `http://www.ma.huji.ac.il/~drorbn`.

[B-NG] ——— and S. Garoufalidis, *On the Melvin-Morton-Rozansky Conjecture*, Invent. Math. **125** (1996) 103–133.

[BM] J. A. Bondy and U. S. R. Murty, *Graph theory with applications*, Elsevier, New-York 1976.

[GO] S. Garoufalidis and T. Ohtsuki, *On finite type 3-manifold invariants III: manifold weight systems*, MIT and Tokyo Inst. of Tech. preprint, August 1995.

[Ka1] L. H. Kauffman, *Map coloring and the vector cross-product*, J. Comb. Theo. B **48** (1990) 145–154.

[Ka2] ———, *Map coloring, q-deformed spin networks and Turaev-Viro invariants for 3-manifolds*, Int. J. of Mod. Phys. B **6** (1992) 1765–1794.

[KS] ——— and H. Saleur, *An algebraic approach to the planar coloring problem*, Commun. Math. Phys. **152** (1993) 565–590.

[L] A. S. Lipson, *Link signature, Goeritz matrices, and polynomial invariants*, L'Enseignement Math. **36** (1990) 93–114.

[O] T. Ohtsuki, Finite type invariants of integral homology 3-spheres, Jour. of Knot Theory and its Ramifications **5(1)** (1996) 101–115.

[P] R. Penrose, *Applications of negative dimensional tensors*, Combinatorial mathematics and its applications (D. J. A. Welsh, ed.), Academic Press, San-Diego 1971, 221–244.

[R] L. Rozansky, *The trivial connection contribution to Witten's invariant and finite type invariants of rational homology spheres*, q-alg/9503011 preprint, March 1995.
[SK] T. L. Saaty and P. C. Kainen, *The four color problem, assaults and conquest*, McGraw-Hill, New-York 1977.
[T] P. G. Tait, *Remarks on colourings of maps*, Proc. Royal Soc. Edinburgh Ser. A., **10** (1880) 729.
[W1] H. Whitney, *A set of topological invariants for graphs*, Amer. J. Math. **55** (1933) 231–235.
[W2] ______, *On the classification of graphs*, Amer. J. Math. **55** (1933) 236-244.

Institute of Mathematics, Århus University, 8000 Århus C, Denmark
Current address: Institute of Mathematics, The Hebrew University, Giv'at-Ram, Jerusalem 91904, Israel
E-mail address: drorbn@math.huji.ac.il