

## Confluence theory for graphs

ADAM S SIKORA  
BRUCE W WESTBURY

We develop a theory of confluence of graphs. We describe an algorithm for proving that a given system of reduction rules for abstract graphs and graphs in surfaces is locally confluent. We apply this algorithm to show that each simple Lie algebra of rank at most 2, gives rise to a confluent system of reduction rules of graphs (via Kuperberg’s spiders) in an arbitrary surface. As a further consequence of this result, we find canonical bases of  $SU_3$ -skein modules of cylinders over orientable surfaces.

57M15, 57M27; 05C10, 16S15

### 1 Introduction

This paper is motivated by the following problem appearing in representation theory of Lie algebras and of quantum groups, in the study of moduli spaces, in knot theory, and in other areas of mathematics. We state it first for abstract graphs and, later, for graphs in manifolds.

Let  $R$  be a ring. An  $R$ -linear graph is a formal  $R$ -linear combination of graphs  $\Gamma = \sum_{i=1}^k r_i \Gamma_i$ , such that the graphs  $\Gamma_i$  have distinguished sets  $E_i$  of 1-valent vertices (called *external*) and there are specified bijections  $E_1 \simeq E_2 \simeq \cdots \simeq E_k$ . For any graph  $\Gamma'$  with a distinguished set of 1-valent external vertices  $E'$  in a bijection with  $E_1$  (and, consequently, in a bijection with  $E_i$  for all  $i$ ), let  $\langle \Gamma_i, \Gamma' \rangle$  denote the contraction of  $\Gamma_i$  and  $\Gamma'$  along their external vertices, respecting the specified bijections. In the process of the contraction these 1-valent vertices are removed and adjacent edges identified. Finally, let  $\langle \Gamma, \Gamma' \rangle = \sum_{i=1}^k r_i \langle \Gamma_i, \Gamma' \rangle$ .

Let  $\mathcal{G}$  be a set of graphs,  $\{\Gamma_i\}_{i \in I}$  be a set of  $R$ -linear graphs, and let  $\mathcal{R}(\Gamma_i, i \in I) \subset R\mathcal{G}$  be the submodule generated by contractions  $\langle \Gamma_i, \Gamma' \rangle$  for all  $i \in I$  and all graphs  $\Gamma'$  as above.

- (1) Is  $R\mathcal{G}/\mathcal{R}(\Gamma_i, i \in I)$  a free  $R$ -module? If so, then find an explicit basis of it.
- (2) Can a basis be given by taking all graphs in  $\mathcal{G}$  satisfying a certain “natural” property?

Examples appear in Section 8.

The topological version of this problem in dimension  $n$  involves topological graphs embedded in  $n$ -dimensional manifolds. An  $R$ -linear topological graph is  $\Gamma = \sum_i^k r_i \Gamma_i$ , such that  $\Gamma_1, \dots, \Gamma_k$  lie in a manifold  $M$  of dimension  $n$  and there is a finite set  $E \subset \partial M$ , such that  $\Gamma_i \cap \partial M = E$  for every  $i$  and this set is composed of 1-valent vertices of  $\Gamma_i$ . If  $\iota: M \rightarrow N$  is an embedding into a manifold of equal dimension and  $\Gamma'$  is a graph in  $\overline{N \setminus \iota(M)}$  such that points of  $E$  are 1-valent vertices of  $\Gamma'$  then  $\langle \Gamma_i, \Gamma' \rangle$  denotes the contraction of graphs  $\Gamma_i$  and  $\Gamma'$  along the vertices in  $E$ . As before,  $\langle \Gamma, \Gamma' \rangle = \sum_i r_i \langle \Gamma_i, \Gamma' \rangle$ .

Now, let  $\mathcal{G}$  be a set of topological graphs in  $N$  and  $\Gamma_i$  be an  $R$ -linear graph in  $M_i$ , for every  $i$  in some index set  $I$ . As before, let  $\mathcal{R}(\Gamma_i, i \in I) \subset R\mathcal{G}$  be the submodule generated by  $\langle \Gamma_i, \Gamma' \rangle$  for all  $i$ 's and all embeddings  $\iota: M_i \rightarrow N$  and all graphs  $\Gamma'$  as above. In this setting we ask again questions (1) and (2) above.

The flavor of these questions depends on the dimension of the manifold  $N$ :

**Dimension 2** Interesting examples come from Kuperberg's spider webs [36], which provide a convenient graphical description of representation theory of Lie algebras and associated quantum groups of rank  $\leq 2$ . (Spider webs for  $SO(7)$ , which has rank 3, were worked out by the second author in [65].) These are spaces of graphs in  $D^2$  considered modulo certain relations, of the type defined above. The classes of graphs considered and the relations between them depend on the Lie algebra in question. Because of their applications to quantum invariants, it is important to consider Kuperberg's webs in surfaces other than  $D^2$  as well. We answer (1) and (2) for these graphs in Sections 4–7. Our approach is based on theory of confluence of graphs developed in Section 2 and an algorithm for finding confluent reduction rules for graphs described in Section 3.6. As an application, we will find canonical bases of skein modules of skein modules of  $[0, 1]$ -bundles over surfaces for all simple Lie groups of rank 1 and 2. This reproves a theorem of Przytycki [52, Theorem 3.1] for the Kauffman bracket ( $SU_2$ ) skein modules of  $[0, 1]$ -bundles over surfaces and answers the question for  $SU_3$ -skein modules, cf Frohman–Zhong [18] and Sikora [62].

**Dimension 3** The three-dimensional version of this problem appears in knot theory, for example, in connection with Vassiliev invariants and skein modules. In both cases, (1) and (2) are open in general.

**Dimension  $> 3$**  In dimensions greater than 3 homotopic graphs are isotopic, and therefore the problem of describing  $R\mathcal{G}/\mathcal{R}(\Gamma_i, i \in I)$  can be reduced to purely algebraic form depending on  $\pi_1(N)$  only, since every  $\Gamma \subset N$  is determined by a labeling all cycles of  $\Gamma$  by conjugacy classes of  $\pi_1(N)$ . In particular, if  $\pi_1(N)$  is trivial then

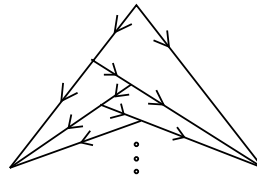


Figure 1

graphs in  $N$  can be thought as abstract graphs. For that reason, it is enough to consider questions (1) and (2) for abstract graphs only.

## 2 Confluence

We will approach the problems outlined in Introduction, by the method of confluence. To introduce it in its most abstract form, consider a set of objects  $V$  and a set of reduction rules,  $E$ , composed of pairs of elements of  $V$ , denoted by  $v \rightarrow v'$ . In other words,  $(V, E)$  is an arbitrary directed graph. A sequence of its vertices  $v_1 \rightarrow v_2 \rightarrow \dots \rightarrow v_n$  is called a *descending path* and its existence is denoted by  $v_1 \xrightarrow{*} v_n$ . We say that  $v_n$  is a *descendant* of  $v_1$ . We allow the empty path,  $v \xrightarrow{*} v$ , for any  $v$ . Consequently,  $v \xrightarrow{*} w$  is a relation on  $V$  which is reflexive and transitive but not necessarily symmetric. We write  $v \sim w$  if there is a finite path connecting  $v$  and  $w$ . (The edges of this path may have arbitrary directions.)

The reduction rules  $E$  are (*globally*) *confluent* if all  $v_1 \sim v_2$  have a common descendant, that is,  $w \in V$  such that  $v_1 \xrightarrow{*} w$  and  $v_2 \xrightarrow{*} w$ . Finally, rules  $E$  are *locally confluent* if for any  $v, w_1, w_2$  such that  $v \rightarrow w_1, v \rightarrow w_2$ , the elements  $w_1, w_2$  have a common descendant. Clearly, global confluence implies local confluence. However, the opposite implication fails, as shown in the graph in Figure 1, containing infinitely many vertices and edges.

Nonetheless, under certain mild conditions on reduction rules, local confluence implies global confluence. We say that reduction rules are *terminal* if all descending paths are finite.

**Diamond Lemma** (Newman [43, Theorem 3]) *If reduction rules are terminal then local confluence implies global confluence.*

An example of an application of the Diamond Lemma is the Jordan-Hölder theorem, which follows directly from this result. Other applications of the Diamond Lemma to ring theory and group theory are discussed by Bergman [4] and Sims [63]. Furthermore,

the Diamond Lemma and the notion of confluence is used in mathematical logic: in Church calculus (see Lalement [37] and Newman [43]), in lambda calculus (see Baader–Nipkow [3], Curien [12], Ohlebusch [45], Lalement [37] and Mitchel [41]), and in equational logic (see O’Donnell [44]). Additionally, it appears in computer science, in the theory of rewriting systems and in the study of graph grammars (see Ehrig [13; 14] and Nagl [42]).

## 2.1 Confluence of linear objects

For our applications we need a generalization of the notion of confluence to linear objects. For a ring  $R$ , an  $R$ -linear reduction rule on a set  $V$  is a pair  $S: v \rightarrow \sum_{i=1}^n r_i v_i$  where  $v, v_1, \dots, v_n \in V$  and  $r_1, \dots, r_n \in R$ . Denote the free  $R$ -module over  $V$  by  $RV$ . For  $X, Y \in RV$ , we write  $X \xrightarrow{S} Y$  if  $v$  appears with a non-zero coefficient in  $X$  and  $Y$  is obtained from  $X$  by replacing  $v$  by  $\sum_{i=1}^n r_i v_i$ . Finally, given a family of reduction rules,  $\{S_i\}_{i \in I}$ , we write  $X \xrightarrow{*} Y$  if there is a sequence of reduction rules leading from  $X$  to  $Y$ . Denote the  $R$ -submodule of  $RV$  generated by  $X - Y$  for all  $X \xrightarrow{S_i} Y$  by  $\mathcal{R}(S_i, i \in I)$  and write  $X_1 \sim X_2$  if  $X_1 - X_2 \in \mathcal{R}(S_i, i \in I)$ . As before, we say that rules  $\{S_i\}_{i \in I}$  are (globally) confluent if any  $X_1 \sim X_2$  have a common descendant, that is, there is  $Y$  such that  $X_1 \xrightarrow{*} Y$  and  $X_2 \xrightarrow{*} Y$ . Finally, rules  $\{S_i\}_{i \in I}$  are locally confluent on  $V$  (respectively, on  $RV$ ) if, for any  $X \in V$  (respectively, any  $X \in RV$ ) and any  $Y_1, Y_2 \in RV$  such that  $X \xrightarrow{S_i} Y_1$ ,  $X \xrightarrow{S_j} Y_2$ ,  $Y_1$  and  $Y_2$  have a common descendant. Clearly, global confluence implies local confluence on  $RV$ , and terminal local confluence on  $RV$  implies global confluence. However local confluence on  $V$  does not imply local confluence on  $RV$ ! For example, let  $V = \{v_1, v_2\}$  and let  $S_1: v_1 \rightarrow v_1 + 2v_2$ ,  $S_2: v_2 \rightarrow v_2 + 2v_1$ . Obviously,  $S_1, S_2$  are locally confluent on  $V$ , since for no  $v \in V$ ,  $S_1(v)$  and  $S_2(v)$  are simultaneously defined. However  $S_1$  and  $S_2$  are not confluent on  $\mathbb{R}V$ !

**Lemma 2.1**  $S_1(v_1 + \sqrt{2}v_2)$  and  $S_2(v_1 + \sqrt{2}v_2)$  have no common descendant in  $\mathbb{R}V$ .

**Proof** Notice that  $S_i$  sends  $a_1v_1 + a_2v_2$  to  $b_1v_1 + b_2v_2$ , where

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = M_i \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \text{and} \quad M_1 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

If  $S_1(v_1 + \sqrt{2}v_2)$  and  $S_2(v_1 + \sqrt{2}v_2)$  have a common descendant  $c_1v_1 + c_2v_2$  for some  $c_1, c_2 \in \mathbb{R}$ , then for certain products  $N_1, N_2$  of matrices  $M_1, M_2$ ,

$$N_1 M_1 \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix} = N_2 M_2 \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

Irrationality of  $\sqrt{2}$  implies  $N_1M_1 = N_2M_2$  as matrices in  $SL(2, \mathbb{Z})$ . However,  $M_1, M_2$  generate a free semigroup in  $SL(2, \mathbb{Z})$ . Therefore,  $N_1M_1 \neq N_2M_2$  for any  $N_1, N_2$ .  $\square$

Nonetheless, we have

**Theorem 2.2** (Linear Diamond Lemma) *Let  $\{V_j\}_{j \in J}$  be a family of subsets of  $V$ , such that  $J$  is a well ordered set and  $V_j \subset V_{j'}$  for  $j < j'$  and  $\bigcup_{j \in J} V_j = V$ . Let  $\deg(v) = \min \{j : v \in V_j\}$ . Consider a family of linear reduction rules on  $V$  such that each of them sends an element of  $V$  to a linear combination of elements of smaller degree. Then*

- (1) *these reduction rules are terminal,*
- (2) *if this family is locally confluent on  $V$  then it is also locally confluent on  $RV$ .*

Therefore, by the Diamond Lemma, such a family of reduction rules is globally confluent on  $RV$ .

**Proof** (1) Assume that there is an infinite chain  $X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow \dots$ . Let  $X_i = \sum_j^{n_i} c_{ij} v_{ij}$  and let  $d_{ik}$  denote the  $k$ th highest degree among degrees of  $v_{i,1}, \dots, v_{i,n_i}$ . Since  $d_{11} \geq d_{21} \geq d_{31} \geq \dots$ , the sequence stabilizes at certain place, which we denote by  $N_1$ . In other words  $d_{k,1} = d_{N_1,1}$ , for all  $k \geq N_1$ . Let  $e_1 = d_{N_1,1}$ . Similarly,  $d_{N_1,2} \geq d_{N_1+1,2} \geq d_{N_1+2,2} \geq \dots$ , stabilizes, let us say, at  $N_2$ th place. Let  $e_2 = d_{N_2,2}$ . By continuing this process, we construct  $e_1 \geq e_2 \geq e_3 \geq \dots$ . This sequence stabilizes at some point as well – let us say at  $s$ . Then for any  $k \geq N_s$ , the elements of  $V$  appearing in  $X_k = \sum_j^{n_i} c_{kj} v_{kj}$  have degrees  $e_1, \dots, e_s$  (each of them may be appearing many times). This, however, leads to contradiction since any reduction transformation replaces some  $v$  by a linear combination of elements of  $V$  of lower degree.

(2) Let  $S_1: v_1 \rightarrow \sum_{i=1}^{n_1} b_i w_i$  and  $S_2: v_2 \rightarrow \sum_{i=1}^{n_2} c_i z_i$ . Assume that  $\deg(v_1) < \deg(v_2)$ . We need to prove that for any  $X$ ,  $S_1(X), S_2(X)$  have a common descendant. Let  $X = a_1 v_1 + a_2 v_2 + X'$ , where  $X'$  is a linear combination of elements of  $V \setminus \{v_1, v_2\}$ . Since degrees of  $w_1, \dots, w_{n_1}$  are smaller than that of  $v_2$ , the elements  $w_1, \dots, w_{n_1}$  are different than  $v_2$ . If, additionally,  $z_1, \dots, z_{n_2} \neq v_1$  then

$$(1) \quad S_2 S_1(X) = a_1 \sum_{i=1}^{n_1} b_i w_i + a_2 \sum_{i=1}^{n_2} c_i z_i + X' = S_1 S_2(X)$$

is a common descendant of  $S_1(X)$  and  $S_2(X)$  and the proof is complete. Therefore, assume now that one of the  $z_i$ 's, say  $z_1$  is equal to  $v_1$ . If  $a_2 c_1 = 0$  then  $S_1 S_2(X) =$

$S_2S_1(X)$  again. However, this may not be the case if  $a_2c_1 \neq 0$ , since then

$$S_1S_2(X) = (a_1 + a_2c_1) \sum_{i=1}^{n_1} b_i w_i + a_2 \sum_{i=2}^{n_2} c_i z_i + X'$$

and  $S_2S_1(X)$  is as in (1). Now, however,  $S_1S_2(X) = S_1S_2S_1(X)$  is a common descendant of  $S_1(X)$  and  $S_2(X)$ .  $\square$

$X \in V$  is *irreducible* with respect to a given set of reduction rules if none of these rules applies to  $X$ . Denote the set of irreducible elements by  $V_{\text{irr}}$ . Note that if  $\{S_i\}_{i \in I}$  are terminal then  $RV/\mathcal{R}(S_i, i \in I)$  is spanned by  $V_{\text{irr}}$ . The opposite implication does not hold in general.

The combination of confluence and termination is a very strong property of reduction rules.

**Theorem 2.3** (1) *For any terminal rules  $\{S_i\}_{i \in I}$  for  $RV$  the following conditions are equivalent:*

- (a)  $S_i, i \in I$ , are locally confluent in  $RV$ ;
- (b)  $S_i, i \in I$ , are confluent in  $RV$ ;
- (c) For any  $x \in RV$  there is a unique element  $\psi(x) \in RV_{\text{irr}}$  such that  $x \xrightarrow{*} \psi(x)$ .

(2) *If any of the above conditions holds then  $\psi: RV \rightarrow RV_{\text{irr}}$  is an  $R$ -linear map which factors to an isomorphism*

$$\bar{\psi}: RV/\mathcal{R}(S_i, i \in I) \rightarrow RV_{\text{irr}}.$$

*Further,  $\psi$  is the identity on  $RV_{\text{irr}}$ , and consequently  $V_{\text{irr}}$  is a basis of  $RV/\mathcal{R}(S_i, i \in I)$ .*

**Proof** (a)  $\Rightarrow$  (b) This follows by the Diamond Lemma.

(b)  $\Rightarrow$  (c) Since the reduction rules are terminal, every  $x \in RV$  has a descendant  $y \in RV_{\text{irr}}$ . By confluence,  $y$  is unique – indeed, if  $x \xrightarrow{*} y' \neq y$  and  $y' \in RV_{\text{irr}}$  then  $y \sim y'$  but they have no common descendants, contradicting the confluence assumption.

(c)  $\Rightarrow$  (a) This is obvious.

(c)  $\Rightarrow$  (2) If  $x \xrightarrow{*} y$  then  $\psi(x) = \psi(y)$ . Since the relation  $\sim$  defined at the beginning of Section 2.1 is the smallest equivalence relation on  $RV$  generated by  $\xrightarrow{*}$ ,  $x \sim y$  implies that  $\psi(x) = \psi(y)$ . Therefore  $\psi$  factors to

$$\bar{\psi}: RV/\mathcal{R}(S_i, i \in I) = RV/\sim \rightarrow RV_{\text{irr}}.$$

If we denote the “obvious” map  $RV_{\text{irr}} \rightarrow RV \rightarrow RV/\mathcal{R}(S_i, i \in I)$  by  $\iota$  then clearly both  $\iota\psi$  and  $\psi\iota$  are identities on their respective domains. Therefore  $\bar{\psi}$  is a bijection and an  $R$ -linear map. Finally,  $\psi$  is also  $R$ -linear, since it is a composition of linear maps

$$RV \rightarrow RV/\mathcal{R}(S_i, i \in I) \xrightarrow{\bar{\psi}} RV_{\text{irr}}. \quad \square$$

Only a few interesting terminal and confluent reduction systems on sets of graphs are known. Most of them appear in the context of representation theory of Lie algebras of rank  $\leq 2$  and of associated quantum groups, cf Sections 4–7. See Section 8 for other examples.

### 3 Graphs

#### 3.1 Abstract graphs

In a most general setting, a *labeled graph* is  $\Gamma = (\mathcal{V}, \mathcal{E}, t, \tau, \Lambda, \lambda, \nu)$ , where  $\mathcal{V}$  is a vertex set,  $\mathcal{E}$  is the set of *edge directions*,  $t: \mathcal{E} \rightarrow \mathcal{V}$  is the *tail map*,  $\tau: \mathcal{E} \rightarrow \mathcal{E}$  is the *change of direction involution* which is fixed-point free.  $\Lambda$  is a set of labels and  $\lambda: \mathcal{E} \rightarrow \Lambda$  is a labeling function.  $\nu: \Lambda \rightarrow \Lambda$  is an involution such that  $\nu\lambda = \lambda\tau: \mathcal{E} \rightarrow \Lambda$ . The function  $t\tau: \mathcal{E} \rightarrow \mathcal{V}$  is called the *head map*.

An edge is a two-element set  $\{e, \tau(e)\}$ . The *valency* of  $v \in \mathcal{V}$  is the number of edge directions  $e$  such that  $t(e) = v$ . As mentioned in Introduction, we will sometimes specify a set of 1-valent vertices  $V_{\text{ext}}(\Gamma) \subset V(\Gamma)$ , called *external vertices*, and consider it as part of graph structure of  $\Gamma$ . The remaining vertices,  $V_{\text{int}}(\Gamma) = V(\Gamma) \setminus V_{\text{ext}}(\Gamma)$ , are *internal*.

Most definitions of graphs can be deduced from this one. For example, a *partially directed graph* is  $\Gamma = (\mathcal{V}, \mathcal{E}, t, \tau, \emptyset, \Lambda, \lambda, \nu)$ , such that  $\Lambda = \{\pm 1, 0\}$  and  $\nu(x) = -x$ . An edge  $\{e, \tau(e)\}$  with  $\lambda(e) = 0$  is *undirected*. Otherwise, its *direction* is either  $e$  or  $\tau(e)$  depending on whether  $\lambda(e) = 1$  or  $-1$ .

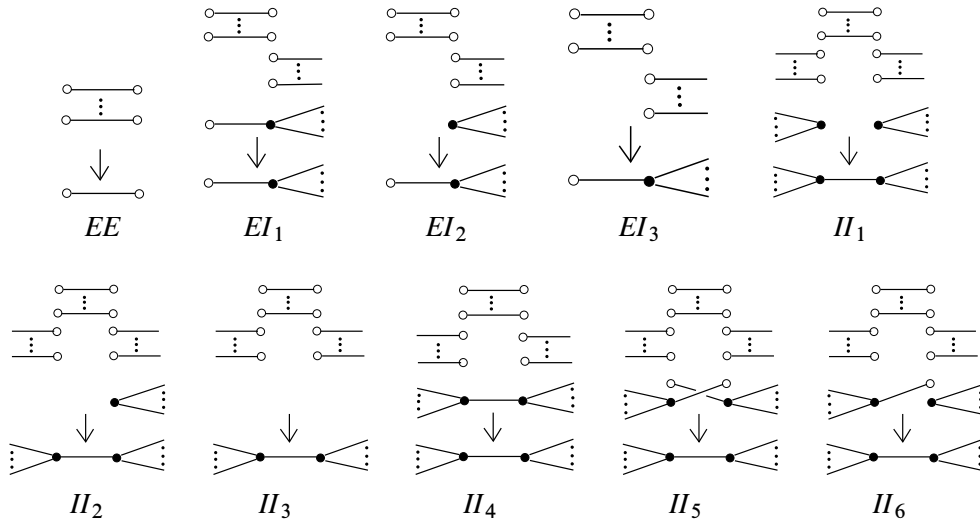
The reason for using edge directions, instead of edges, is that in representation theory one considers graphs whose edges are labeled by representations and have no canonical orientation. If an edge direction is labeled by a representation  $V$  then the opposite direction is labeled by the dual of  $V$ .

An embedding of  $\Gamma_1$  into  $\Gamma_2$  is

- (1) a map  $f: V_1 \rightarrow V_2$  which is an embedding of internal vertices of  $\Gamma_1$  into internal vertices of  $\Gamma_2$ ,

- (2) a map  $g: \mathcal{E}_1 \hookrightarrow \mathcal{E}_2$ , such that  $t_2g = ft_1$ ,  $\tau_2g = g\tau_1$ , and  $g$  restricted to  $\{e \in \mathcal{E}_1: t(e) \text{ is an internal vertex}\}$  is an embedding.
- (3) an embedding  $h: \Lambda_1 \hookrightarrow \Lambda_2$  such that  $\lambda_2h = h\lambda_1$  and  $v_2h = hv_1$ .

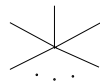
For any embedding  $f: \Gamma \hookrightarrow \Gamma'$  and  $e \in E(\Gamma')$ , neighborhood of  $f^{-1}(e)$  has one of the following forms:<sup>1</sup>



Above, black dots denote internal vertices and white dots the external ones. Triple dots denote several parallel copies (possibly zero). For the purpose of this classification we ignore edge directions.

### 3.2 Graphs in manifolds

Throughout the paper all manifolds are smooth. A graph in a manifold  $M$ , or *manifold graph*, is a subspace  $\Gamma \subset M$  which looks locally like 1–dimensional submanifold of  $M$  (possibly with boundary) except for *internal vertices*:



For simplicity, we do not allow 2–valent vertices.

The points of  $\Gamma \cap \partial M = V_{\text{ext}}(\Gamma)$  are called *external vertices*. Manifold graphs are considered up to isotopy of  $M$  fixing  $\partial M$ . We denote the set of all vertices of  $\Gamma$

<sup>1</sup>To be precise, one considers the topological realization of  $\Gamma$  and the topological neighborhood of  $f^{-1}(e)$ .



by  $V(\Gamma)$ . Note that  $\Gamma \setminus V(\Gamma)$  is composed of open intervals and circles (also called loops).

A manifold graph  $\Gamma \subset M$  is *labeled* if there is specified a set  $\Lambda$  with an involution  $\tau: \Lambda \rightarrow \Lambda$  and a labeling function

$$\lambda: \{\text{orientations of connected components of } \Gamma \setminus V(\Gamma)\} \rightarrow \Lambda.$$

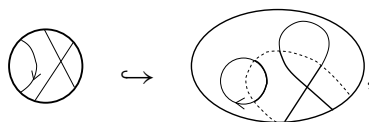
We require that if  $o, \bar{o}$  are opposite orientations of the same edge or circle in  $\Gamma \setminus V(\Gamma)$  then  $\lambda(\bar{o}) = \tau(\lambda(o))$ . A graph labeled by  $\Lambda = \{0, \pm 1\}$ , with  $\tau(x) = -x$ , is called *partially oriented*. An edge or circle  $e$  of  $\Gamma$  is unoriented if  $\lambda(e) = 0$  and oriented otherwise. Its orientation is the one labeled by 1.

Note that if  $M$  is connected, simply-connected, has connected boundary, and  $\dim M \geq 4$  then each graph in  $M$  can be thought as a geometric realization of an abstract graph. Such abstract graph is unique up to an insertion or deletion of 2-valent internal vertices into edges or from edges.

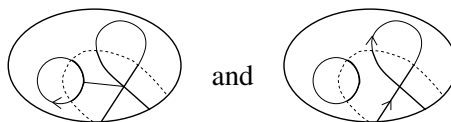
An *embedding* of manifold graph  $\Gamma_1 \subset M_1$  into  $\Gamma_2 \subset M_2$  is an embedding  $f: M_1 \hookrightarrow M_2$  of manifolds of equal dimensions, which embeds a certain representative  $\bar{\Gamma}_1$  of the isotopy class of  $\Gamma_1 \subset M_1$  into a certain representative  $\bar{\Gamma}_2$  of the isotopy class of  $\Gamma_2 \subset M_2$  such that  $f$  restricted to  $\Gamma_1 \setminus V_{\text{ext}}(\Gamma_1)$  is an open map into  $\bar{\Gamma}_2$ . We identify isotopic embeddings. This definition implies that edges of  $\Gamma_1$  are mapped either into edges or circles of  $\Gamma_2$ .

If  $\Gamma_1 \subset M_1, \Gamma_2 \subset M_2$  are labeled manifold graphs, then an embedding of  $\Gamma_1$  into  $\Gamma_2$  consists of a map  $f: M_1 \rightarrow M_2$  as above together with an embedding of the set of labels  $\iota: \Lambda_1 \hookrightarrow \Lambda_2$  such that  $f$  maps every edge or circle with some orientation,  $e_1$ , of  $\Gamma_1$  into an edge or circle of  $\Gamma_2$ , denoted by  $e_2$  with coinciding orientation such that  $\lambda_2(e_2) = \iota\lambda_1(e_1)$ .

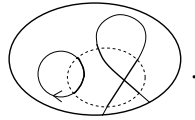
**Example 3.1** A graph embedding



and two non-embeddings



The above graph embedding is isotopic and, hence, identified with the embedding



The theory presented in this paper comes in two flavors: oriented and unoriented. In the first case all manifolds are oriented and all embeddings preserve orientations of manifolds. In the latter case, orientations of manifolds do not play any role. In both cases, labelings of edge and circle orientations are preserved. We will stress the difference between oriented and unoriented case whenever necessary, for example in Section 5.

### 3.3 Linear graphs

Let  $R$  be a ring. An  $R$ -linear graph is a formal  $R$ -linear combination of graphs  $\Gamma = \sum_{i=1}^k r_i \Gamma_i$ , together with specified bijections  $V_{\text{ext}}(\Gamma_1) \simeq V_{\text{ext}}(\Gamma_2) \simeq \dots \simeq V_{\text{ext}}(\Gamma_k)$  such that the corresponding external edge directions have identical labels. (Since each external vertex is 1-valent, its adjacent *external edge direction* is well defined.)

An  $R$ -linear manifold graph in  $M$  is a formal linear combination  $\Gamma = \sum_{i=1}^n r_i \Gamma_i$  of graphs in  $M$  such that their external vertices coincide and the outward orientations of the corresponding external edges have identical labels. These graphs are considered up to isotopy of  $M$  fixing  $\partial M$ . For example,

$$T = \text{[Diagram 1]} - \text{[Diagram 2]}$$

is a non-zero linear graph in  $D^2$ .

### 3.4 Reduction rules on graphs

We are going to apply the theory of confluence to the problems stated in Introduction, with a particular focus on graphs arising as Kuperberg’s spiders webs. Certain versions of this method were used implicitly already by Jaeger [28], Kuperberg [35; 36] and Yetter [66]. Nonetheless, to our knowledge the subtle difference between local confluence on  $V$  and on  $RV$  discussed in Section 2.1 has never been observed.

For our purposes, the set of objects,  $V$ , considered in Section 2 is either a set of abstract graphs or a set of graphs in a given manifold  $M$ .

In the first case, for a given ring  $R$ , a *graph reduction rule* is a pair denoted by  $T_0 \xrightarrow{S} T$ , where  $T_0$  is a graph,  $T = \sum_{i=1}^k r_i T_i$  is an  $R$ -linear graph and the external vertices

of  $T_0$  and  $T_i$ 's are identified via a bijection such that the corresponding external edge directions have identical labels. Any graph reduction rule  $T_0 \xrightarrow{S} T$  defines reductions of graphs  $\Gamma$  as follows: For any embedding  $\Gamma_0 \hookrightarrow \Gamma$  we obtain new graphs  $\Gamma_i$  by replacing  $T_0$  in  $\Gamma$  by  $T_i$ . We say that reduction  $\Gamma_0 \rightarrow \sum_{i=1}^k r_i \Gamma_i$  is induced by  $T_0 \xrightarrow{S} T$  and we denote that fact by putting  $S$  above the arrow,  $\Gamma \xrightarrow{S} \sum_{i=1}^k r_i \Gamma_i$ . Therefore, unlike in Section 2, we use one symbol (here,  $S$ ) to denote many reduction rules arising from  $T_0 \xrightarrow{S} T$ .

Similarly, a *graph reduction rule* for manifold graphs is a pair  $T_0 \xrightarrow{S} T$ , where  $T_0$  is a graph in some manifold  $M_0$  and  $T = \sum_{i=1}^k r_i T_i$  is an  $R$ -linear graph in the same manifold such that the external vertices of  $T_0$  and  $T_i$ 's coincide and the corresponding external edge orientations have identical labels. For any embedding  $M_0 \hookrightarrow M$  and a graph  $\Gamma \subset M$  such that  $\Gamma \cap M_0 = T_0$ , we obtain new graphs  $\Gamma_i \subset M$  by replacing  $T_0$  in  $\Gamma$  by  $T_i$ . We say that the reduction  $\Gamma_0 \rightarrow \sum_{i=1}^k r_i \Gamma_i$  is induced by  $T_0 \xrightarrow{S} T$  and we denote that fact by  $\Gamma \xrightarrow{S} \sum_{i=1}^k r_i \Gamma_i$ .

By analogy to the notation in Section 2, we use  $\mathcal{R}(S_i, i \in I) \subset RV$  to denote the submodule generated by all linear graphs  $\Gamma - \sum_{i=1}^k r_i \Gamma_i$  coming from graph reductions  $\Gamma \xrightarrow{S_i} \sum_{i=1}^k r_i \Gamma_i$ , for  $i \in I$ .

### 3.5 Proving confluence of reduction rules of abstract graphs

An *overlap* of graphs  $\Gamma_1$  and  $\Gamma_2$  is a graph  $\Gamma$  and pair  $(\Gamma_1 \hookrightarrow \Gamma, \Gamma_2 \hookrightarrow \Gamma)$  of graph embeddings. If  $\{T_{i0} \xrightarrow{S_i} \sum_k r_{ik} T_{ik}\}_{i \in I}$  is a set of reduction rules of abstract graphs with coefficients in  $R$ , then each overlap  $O = (\iota_1: T_{i0} \hookrightarrow \Gamma, \iota_2: T_{j0} \hookrightarrow \Gamma)$  leads to two different reductions of  $\Gamma$ . We say that reduction rules  $\{S_i\}_{i \in I}$  are *locally confluent on*  $O = (\Gamma_1 \hookrightarrow \Gamma, \Gamma_2 \hookrightarrow \Gamma)$  if, for any  $i, j$  such that  $\Gamma_{i0} = \Gamma_1, \Gamma_{j0} = \Gamma_2$ , the two reductions  $\Gamma \xrightarrow{S_i} \sum_k r_{ik} \Gamma_k$  and  $\Gamma \xrightarrow{S_j} \sum_k r_{jk} \Gamma'_k$ , arising from this overlap have a common descendant.

Note that  $\{S_i\}_{i \in I}$  are locally confluent on  $V$  if and only if they are locally confluent on all overlaps of graphs  $T_{i0}, i \in I$ .

We say that  $O = (\iota_1: \Gamma_1 \hookrightarrow \Gamma, \iota_2: \Gamma_2 \hookrightarrow \Gamma)$  factors through  $O' = (\iota'_1: \Gamma_1 \hookrightarrow \Gamma', \iota'_2: \Gamma_2 \hookrightarrow \Gamma')$  if there is an embedding  $f: \Gamma' \hookrightarrow \Gamma$ , such that  $\iota_1 = f \iota'_1, \iota_2 = f \iota'_2$ . If reduction rules  $\{S_i\}_{i \in I}$  are locally confluent on  $O$  then they are locally confluent on all overlaps which factor through  $O$ . An overlap with no factorizations other than the identity is *irreducible*.

We are going to show that factorizations of the type

$$(\iota'_1: \Gamma_1 \hookrightarrow \Gamma', \iota'_2: \Gamma_2 \hookrightarrow \Gamma') \xrightarrow{f} (\iota_1: \Gamma_1 \hookrightarrow \Gamma, \iota_2: \Gamma_2 \hookrightarrow \Gamma)$$

reduce every overlap to an irreducible one:

- (1) If  $V(\Gamma_1) \cup V(\Gamma_2)$  is a proper subset of  $V(\Gamma)$  then let  $\Gamma'$  be a graph obtained from  $\Gamma$  by removing vertices in  $V(\Gamma) \setminus (V(\Gamma_1) \cup V(\Gamma_2))$ . Let  $\iota'_1 = \iota_1$ ,  $\iota'_2 = \iota_1$ , and let  $f$  be the obvious embedding.
- (2) If  $V_{\text{int}}(\Gamma_1) \cup V_{\text{int}}(\Gamma_2)$  is a proper subset<sup>2</sup> of  $V_{\text{int}}(\Gamma)$ , then let  $\Gamma'$  be a graph obtained from  $\Gamma$  by changing the internal vertices in  $V_{\text{int}}(\Gamma) \setminus (V_{\text{int}}(\Gamma_1) \cup V_{\text{int}}(\Gamma_2))$  to external ones. Let  $f$  be the obvious embedding.
- (3) If  $E(\Gamma_1) \cup E(\Gamma_2) \subset E(\Gamma)$  is a proper subset, then let  $\Gamma'$  be  $\Gamma$  with the edge directions in  $E(\Gamma) \setminus (E(\Gamma_1) \cup E(\Gamma_2))$  removed.  $f$  is the obvious embedding.
- (4) Let  $E_{\text{ext}}(\Gamma)$  denote  $\{e \in E(\Gamma) : t(e), h(e) \in V_{\text{ext}}(\Gamma)\}$ . If  $e_1 \in E_{\text{ext}}(\Gamma_i)$ ,  $\iota_i(e_1) = \iota_j(e_2)$ ,  $e_1 \neq e_2$ , for some  $i, j \in \{1, 2\}$ , then let  $\Gamma'$  be  $\Gamma$  with two extra external vertices  $v_1, v_2$  and two new edge directions  $e', \tau(e')$  forming an edge connecting  $v_1$  and  $v_2$ . Let  $\iota'_1, \iota'_2$  coincide with  $\iota_1, \iota_2$ , except for  $\iota'_i$  sending  $e_1, \tau(e_1)$  to  $e', \tau(e')$  and sending  $t(e_1), h(e_1)$  to  $v_1, v_2$ . Let  $f(v_1) = t(e)$ ,  $f(v_2) = h(e)$ ,  $f(e') = e$ , and let  $f$  be the identity on the remaining vertices and edges.
- (5) If  $e_1 \in E_{\text{ext}}(\Gamma_i)$  and there is no edge  $e_2$  as in (4) but  $\iota_i(t(e_1)) = \iota_j(v)$ ,  $t(e_1) \neq v$ , for some  $i, j \in \{1, 2\}$ , then let  $\Gamma'$  be  $\Gamma \setminus \{\iota_i(e_1), \tau(\iota_i(e_1))\}$  with an extra external vertex  $w$  and two new edge directions  $e', \tau(e')$  forming an edge connecting  $w$  and  $\iota_i(h(e_1))$ . Let  $\iota'_1, \iota'_2$  coincide with  $\iota_1, \iota_2$ , except for  $\iota'_i$  sending  $e_1, \tau(e_1)$  to  $e', \tau(e')$  and sending  $t(e_1)$  to  $w$ . Let  $f(w) = \iota_i(t(e_1))$ ,  $f(e') = \iota_i(e_1)$ ,  $f(\tau(e')) = \iota_i(\tau(e_1))$ , and let  $f$  be the identity on the remaining vertices and edge directions.
- (6) If  $\iota_1^{-1}(e)$  and  $\iota_2^{-1}(e)$  are of type  $II_5$  (as defined at the end of Section 3.1) then consider factorization in Figure 2. (For simplicity the edge directions are ignored.)  
Analogously, if  $\iota_1^{-1}(e)$  and  $\iota_2^{-1}(e)$  are (in some order) of types  $(II_1, II_5)$ ,  $(II_1, II_6)$ ,  $(II_2, II_5)$ ,  $(II_2, II_6)$ ,  $(II_3, II_5)$ ,  $(II_3, II_6)$ ,  $(II_5, II_6)$  or  $(II_6, II_6)$  then we perform similar factorizations.

**Theorem 3.2** An overlap  $O = (\iota_1: \Gamma_1 \hookrightarrow \Gamma, \iota_2: \Gamma_2 \hookrightarrow \Gamma)$  is irreducible iff

- (1)  $V(\Gamma) = V(\Gamma_1) \cup V(\Gamma_2)$ .
- (2)  $V_{\text{int}}(\Gamma) = V_{\text{int}}(\Gamma_1) \cup V_{\text{int}}(\Gamma_2)$ .
- (3)  $E(\Gamma) = E(\Gamma_1) \cup E(\Gamma_2)$
- (4) If  $e_1 \in E_{\text{ext}}(\Gamma_i)$  and  $\iota_i(e_1) = \iota_j(e_2)$  then  $j = i$  and  $e_1 = e_2$ .
- (5) If  $e \in E_{\text{ext}}(\Gamma_i)$ ,  $\iota_i(t(e)) = \iota_j(v)$ , then  $i = j$  and  $t(e) = v$ .

<sup>2</sup>Since graph embeddings send internal vertices to internal vertices,  $V_{\text{int}}(\Gamma_1) \cup V_{\text{int}}(\Gamma_2) \subset V_{\text{int}}(\Gamma)$ .

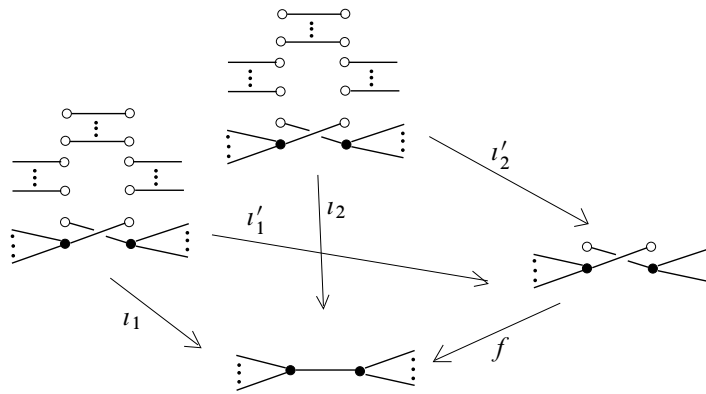


Figure 2

- (6)  $\Gamma$  has no edges of types  $(II_1, II_5)$ ,  $(II_1, II_6)$ ,  $(II_2, II_5)$ ,  $(II_2, II_6)$ ,  $(II_3, II_5)$ ,  $(II_3, II_6)$ ,  $(II_5, II_5)$ ,  $(II_5, II_6)$ ,  $(II_6, II_6)$ .

**Proof**  $\Rightarrow$  If one of the conditions does not hold then  $\mathcal{O}$  admits one of the factorizations described above.

$\Leftarrow$  Suppose that  $\mathcal{O}$  satisfies (1)–(6) above and

$$f: (i'_1: \Gamma_1 \hookrightarrow \Gamma', i'_2: \Gamma_2 \hookrightarrow \Gamma') \rightarrow (i_1: \Gamma_1 \hookrightarrow \Gamma, i_2: \Gamma_2 \hookrightarrow \Gamma) = \mathcal{O}$$

is a factorization of  $\mathcal{O}$ .

**Lemma 3.3**  $f: V(\Gamma') \rightarrow V(\Gamma)$  is a bijection.

**Proof**  $f(V(\Gamma')) \supset f i'_1 V(\Gamma_1) \cup f i'_2 V(\Gamma_2) = i_1 V(\Gamma_1) \cup i_2 V(\Gamma_2)$  is by (1) equal to  $V(\Gamma)$ . Therefore,  $f: V(\Gamma') \rightarrow V(\Gamma)$  is onto.

Suppose  $f(v_1) = f(v_2) = v$  for  $v_1 \neq v_2$ ,  $v_1, v_2 \in V(\Gamma')$ . By definition of graph embedding,  $f$  is 1–1 on  $V_{\text{int}}(\Gamma')$ . Hence, at least one of the vertices  $v_1, v_2$  is external, say  $v_1$ . Consider two cases:

**Case 1**  $v$  is external Then  $v_2$  is external as well. Let  $e_1, e_2$  be edge directions with tails  $v_1, v_2$ . Since  $v$  is 1–valent,  $f(e_1) = f(e_2) := e \in E(\Gamma)$ . Since  $f(h(e_1)) = f(h(e_2)) = h(e)$ , at least one of the vertices  $h(e_1), h(e_2)$  is external, contradicting (4).

**Case 2**  $v$  is internal By (2), we can choose  $v_2$  to be an internal vertex. Since  $v_1$  has valency one, there is a unique vertex  $w$  in  $\Gamma'$  connected with  $v$  by an edge  $e'$ . By (5),  $w$  is internal. Let  $e = f(e')$ . Then the preimage,  $f^{-1}$ , of the neighborhood of  $e$  must be of type  $II_5$  or  $II_6$ . Since  $i_1, i_2$  satisfy (1)–(4) and factorize through  $f$ , they must be of one of the types listed in (6). Contradiction.  $\square$

**Corollary 3.4**  $f: V_{\text{int}}(\Gamma') \rightarrow V_{\text{int}}(\Gamma)$  and  $f: V_{\text{ext}}(\Gamma') \rightarrow V_{\text{ext}}(\Gamma)$  are bijections.

**Proof** Since  $f$  is 1–1 it is enough to prove that (a)  $f(V_{\text{int}}(\Gamma')) = V_{\text{int}}(\Gamma)$  and (b)  $f(V_{\text{ext}}(\Gamma')) = V_{\text{ext}}(\Gamma)$ .

(a) Since  $f$  is a graph embedding,  $f(V_{\text{int}}(\Gamma')) \subset V_{\text{int}}(\Gamma)$ . By (2),  $\iota_1(V_{\text{int}}(\Gamma_1)) \cup \iota_2(V_{\text{int}}(\Gamma_2)) = V_{\text{int}}(\Gamma)$ , and since  $\iota_1, \iota_2$  factor through  $f$ ,  $f(V_{\text{int}}(\Gamma')) = V_{\text{int}}(\Gamma)$ .

(b) By the previous lemma and by (a),  $|V(\Gamma')| = |V(\Gamma)|$ ,  $|V_{\text{int}}(\Gamma')| = |V_{\text{int}}(\Gamma)|$ . Hence,  $|V_{\text{ext}}(\Gamma')| = |V_{\text{ext}}(\Gamma)|$  and the statement follows from the fact that  $f$  is 1–1.  $\square$

**Proposition 3.5**  $f: E(\Gamma') \rightarrow E(\Gamma)$  is a bijection.

**Proof** By (3),  $f$  is onto. Suppose that  $f(e_1) = f(e_2) = e$ ,  $e_1 \neq e_2$ . Since  $f$  is a bijection on vertices,  $t(e_1) = t(e_2)$ ,  $h(e_1) = h(e_2)$ . Since  $t(e_1), h(e_2)$  are at least 2 valent, they are internal and, consequently,  $f$  maps two internal edge directions to a single edge direction and, therefore, it is not a graph embedding.  $\square$

Therefore  $f$  is the identity and the proof of Theorem 3.2 is completed.  $\square$

Since each overlap of  $\Gamma_1$  and  $\Gamma_2$  satisfying Theorem 3.2(1)–(3) is obtained as a quotient of the disjoint union of  $\Gamma_1$  and  $\Gamma_2$ , the number of such overlaps is finite.

**Corollary 3.6** Any two abstract graphs have a finite number of irreducible overlaps only.

Consider an overlap of  $\Gamma_1$  and  $\Gamma_2$ . By applying factorizations of types (1)–(5), we obtain an overlap satisfying conditions (1)–(5) of Theorem 3.2. Observe, that if an overlap  $O$  satisfies these conditions then for every factorization  $f: O' \rightarrow O$  of type (6),  $O'$  satisfies (1)–(5) as well. Furthermore, observe that for each factorization

$$f: (\iota'_1: \Gamma_1 \hookrightarrow \Gamma', \iota'_2: \Gamma_2 \hookrightarrow \Gamma') \rightarrow (\iota_1: \Gamma_1 \hookrightarrow \Gamma, \iota_2: \Gamma_2 \hookrightarrow \Gamma) = O$$

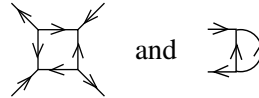
of type (6), either (a) the number of connected components of  $\Gamma'$  is larger than the number of components of  $\Gamma$ , or (b) the number of cycles of  $\Gamma'$  is lower than the number of cycles of  $\Gamma$ . Since the number of connected components is bounded above by the (unchanging) number of vertices, every sequence of factorizations of type (6) is finite. Therefore, we proved:

**Corollary 3.7** (1) Each overlap can be reduced to an irreducible one by a finite number of factorizations of types (1)–(6).

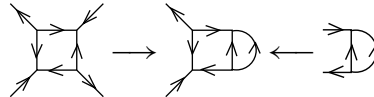
(2)  $\{T_{i0} \xrightarrow{S_i} T_i\}_{i \in I}$  are locally confluent if they are locally confluent on all irreducible overlaps of pairs of graphs in  $\{T_{i0}\}_{i \in I}$ .

Therefore, one has an effective procedure for deciding whether any finite set of reduction rules on graphs is locally confluent.

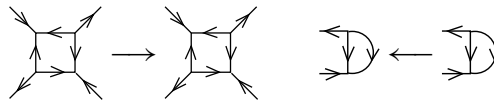
**Example 3.8** The graphs



have 5 different irreducible overlaps: four overlaps of the form



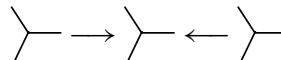
and one “trivial” overlap



**Example 3.9** The graphs

$$\Gamma_1 = \Gamma_2 = \text{Y-shape}$$

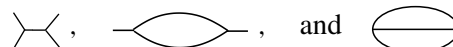
have three different irreducible overlaps of the form



and one trivial overlap,



Notice that the embeddings of  $\Gamma_1, \Gamma_2$  into



are not irreducible since they factor through the trivial overlap.

### 3.6 Proving confluence of reduction rules of surface graphs

An *overlap* of manifold graphs  $\Gamma_1 \subset M_1, \Gamma_2 \subset M_2$  is a graph  $\Gamma$  in a manifold  $M$  together with isotopy classes of embeddings  $(\Gamma_1, M_1) \hookrightarrow (\Gamma, M), (\Gamma_2, M_2) \hookrightarrow (\Gamma, M)$ .

$$f: \mathcal{O}' = ((\Gamma_1, M_1) \hookrightarrow (\Gamma', M'), (\Gamma_2, M_2) \hookrightarrow (\Gamma', M')) \rightarrow ((\Gamma_1, M_1) \hookrightarrow (\Gamma, M), (\Gamma_2, M_2) \hookrightarrow (\Gamma, M)) = \mathcal{O}$$

is a factorization of  $O$ , if for certain representatives

$$\iota'_1: (\Gamma_1, M_1) \hookrightarrow (\Gamma, M) \quad \text{and} \quad \iota'_2: (\Gamma_2, M_2) \hookrightarrow (\Gamma, M),$$

of embeddings of  $O'$ ,  $f\iota_1$  and  $f\iota_2$  belong to isotopy classes of embeddings of  $O$ .

As before, given reduction rules,  $\{T_{i0} \xrightarrow{S_i} \sum_k r_{ik} T_{ik}\}_{i \in I}$ , where  $S_i$  takes place in a manifold  $M_i$ , each overlap  $O = (\iota_1: (T_{i0}, M_i) \hookrightarrow (\Gamma, M), \iota_2: (T_{j0}, M_j) \hookrightarrow (\Gamma, M))$  leads to two different reductions of  $\Gamma$ . Rules  $\{S_i\}_{i \in I}$  are locally confluent if they are locally confluent on all overlaps of graphs  $T_{i0}$ ,  $i \in I$ . As before we consider factorization of overlaps and observe that if rules  $\{S_i\}_{i \in I}$  are locally confluent on  $O$  then they are locally confluent on all overlaps which factor through  $O$ .

A factorization  $f: O' \rightarrow O$  is trivial if  $f: M' \rightarrow M$  is isotopic to a homeomorphism. As before, an overlap is *irreducible* if it does not admit a non-trivial factorization.

In this section we are going to develop an algorithm for proving local confluence of overlaps of surface graphs. Observe that we cannot apply verbatim the method of the previous section to our current setting since Corollary 3.6 and Corollary 3.7(1) and (2) fail for surface graphs:

**Lemma 3.10** (1) *If every component of  $F$  has a non-empty boundary then no overlap  $((\Gamma_1, F_1) \hookrightarrow (\Gamma, F), (\Gamma_2, F_2) \hookrightarrow (\Gamma, F))$  is irreducible.*

(2) *If every component of  $F$  has a non-empty boundary then no overlap  $((\Gamma_1, F_1) \hookrightarrow (\Gamma, F), (\Gamma_2, F_2) \hookrightarrow (\Gamma, F))$  factors through an irreducible one.*

(3) *If  $F$  is closed,  $F \neq S^2, RP^2$ , and  $\Gamma$  is either empty or it is a contractible loop in  $F$ , then  $(\Gamma, F)$  has infinitely many irreducible overlaps with itself.*

**Proof** (1) Let  $F'$  be  $F$  with a disk removed from one of its components,  $C$ . By imagining the disk lying “very close” to  $\partial C$ , one can isotope  $\iota_1, \iota_2$  to  $\iota'_1, \iota'_2$  so that  $\iota'_1(F_1) \cup \iota'_2(F_2) \subset F'$ . Consequently,  $(\iota_1, \iota_2)$  factors through  $(\iota'_1: (\Gamma_1, F_1) \hookrightarrow (\Gamma', F'), \iota'_2: (\Gamma_2, F_2) \hookrightarrow (\Gamma', F'))$  via the embedding  $f: F' \rightarrow F$ . This is a non-trivial factorization, contradicting the initial assumption.

(2) If an overlap as above factors through  $(\iota_1: (\Gamma_1, F_1) \hookrightarrow (\tilde{\Gamma}, \tilde{F}), \iota_2: (\Gamma_2, F_2) \hookrightarrow (\tilde{\Gamma}, \tilde{F}))$  then  $\tilde{F} \subset F$  and consequently, every component of  $\tilde{F}$  has a non-empty boundary.

(3) For any diffeomorphism  $f: (\Gamma, F) \rightarrow (\Gamma, F)$ , diffeomorphisms  $\iota_1 = f: (\Gamma, F) \rightarrow (\Gamma, F)$  and the identity map  $\iota_2 = f: (\Gamma, F) \rightarrow (\Gamma, F)$  form an irreducible overlap which we denote by  $O_f$ . Notice that  $O_f = O_{f'}$  if and only if  $f'$  is isotopic to  $f$ . Since the mapping class group of  $F$  is infinite, there are infinitely many irreducible overlaps of this type.  $\square$



We will attempt to resolve these difficulties now.  $\{O_j\}_{j \in J}$  is a *basis* of overlaps of  $(\Gamma_1, F_1)$  and  $(\Gamma_2, F_2)$  if every overlap of these surface graphs factors through  $O_j$  for some  $j \in J$ .

**Corollary 3.11** *The rules  $S_i: \Gamma_{i0} \rightarrow \sum_k r_{ik} \Gamma_{ik}$ ,  $i \in I$ , are locally confluent, if they are locally confluent on a certain basis of overlaps of pairs of graphs in  $\{(\Gamma_{i0}, F_i)\}_{i \in I}$ .*

Lemma 3.10(2) shows that not every pair of graphs in surfaces has a finite basis of overlaps. Furthermore, basis of overlaps are generally not unique. Nonetheless, we are going to show that any two simple graphs in surfaces have a finite basis of overlaps. We say that  $\Gamma \subset F$  is *simple* if  $\Gamma$  is connected and every component  $C$  of  $F \setminus \Gamma$  is either  $D^2$  or an annulus whose one boundary component lies in  $\Gamma$  and the other in  $\partial F$ .

**Theorem 3.12** *Any two simple graphs have a finite basis of overlaps.*

Our proof is also an algorithm for finding such a finite basis.

For any graph  $\Gamma \hookrightarrow F$  there is an  $\varepsilon_0 > 0$  such that  $\varepsilon$ -neighborhoods of  $\Gamma$  in  $F$  are diffeomorphic to each other for all  $\varepsilon < \varepsilon_0$ . Denote such  $\varepsilon$ -neighborhood by  $\nu(\Gamma)$  and we call it a framing of  $\Gamma$ . Each framing of  $\Gamma$  retracts onto  $\Gamma$  and each finite abstract graph has finitely many different framings only.

Any overlap of  $(\Gamma_1, \nu(\Gamma_1))$  and  $(\Gamma_2, \nu(\Gamma_2))$  factors through an overlap

$$\iota_1: (\Gamma_1, \nu(\Gamma_1)) \hookrightarrow (\Gamma, F), \iota_2: (\Gamma_2, \nu(\Gamma_2)) \hookrightarrow (\Gamma, F),$$

such that  $(\iota_1: \Gamma_1 \hookrightarrow \Gamma, \iota_2: \Gamma_2 \hookrightarrow \Gamma)$  is an irreducible overlap of abstract graphs and  $F$  is a framing of  $\Gamma$ . Consequently, such overlaps form a basis of overlaps of  $(\Gamma_1, \nu(\Gamma_1))$  and  $(\Gamma_2, \nu(\Gamma_2))$ . Denote them by  $O_1, \dots, O_d$ .

Now assume that  $\Gamma_1$  and  $\Gamma_2$  are embedded into  $F_1, F_2$  in such way that they are simple graphs. We extend each basic overlap

$$O_i = (\iota_1: (\Gamma_1, \nu(\Gamma_1)) \hookrightarrow (\Gamma, F), \iota_2: (\Gamma_2, \nu(\Gamma_2)) \hookrightarrow (\Gamma, F))$$

constructed above to an overlap of  $(\Gamma_1, F_1)$  and  $(\Gamma_2, F_2)$  as follows: Every component  $B$  of  $\partial F$ , disjoint from  $\Gamma$ , is parallel to a unique cycle  $\alpha_B \subset \Gamma$ . If the preimage  $\iota_i^{-1}(\alpha_B)$  for either  $i = 1$  or  $2$  is a circle in  $F_i$  which bounds a disk  $D_i \subset F_i$  containing  $\iota_i^{-1}(B)$  then we attach a disk to  $F$  along  $B$  and we extend  $\iota_i$  over  $D_i$  for those  $i = 1, 2$  which satisfy the above condition. By performing these operations for all components of  $\partial F$  disjoint from  $\Gamma$ , we extend  $O_i$  to an overlap  $\bar{O}_i$  of  $(\Gamma_1, F_1)$  and  $(\Gamma_2, F_2)$ . Notice that every overlap of these graphs which restricts to  $O_i$  must factor through  $\bar{O}_i$ . Therefore we proved

**Corollary 3.13**  $\bar{O}_1, \dots, \bar{O}_d$  is a basis of overlaps of  $(\Gamma_1, F_1)$  and  $(\Gamma_2, F_2)$ .

### 4 $A_1$ -webs

Interesting examples of confluent and terminal reduction rules come from Kuperberg’s spider webs associated with simple Lie algebras of rank  $\leq 2$ . These are spaces of graphs in  $D^2$  considered modulo certain relations, of the type defined in Introduction. The classes of graphs considered and the relations between them depend on the Lie algebra in question. Because of their relations to quantum invariants, it is important to consider Kuperberg’s webs in surfaces other than  $D^2$ , even though they are not spiders anymore, since the join operation is no longer defined. For that reason Kuperberg’s graphs in surfaces other than  $D^2$  will simply be called *webs*. Although Kuperberg’s original reduction rules are not confluent for webs we will show that these rules can be extended to finite, confluent, and terminal sets of reduction rules.

The  $A_1$ -webs without external vertices are unoriented link diagrams. To put such diagrams in the framework of surface graphs, we define *crossings* as marked 4-valent vertices depicted as



We require that opposite edges meeting at any crossing have equal labels, when taken with coinciding orientations. The notions of linear graphs, reduction rules, local and global confluence extend to graphs with crossings in an obvious way. Furthermore, the method of proving local confluence discussed in Section 3.6 holds for graphs with crossings as well.

Consider now a surface  $F$  (not necessarily oriented) with a distinguished set of base points  $B \subset \partial F$  (possibly empty). An  $A_1$ -web in  $(F, B)$  is an unoriented graph all of whose internal vertices are crossings and all of whose external vertices are points of  $B$ . (Such graphs in  $D^2$  are called *unoriented tangle diagrams* with endpoints in  $B$ .) We denote the set of all  $A_1$ -webs in  $(F, B)$  by  $\mathcal{W}_{A_1}(F, B)$ . Let  $R$  be a fixed ring with a distinguished invertible element  $A$ . The  $A_1$ -web space over  $R$  is the  $R$ -module

$$\mathbb{A}_1(F, B, R) = R\mathcal{W}_{A_1}(F, B)/\mathcal{R}(T_1, T_2),$$

where

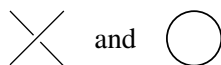
$$T_1 = \left( \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} - A \right) \left( -A^{-1} \begin{array}{c} \frown \\ \smile \end{array} \right)$$

$$T_2 = \bigcirc + (A^2 + A^{-2})\emptyset.$$

Here and further on, all relations take place in  $D^2$  and all 1-valent vertices are external, unless stated otherwise. The above relations suggest the obvious reduction rules:

$$S_1: \begin{array}{c} \diagup \\ \diagdown \end{array} \rightarrow A \left( + A^{-1} \begin{array}{c} \frown \\ \smile \end{array} \right), \quad S_2: \bigcirc \rightarrow -(A^2 + A^{-2})\emptyset.$$

Denote the number of crossings and connected components of  $\Gamma \in \mathcal{W}_{A_1}(F, B)$  by  $v(\Gamma)$  and  $c(\Gamma)$ , respectively. If  $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$  is given the lexicographic ordering then these reduction rules replace each graph  $\Gamma$  by a combination of graphs  $\Gamma_i$  such that  $(v(\Gamma_i), c(\Gamma_i)) < (v(\Gamma), c(\Gamma))$ . Therefore, the rules  $S_1, S_2$  are terminal. The irreducible graphs are those with no crossings and no contractible loops. Since



have no non-trivial overlaps, they are locally confluent and, hence, also globally confluent. Now Theorem 2.3(2) provides answers to questions (1) and (2) of Section 1:

**Corollary 4.1** For any  $F, R$  and  $q$ ,  $\mathbb{A}_1(F, B, R)$  is the free  $R$ -module with a basis given by finite collections of disjoint non-trivial simple closed loops in  $F$ , including  $\emptyset$ .

## 5 $A_2$ -webs

Let  $F$  be a surface with a distinguished set of base points  $B \subset \partial F$  (possibly empty) which are marked by  $\pm 1$ . An  $A_2$ -web in  $(F, B)$  is an oriented graph  $\Gamma$  in  $F$  all of whose internal vertices are either 3-valent sinks or sources or 4-valent crossings:



and such that all external vertices are points of  $B$ . Furthermore, we require that the external edge adjacent to  $b \in B$  is oriented inwards or outwards according to the labeling of  $b$  by 1 or  $-1$ .

Denote the set of  $A_2$ -webs in  $F$  by  $\mathcal{W}_{A_2}(F, B)$ . The  $A_2$ -web space is

$$\mathbb{A}_2(F, B, R) = R\mathcal{W}_{A_2}(F, B)/\mathcal{R}(T_1, T_2, T_3, T_4, T_5, T_6),$$

where

$$\begin{aligned}
 T_1 &= \text{circle with arrow} - (q + 1 + q^{-1})\emptyset, & T_2 &= \text{circle with arrow} - (q + 1 + q^{-1})\emptyset, \\
 T_3 &= \text{circle with two arrows} + (q^{\frac{1}{2}} + q^{-\frac{1}{2}})\text{circle with arrow}, & T_4 &= \text{square with arrows} - \text{crossing} - \text{crossing}, \\
 T_5 &= \text{trivalent vertex} - q^{\frac{1}{6}} \text{trivalent vertex} - q^{-\frac{1}{3}} \text{trivalent vertex}, & T_6 &= \text{trivalent vertex} - q^{-\frac{1}{6}} \text{trivalent vertex} - q^{\frac{1}{3}} \text{trivalent vertex}.
 \end{aligned}$$

and  $R$  is an arbitrary ring with a distinguished invertible element denoted by  $q^{\pm\frac{1}{6}} \in R$ . ( $R = \mathbb{C}[q^{\pm\frac{1}{6}}]$  in Kuperberg [36].)

The above relations suggest “obvious” reduction rules

$$\begin{aligned}
 S_1: \text{circle with arrow} &\rightarrow (q + 1 + q^{-1})\emptyset, & S_2: \text{circle with arrow} &\rightarrow (q + 1 + q^{-1})\emptyset, \\
 S_3: \text{circle with two arrows} &\rightarrow -(q^{\frac{1}{2}} + q^{-\frac{1}{2}})\text{circle with arrow}, & S_4: \text{square with arrows} &\rightarrow \text{crossing} + \text{crossing}, \\
 S_5: \text{trivalent vertex} &\rightarrow q^{\frac{1}{6}} \text{trivalent vertex} + q^{-\frac{1}{3}} \text{trivalent vertex}, & S_6: \text{trivalent vertex} &\rightarrow q^{-\frac{1}{6}} \text{trivalent vertex} + q^{\frac{1}{3}} \text{trivalent vertex}.
 \end{aligned}$$

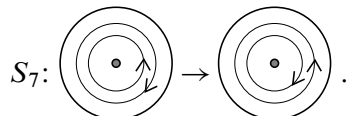
Denote the number of connected components, 3-valent vertices, and crossings of any  $\Gamma \in \mathcal{W}_{A_2}(F, B)$  by  $c(\Gamma)$ ,  $v_3(\Gamma)$ , and  $v_4(\Gamma)$ , respectively. If  $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$  is given the lexicographic ordering, then the above reduction rules replace  $\Gamma$  by a linear combination of graphs  $\Gamma_i$  such that  $(v_4(\Gamma_i), v_3(\Gamma_i), c(\Gamma_i)) < (v_4(\Gamma), v_3(\Gamma), c(\Gamma))$ . Consequently, these reduction rules are terminal. However, they are not confluent for  $F \neq D^2, S^2$ ! Indeed, any surface  $F \neq D^2, S^2$  contains an annulus whose core is not contractible in  $F$  and the two possible applications of  $S_4$  to

$$\Gamma = \text{annulus with arrows},$$

followed by  $S_1, S_2$ , reduce it to  $X_1 + (q + 1 + q^{-1})\emptyset$  and  $X_2 + (q + 1 + q^{-1})\emptyset$ , where

$$X_1 = \text{annulus with arrows} \quad \text{and} \quad X_2 = \text{annulus with arrows}.$$

Since  $X_1$  and  $X_2$  are irreducible and not isotopic, the reduction rules are not confluent. In order to remedy this imperfection, we need to consider an additional reduction rule taking place in an annulus:



Note that  $S_1, \dots, S_7$  are terminal in the oriented case but not terminal in the unoriented case, c.f. last paragraph of Section 3.2. Indeed,  $S_7$  is its own inverse in the unoriented case! Therefore, we consider  $A_2$ -webs in oriented surfaces only and work in the category of oriented surfaces from now on. (We assume that all surfaces appearing in reduction rules  $S_1, \dots, S_7$  have counterclockwise orientation.)

One checks all overlaps for  $S_1, \dots, S_7$  and concludes that  $S_1, \dots, S_7$  are locally confluent on all of them<sup>3</sup>. Therefore, by Theorem 2.3, we conclude:

**Corollary 5.1** *The reduction rules  $S_1, \dots, S_7$  are both terminal and confluent for graphs in  $\mathcal{W}_{A_2}(F, B)$ , for any oriented surface  $F$  and any set of marked base points  $B \subset \partial F$ . Consequently,  $\mathbb{A}_2(F, B, R)$  is a free  $R$ -module with a basis composed of irreducible graphs in  $\mathcal{W}_{A_2}(F, B)$ .*

Observe that irreducible  $A_2$ -webs in  $D^2$  are those which have no  $S^1$ 's, no internal bi-gons, and no internal 4-gons. (Such graphs in  $D^2$  are called *non-elliptic* by Kuperberg [36].) While these terms are intuitively obvious for graphs in  $D^2$ , they do require clarification for graphs in other surfaces.

Components of  $F \setminus \Gamma$  are *faces of  $\Gamma$* . A face is *internal* if it is disjoint from  $\partial F$ . An internal face is called an  $n$ -gon if it is a disk bounded by a sequence of  $n$  edges of  $\Gamma$ . (The orientations of the edges are irrelevant.) An  $n$ -gon is *true* if all its boundary edges are distinct; otherwise it is *fake*. For example, the 4-gon in  $S^1 \times I$  bounded by the edges  $E_1, E_2, E_3, E_2$  depicted in Figure 3 is fake.

A loop in  $\Gamma$  bounding a disk in  $F \setminus \Gamma$  is called a 0-gon. The next statement follows directly from  $A_2$ -web reduction rules:

**Corollary 5.2** *The irreducible graphs in  $\mathcal{W}_{A_2}(F, B)$  are precisely those with no 0-gons, no true bi-gons, and no true 4-gons.*

<sup>3</sup> Recall that irreducible overlaps of  $S_3$  and  $S_4$  (as abstract graphs) were classified in Example 3.8.

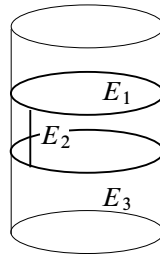


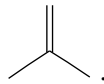
Figure 3

### 6 $B_2$ -webs

Let  $F$  be a surface together with a specified finite set of base points  $B \subset \partial F$  (possibly empty), each of them marked by 1 or 2. Throughout this section we work in the category of unoriented surfaces.

**Definition 6.1** Let  $\mathcal{W}_{B_2}(F, B)$  be the set of all labeled graphs  $\Gamma$  in  $F$ , with  $\Lambda = \{1, 2\}$  and  $\nu$  the identity, c.f. Section 3, such that

- (1) the labels of edges adjacent to points of  $B$  coincide with their labels, and
- (2) all internal vertices of  $\Gamma$  are of the form



(The labels are depicted either by a single or double line.)

The  $B_2$ -web space is

$$\mathbb{B}_2(F, B, R) = R\mathcal{W}_{B_2}(F, B)/\mathcal{R}(T_1, \dots, T_6),$$

where

$$\begin{aligned} T_1 &= \bigcirc + (q^2 + q + q^{-1} + q^{-2})\emptyset & T_2 &= \bigcirc\bigcirc - (q^3 + q + 1 + q^{-1} + q^{-3})\emptyset \\ T_3 &= \text{double line to circle} & T_4 &= \text{double line to circle to double line} + (q + 2 + q^{-1})\text{double line} \\ T_5 &= \text{triangle with double lines} & T_6 &= \text{crossing} - \text{crossing} + \text{arc} - \text{arc} \end{aligned}$$

and  $R$  is an arbitrary ring with a distinguished invertible element  $q \in R$ .  $B_2$ -webs with crossings are discussed in the next subsection.

While  $T_1, \dots, T_5$  yield “obvious” reduction relations, which we denote by  $S_1, \dots, S_5$ , relation  $T_6$  creates a problem since the rule

$$\begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array} \rightarrow \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array} - \left( \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array} + \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array} \right)$$

is its own inverse and, hence, it is not terminal. Following Kuperberg’s idea, we remedy this problem by allowing  $B_2$ -webs to have 4-valent vertices subject to a relation  $T'_6 = 0$ , where

$$T'_6 = \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array} - \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array} - \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array}.$$

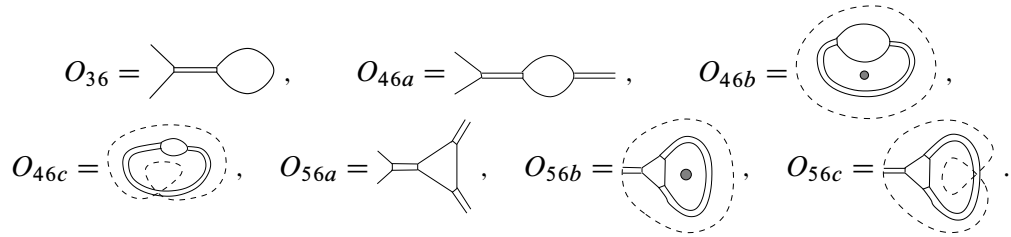
We denote this extended family of webs by  $\mathcal{W}'_{B_2}(F, B)$ . Note that  $T'_6$  does not introduce any new relations and, therefore,

$$\mathbb{B}_2(F, B, R) = R\mathcal{W}_{B_2}(F, B)/\mathcal{R}(T_1, \dots, T_6) = R\mathcal{W}'_{B_2}(F, B)/\mathcal{R}(T_1, \dots, T_5, T'_6).$$

Now,  $T'_6$  suggests the reduction rule

$$S_6: \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array} \rightarrow \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array} + \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array}$$

Since each of the reduction rules  $S_1, \dots, S_6$  either decreases the number of vertices or decreases the number of connected components without increasing the number of vertices, these reduction rules are terminal. However, they are not confluent in general<sup>4</sup>.  $S_1, \dots, S_6$  have the following basis of overlaps:



Dashed lines denote boundaries of surfaces. (No dashed line is drawn for diagrams in  $D^2$ .) Overlaps  $O_{46b}$  and  $O_{56b}$  take place in annuli and  $O_{46c}$  and  $O_{56c}$  in Möbius bands. Unfortunately,  $S_1, \dots, S_6$  are not locally confluent on these overlaps. For example,

$$0 \xleftarrow{S_3} \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array} \xrightarrow{S_6, S_1} \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array} - (q^2 + q + q^{-1} + q^{-2}),$$

and both of these linear graphs are irreducible with respect to  $S_1, \dots, S_6$ . In order to remedy that, we introduce the following new rules (preserving relations  $S_1 - S_6$ ):

<sup>4</sup> These rules may be confluent for certain choices of  $R$  and  $q$ , but not for all.

$$\begin{aligned}
 S_7: & \text{ } \bigcirc \rightarrow (q^2 + q + q^{-1} + q^{-2}) \bigcirc, & S_8: & \text{ } \bigcirc = \rightarrow -(q + 2 + q^{-1}) \bigcirc =, \\
 S_9: & \text{ } \bigtriangle \rightarrow (q + 2 + q^{-1}) \bigtriangle, \\
 S_{10}: & \text{ } \bigcirc \rightarrow -(q + 2 + q^{-1}) \bigcirc + (q^2 + q + q^{-1} + q^{-2}) \bullet, \\
 S_{11}: & \text{ } \bigcirc \rightarrow -(q + 2 + q^{-1}) \bigcirc + (q^2 + q + q^{-1} + q^{-2}) \bigcirc, \\
 S_{12}: & \text{ } \bigcirc \rightarrow 0, & S_{13}: & \text{ } \bigcirc \rightarrow 0.
 \end{aligned}$$

Rules  $S_{10}, S_{12}$  take place in annuli and rules  $S_{11}, S_{13}$  in Möbius bands. (The graph on the left side of rule  $S_{11}$  has a single vertex, and the one on the left side of  $S_{13}$  has two vertices.)

Now  $S_1, \dots, S_{13}$  are locally confluent on  $O_{36}, \dots, O_{56c}$ , but the new rules,  $S_7-S_{13}$ , lead to new overlaps:

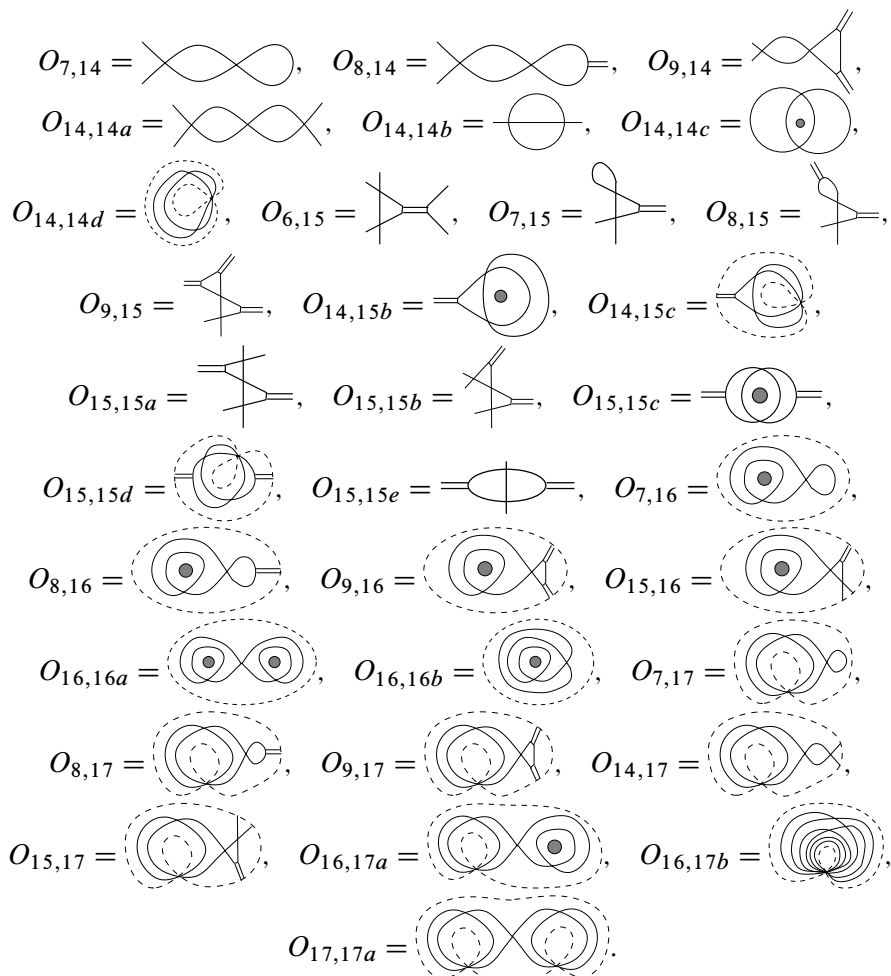
$$\begin{aligned}
 O_{77} &= \infty, & O_{68} &= \bigcirc \bigtriangle, & O_{78} &= \infty =, & O_{88} &= = \infty =, \\
 O_{69a} &= \bigtriangle, & O_{69b} &= \bigtriangle \bullet, & O_{69c} &= \bigtriangle \bigcirc, & O_{79} &= \bigtriangle \bigtriangle, \\
 O_{89} &= = \bigtriangle, & O_{99} &= \bigtriangle \bigtriangle, & O_{6,12} &= \bigcirc \bigcirc, & O_{6,13} &= \bigcirc \bigcirc.
 \end{aligned}$$

Now, we make  $S_1, \dots, S_{13}$  locally confluent on these overlaps by introducing the following new reduction rules:

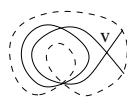
$$\begin{aligned}
 S_{14}: & \text{ } \bigtriangle \rightarrow -(q + 2 + q^{-1}) \bigtriangle - (q^2 + 2q + 2 + 2q^{-1} + q^{-2}) \bigtriangle, \\
 S_{15}: & \text{ } \bigtriangle = \rightarrow (q + 2 + q^{-1}) \left( \bigtriangle + \bigtriangle \right), \\
 S_{16}: & \text{ } \bigcirc \bigtriangle \rightarrow (q + 2 + q^{-1}) \bigcirc \bigtriangle - (q^2 + q + q^{-1} + q^{-2}) \bigcirc \bullet, \\
 S_{17}: & \text{ } \bigcirc \bigtriangle \rightarrow (q + 2 + q^{-1}) \bigcirc \bigtriangle - (q^2 + q + q^{-1} + q^{-2}) \bigcirc.
 \end{aligned}$$

The new overlaps now are:





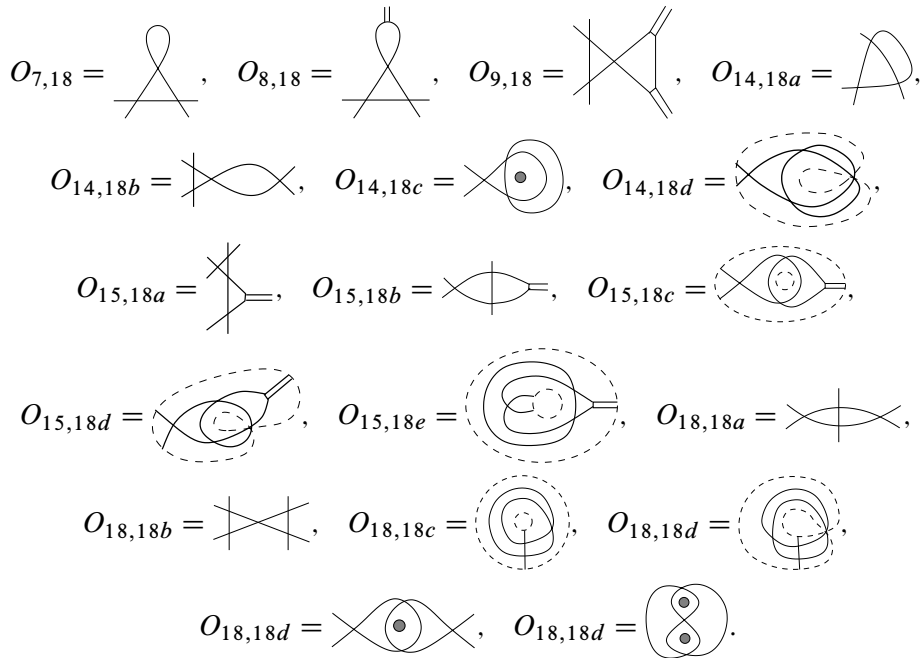
Finally, there is an overlap  $O_{17,17b}$  in a Klein bottle obtained by gluing two Möbius bands along their boundaries and taking a union  $\Gamma$  of two graphs



overlapping at vertex  $v$ . (Hence,  $\Gamma$  has three 4-valent vertices and no vertices of other valences.) Since  $S_1-S_{17}$  are not locally confluent on  $O_{6,15}$ , we add yet another reduction rule:

$$S_{18}: \text{Diagram} \rightarrow (q+2+q^{-1}) \left( \text{Diagram 1} + \text{Diagram 2} + \text{Diagram 3} \right) + (q^2+4q+6+4q^{-1}+q^{-2}) \text{Diagram 4}$$

One can check that  $S_1 - S_{18}$  are locally confluent on all overlaps listed so far. However,  $S_{18}$  leads to new overlaps:



By checking all of the above overlaps, we conclude that  $S_1 - S_{18}$  are locally confluent on all of them! Therefore, we proved:

**Theorem 6.2** *The reduction rules  $S_1, \dots, S_{18}$  are both terminal and confluent for  $\mathcal{W}_{B_2}(F, B)$ , for any surface  $F$  and  $B \subset \partial F$  (both in orientable and unorientable categories).*

For sets  $B \subset \partial F$  all of whose points are marked by 1, denote the set of all graphs in  $\mathcal{W}_{B_2}(F, B)$  with single edges only (that is, edges labeled by 1), by  $\mathcal{W}'_{B_2}(F, B)$ . (Graphs in  $\mathcal{W}'_{B_2}(F, B)$  may include double loops.)

**Proposition 6.3** *For any  $F$  and  $B \subset \partial F$  as above,*

(1) *the embedding  $\mathcal{W}'_{B_2}(F, B) \hookrightarrow \mathcal{W}_{B_2}(F, B)$  induces an isomorphism*

$$\phi: RW'_{B_2}(F, B) / \mathcal{R}(S_1, S_2, S_7, S_{10}, S_{11}, S_{14}, S_{16}, S_{17}) \rightarrow \mathbb{B}_2(F, B, R).$$

(2) *the rules  $S_1, S_2, S_7, S_{10}, S_{11}, S_{14}, S_{16}, S_{17}$  are terminal and confluent for graphs in  $\mathcal{W}'_{B_2}(F, B)$ .*

**Proof** Since internal double edges are resolvable by  $S_6$ ,  $\phi$  is onto. The map  $\phi$  is 1–1 as well: If  $\phi(x) = 0$  in  $\mathcal{W}_{B_2}(F, B)$ , then by confluence of  $S_1$ – $S_{18}$ ,  $x$  can be reduced to 0 by these rules. Since  $x \in R\mathcal{W}'_{B_2}(F, B)$ , the rules which do not contain double edges are sufficient to reduce  $x$  to 0.  $\square$

### 6.1 $B_2$ –webs with crossings

For any  $F$  and  $B \subset \partial F$  as in Section 6, let  $\mathcal{W}_{B_2}^c(F, B)$  be the set of all labeled graphs  $\Gamma$  in  $F$ , with edges labeled by 1 and 2 such that

- (1) the labels of edges adjacent to points of  $B$  coincide with their labels,
- (2) all internal vertices of  $\Gamma$  are of the form



Resolutions of crossings are provided by

$$\begin{aligned}
 C_1: & \text{Crossing} \rightarrow -q^{\frac{1}{2}} \left( -q^{-\frac{1}{2}} \text{Cup} + \frac{1}{q^{\frac{1}{2}} + q^{-\frac{1}{2}}} \text{Crossing} \right) \\
 C_2: & \text{Double Crossing} \rightarrow \frac{q^{-\frac{1}{2}}}{q^{\frac{1}{2}} + q^{-\frac{1}{2}}} \text{Y-junction} + \frac{q^{\frac{1}{2}}}{q^{\frac{1}{2}} + q^{-\frac{1}{2}}} \text{Double Crossing} \\
 C_3: & \text{Double Crossing} \rightarrow \frac{q^{\frac{1}{2}}}{q^{\frac{1}{2}} + q^{-\frac{1}{2}}} \text{Y-junction} + \frac{q^{-\frac{1}{2}}}{q^{\frac{1}{2}} + q^{-\frac{1}{2}}} \text{Double Crossing} \\
 C_4: & \text{Quadruple Crossing} \rightarrow q \left( +q^{-1} \text{Cup} + \frac{1}{q + 2 + q^{-1}} \text{Square} \right)
 \end{aligned}$$

Since the crossing diagrams do not add any new overlaps, we conclude with:

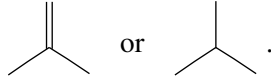
**Corollary 6.4** *If  $q^{\frac{1}{2}} \in R$  is such that  $q + 1$  is invertible in  $R$ , then resolutions  $S_1, \dots, S_{18}$  together with  $C_1, \dots, C_4$  are confluent and terminal for any surface  $F$  and  $B \subset \partial F$ .*

## 7 $G_2$ –webs

Let  $F$  be a surface together with a specified finite set of base points  $B \subset \partial F$  (possibly empty), each of them marked by 1 or 2. Throughout this section we work in the category of unoriented surfaces.

**Definition 7.1** Let  $\mathcal{W}_{G_2}(F, B)$  be the set of all labeled graphs  $\Gamma$  in  $F$ , whose edges are labeled by 1 or 2 and such that

- (1) the labels of edges adjacent to points of  $B$  coincide with their labels,
- (2) all internal vertices of  $\Gamma$  are of the form



The  $G_2$ -web space is

$$\mathbb{G}_2(F, B, R) = R\mathcal{W}_{G_2}(F, B)/\mathcal{R}(S_1, \dots, S_8),$$

where

- $S_1$ :  $\rightarrow (q^5 + q^4 + q + 1 + q^{-1} + q^{-4} + q^{-5})\emptyset$
- $S_2$ :  $\rightarrow (q^9 + q^6 + q^5 + q^4 + q^3 + q + 2 + q^{-1} + q^{-3} + q^{-4} + q^{-5} + q^{-6} + q^{-9})\emptyset$
- $S_3$ :  $\rightarrow 0$
- $S_4$ :  $\rightarrow -(q^3 + q^2 + q + q^{-1} + q^{-2} + q^{-3})$  —
- $S_5$ :  $\rightarrow (q^2 + 1 + q^{-2})$
- $S_6$ :  $\rightarrow -(q + q^{-1}) \left( \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} \right) + (q + 1 + q^{-1}) \left( \begin{array}{c} \diagup \diagup \\ \diagdown \diagdown \end{array} + \begin{array}{c} \diagdown \diagdown \\ \diagup \diagup \end{array} \right)$
- $S_7$ :  $\rightarrow \left( \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} + \begin{array}{c} \diagup \diagup \\ \diagdown \diagdown \end{array} + \begin{array}{c} \diagdown \diagdown \\ \diagup \diagup \end{array} + \begin{array}{c} \diagup \diagup \\ \diagdown \diagdown \end{array} + \begin{array}{c} \diagdown \diagdown \\ \diagup \diagup \end{array} \right) - \left( \begin{array}{c} \diagup \diagup \\ \diagdown \diagdown \end{array} + \begin{array}{c} \diagdown \diagdown \\ \diagup \diagup \end{array} + \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} + \begin{array}{c} \diagup \diagup \\ \diagdown \diagdown \end{array} + \begin{array}{c} \diagdown \diagdown \\ \diagup \diagup \end{array} \right)$
- $S_8$ :  $\rightarrow \left( \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} - \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} - \frac{1}{q^2 - 1 + q^{-2}} \right) \left( + \frac{1}{q + 1 + q^{-1}} \begin{array}{c} \diagup \diagup \\ \diagdown \diagdown \end{array} + \frac{1}{q + 1 + q^{-1}} \begin{array}{c} \diagdown \diagdown \\ \diagup \diagup \end{array} \right)$

(Be advised that reduction rule  $S_7$  has wrong signs in Kuperberg [36].) Reduction rules for crossings are listed in [36].

By checking all overlaps we conclude:

**Theorem 7.2** If  $F$  is orientable and  $q^2 - 1 + q^{-2}, q + 1 + q^{-1}$  are invertible in  $R$ , then reduction rules  $S_1, \dots, S_8$  together with

$$S_9: \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \rightarrow \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} - (q + 1 + q^{-1}) \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array}$$

(taking place in annulus) are confluent and terminal.<sup>5</sup>

### 8 Partition category and dichromatic reduction rules

An example of terminal and confluent reduction rules for abstract graphs comes from the dichromatic polynomial, cf Yetter [66].

Let  $\mathcal{G}_n$  be the set of all unoriented graphs with  $n$  external vertices labeled from 1 to  $n$  and let  $R = \mathbb{Z}[p, q, s, v, w_1, w_2]$ . Consider reduction rules

$$\begin{aligned}
 S_{k,l}: & k \left\{ \begin{array}{c} \text{---} \circ \text{---} \\ \vdots \\ \text{---} \circ \text{---} \end{array} \right\} l \rightarrow p \begin{array}{c} \text{---} \circ \\ \vdots \\ \text{---} \circ \end{array} + q \begin{array}{c} \text{---} \circ \text{---} \\ \vdots \\ \text{---} \circ \text{---} \end{array}, \\
 S_l: & \left\{ \begin{array}{c} \text{---} \circ \text{---} \\ \vdots \\ \text{---} \circ \text{---} \end{array} \right\} l \rightarrow s \begin{array}{c} \text{---} \circ \\ \vdots \\ \text{---} \circ \end{array}, \\
 S_v: & \circ \rightarrow v \emptyset, \\
 S_{bw}: & \bullet \text{---} \circ \rightarrow w_1 \bullet \\
 S_{bwb}: & \bullet \text{---} \circ \text{---} \bullet \rightarrow w_2 \bullet \text{---} \bullet
 \end{aligned}$$

where external (respectively, internal) vertices are denoted by black (respectively, white) nodes and  $k, l \in \mathbb{Z}_{\geq 0}$ .

- Theorem 8.1** (a) *The above reduction rules are terminal and confluent.*  
 (b) *Connected components of irreducible graphs are either isolated external vertices,  $\bullet$ , or  $\bullet\text{---}\bullet$  or at least 3 external vertices connected to a single internal vertex.*

The termination of the above rules is obvious. The proof of confluence is left to the reader.

**Corollary 8.2** *Irreducible graphs in  $\mathcal{G}_n$  are in 1–1 correspondence to partitions of  $\{1, \dots, n\}$ .*

The only irreducible graph with no external vertices is the empty graph  $\emptyset$ . Consequently,  $R\mathcal{G}_0/\mathcal{R}(S_{kl}, S_l, k, l \geq 0, S_0)$  is a cyclic  $R$ -module generated by  $\emptyset$ . The projection  $\mathcal{G}_0 \rightarrow R\mathcal{G}_0/\mathcal{R}(S_{kl}, k, l \geq 0, S_0) \simeq R$  followed by substitution

$$R \xrightarrow{p \rightarrow 1, s \rightarrow 1+q} \mathbb{Z}[q, v]$$

is the dichromatic polynomial of graphs. A generalization of dichromatic polynomial to ribbon graphs is considered by Bollobás and Riordan [6]. It can be defined by reductions rules similar to those above as well.

<sup>5</sup> We did not check confluence for unoriented surfaces.

## 8.1 Partition category

Let  $R$  be a ring with a specified  $\delta^{\pm 1} \in R$ . A version of the “dichromatic” reduction rules appears in the context of the partition category. There are two types of that category: symmetric and planar one. In each of them, objects are non-negative integers.

In the symmetric partition category, the morphisms  $[n] \rightarrow [m]$  are  $R$ -linear combinations of abstract graphs with  $n + m$  external vertices divided into an *input set* and an *output set* of  $n$  and  $m$  vertices, respectively. Vertices in each of these sets are numbered. The internal vertices have valency at most 3. Furthermore, the graphs are subject to the following rules:

$$\begin{array}{lll}
 T_1: \begin{array}{c} \diagup \quad \diagdown \\ \circ \quad \circ \\ \diagdown \quad \diagup \end{array} = \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array}, & T_2: \begin{array}{c} \circ \\ \circ \end{array} = \delta, & T_3: \begin{array}{c} \circ \quad \circ \\ \circ \quad \circ \end{array} = \begin{array}{c} \circ \quad \circ \end{array}, \\
 T_4: \begin{array}{c} \diagup \quad \diagdown \\ \circ \quad \circ \\ \diagdown \quad \diagup \end{array} = \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array}, & T_5: \begin{array}{c} \circ \quad \circ \\ \circ \quad \circ \end{array} = 1, & T_6: \begin{array}{c} \bullet \quad \circ \\ \circ \quad \circ \end{array} = \bullet, \\
 T_7: \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \\ \circ \end{array} = \cup, & T_8: \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \\ \circ \end{array} = \delta \downarrow.
 \end{array}$$

As before, empty/full dots represent internal/external vertices. The composition of morphisms  $\Gamma_1 \in \text{Mor}([n], [m])$ ,  $\Gamma_2 \in \text{Mor}([m], [k])$  is defined by identifying output vertices of  $\Gamma_1$  with corresponding input vertices of  $\Gamma_2$ . (The correspondence between these vertices is established by their numbering.) These vertices become internal in  $\Gamma_2 \circ \Gamma_1$ . The symmetric partition category is monoidal, with the tensor product of morphisms  $\Gamma_1 \in \text{Mor}([n_1], [m_1])$ ,  $\Gamma_2 \in \text{Mor}([n_2], [m_2])$  given by taking their disjoint union and shifting the numbering of the input and output vertices of  $\Gamma_2$  by  $n_1$  and  $m_1$  respectively.

In the planar partition category, one thinks of objects as sets  $[n] = \{\frac{1}{n+1}, \dots, \frac{n}{n+1}\} \subset [0, 1]$ . Morphisms  $[n] \rightarrow [m]$  are  $R$ -linear (manifold) graphs in  $[0, 1] \times [0, 1]$  without  $S^1$ 's, whose external vertices are  $[n] \times \{0\} \cup [m] \times \{1\}$ . These graphs are considered up to relations  $T_1, \dots, T_8$ . We denote the space of morphisms in the planar partition category by  $\text{Mor}^P([n], [m])$ . Compositions (respectively, tensor products) of morphisms are given by vertical (respectively, horizontal) stacking of graphs.

The importance of the planar (respectively, symmetric) partition category stems from the fact that the object  $[1]$  is the universal Frobenius algebra (respectively, commutative Frobenius algebra) in a monoidal (respectively symmetric monoidal) category. Consequently, monoidal (respectively, symmetric monoidal) functors from the planar (respectively, symmetric) partition category into the category of  $R$ -modules are in 1–1 correspondence with Frobenius algebras (respectively, commutative Frobenius

algebras) over  $R$ . Furthermore, the algebra of morphisms  $[n] \rightarrow [n]$  called the *partition algebra* appears in the theory of Potts models (see Jones [29] and Martin [40]).

Reduction relations going from left to right sides of equations  $T_1 - T_8$  are locally confluent but not terminal since  $T_1$  is invertible. (Nonetheless, we are going to prove that they are confluent.) The intuitive way of making these rules terminal is by allowing 4-valent vertices and by adding a reduction rule

$$\begin{array}{c} \diagup \quad \diagdown \\ \circ \quad \circ \\ \diagdown \quad \diagup \end{array} \longrightarrow \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array}.$$

Now to make the new rules terminal and locally confluent one needs to add an additional relation involving 5-valent vertices, and then one involving 6-valent vertices, etc. Finally, one arrives at:

**Theorem 8.3** (1) *The natural embedding of graphs of the symmetric partition category into (abstract) graphs factors to an isomorphism of  $R$ -modules*

$$\phi: \text{Mor}([n], [m]) \rightarrow \mathcal{G}_{n+m} / \mathcal{R}(\bar{S}_{k,l}, \bar{S}_l, \bar{S}_v, \bar{S}_{bw}, \bar{S}_{bwb}),$$

where

$$\begin{array}{l} \bar{S}_{k,l}: \left\{ \begin{array}{c} \diagup \quad \diagdown \\ \circ \quad \circ \\ \diagdown \quad \diagup \end{array} \right\} \rightarrow \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array}, \quad \bar{S}_l: \left\{ \begin{array}{c} \diagup \quad \diagdown \\ \diagdown \quad \diagup \end{array} \right\} \rightarrow \delta \begin{array}{c} \diagdown \quad \diagup \\ \diagup \quad \diagdown \end{array}, \\ \bar{S}_v: \circ \rightarrow \emptyset, \quad \bar{S}_{bw}: \bullet \circ \rightarrow \bullet, \quad \bar{S}_{bwb}: \bullet \circ \bullet \rightarrow \bullet \bullet \bullet. \end{array}$$

(Note that these are the dichromatic reduction rules for  $p = 0, q = v = w_1 = w_2 = 1, s = \delta$ ).

(2) *Similarly, the natural embedding of graphs of planar partition category into planar graphs in  $D^2$  with  $n + m$  external vertices factors to an isomorphism*

$$\phi^p: \text{Mor}^p([n], [m]) \rightarrow \mathcal{G}_{n+m}(D^2) / \mathcal{R}(\bar{S}_{k,l}, \bar{S}_l, \bar{S}_v, \bar{S}_{bw}, \bar{S}_{bwb}).$$

**Sketch of Proof** (1) The irreducible graphs in  $\mathcal{G}_{n+m}$  listed in Theorem 8.1 span  $\mathcal{G}_{n+m} / \mathcal{R}(\bar{S}_{k,l}, \bar{S}_l, \bar{S}_v, \bar{S}_{bw}, \bar{S}_{bwb})$ . Since all of these graphs are values of  $\phi$ , it is an epimorphism. To prove that  $\phi$  is 1-1 observe that connected components of every graph in  $\text{Mor}([n], [m])$  determine a partition of  $\{1, \dots, n + m\}$ . Let  $\text{Mor}^\tau([n], [m])$  be the subspace of  $\text{Mor}([n], [m])$  spanned by graphs associated with the partition  $\tau$ . Since  $T_1, \dots, T_8$  preserve partitions,

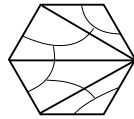
$$\text{Mor}([n], [m]) = \bigoplus_{\text{partitions } \tau} \text{Mor}^\tau([n], [m])$$

and, similarly,  $\mathcal{G}_{n+m} / \mathcal{R}(\bar{S}_{k,l}, \bar{S}_l, \bar{S}_v, \bar{S}_{bw}, \bar{S}_{bwb})$  decomposes into subspaces indexed by partitions, which by Theorem 8.1 are 1-dimensional. Since  $\phi$  preserves partition

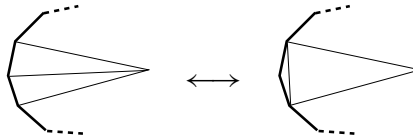
classes, it is enough to prove that  $\text{Mor}^\tau([n], [m]) \simeq R$  as an  $R$ -module. This follows from the following:

**Lemma 8.4** *Any two connected graphs with internal vertices of valency  $\leq 3$  are equivalent via relations  $T_1, \dots, T_8$ .*

**Proof**  $T_4$  and  $T_1$  allow to “slide” edges past 2-valent and 3-valent vertices. Therefore, all cycles in a graph can be transformed into loops, which can be eliminated by  $T_2$  and  $T_8$ . Furthermore, all internal 1-valent vertices can be removed by  $T_5, T_6$  and  $T_7$ . Consequently, each connected graph is equivalent to a tree with no internal 1-valent vertices. All such trees are related by  $T_1$  moves. There is a geometric way to see that. First, any such tree embeds into  $D^2$  and its dual corresponds to a division of a  $(n + m)$ -gon into triangles by non-intersecting diagonals. Here is an example of a planar tree and the corresponding dual triangulation of 6-gon:



Any two such triangulations are related by the move:



which is dual to  $T_1$ . □

The proof of Theorem 8.3(2) is analogous. □

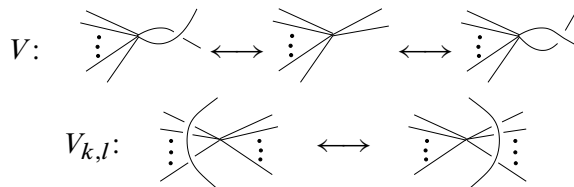
## 9 Application to knots

Now, we turn to spaces of dimension 3, which are the most difficult to deal with in the context of graph embeddings. Graphs in a 3-dimensional manifold  $M$  are the easiest to analyze if  $M$  is an  $I$ -bundle over a surface  $F$ ,  $I = [-1, 1]$ , since then each graph is represented by its diagram in  $F$  and such representations are unique up to Reidemeister moves:

$$RI: \left( \begin{array}{c} | \\ \circlearrowleft \\ | \end{array} \right) \longleftrightarrow \left( \begin{array}{c} | \\ \circlearrowright \\ | \end{array} \right), \quad RII: \left( \begin{array}{c} \diagup \\ \diagdown \end{array} \right) \longleftrightarrow \left( \begin{array}{c} \diagdown \\ \diagup \end{array} \right) \left( \begin{array}{c} \diagup \\ \diagdown \end{array} \right) \left( \begin{array}{c} \diagdown \\ \diagup \end{array} \right), \quad \text{and} \quad RIII: \left( \begin{array}{c} \diagup \\ \diagdown \end{array} \right) \longleftrightarrow \left( \begin{array}{c} \diagdown \\ \diagup \end{array} \right)$$



and the moves



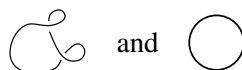
where there are  $k$  edges on the right and  $l$  on the left, for all  $k, l \geq 0$ .

The main problem in knot theory is deciding whether any two link diagrams represent isotopic links. This reduces now to question whether these link diagrams are equal in  $\mathbb{Z}\mathcal{LD}(F)/\mathcal{R}(RI, RII, RIII)$ , where  $\mathcal{LD}(F)$  is the set of all link diagrams in  $F$ . (In classical knot theory  $F = D^2$ , but other surfaces are of interest for us as well.) Notice that this is a version of the problem formulated in Introduction.

The rules

$$r_1: \left. \begin{array}{c} | \\ \circlearrowleft \end{array} \right\} \rightarrow \left. \begin{array}{c} | \\ \circlearrowright \end{array} \right\}, \quad r_2: \left. \begin{array}{c} \diagup \\ \diagdown \end{array} \right\} \rightarrow \left. \begin{array}{c} \diagup \\ \diagdown \end{array} \right\} \left( \right.$$

although terminal are not confluent, since



are both irreducible but not equivalent to each other. However, it is easy to show by the method of Section 3.6 that  $r_1, r_2$  together with

$$r'_1: \left. \begin{array}{c} | \\ \circlearrowleft \end{array} \right\} \rightarrow \left. \begin{array}{c} | \\ \circlearrowright \end{array} \right\}$$

are terminal and confluent. Unfortunately, the rule

$$r_3: \left. \begin{array}{c} \diagup \\ \diagdown \end{array} \right\} \rightarrow \left. \begin{array}{c} \diagdown \\ \diagup \end{array} \right\}$$

is not terminal since it is its own inverse. Furthermore,  $r_1, r'_1, r_2, r_3$  are not confluent.

Nonetheless, inspired by the notion of confluence we ask the following question:

**Question 9.1** For any given surface  $F$ , is there a set  $\mathcal{G}$  of graphs in  $F$  containing  $\mathcal{LD}(F)$  and a finite set of terminal and confluent reduction rules  $S_1, \dots, S_d$ , with coefficients in a ring  $R$  such that the inclusion  $\mathcal{LD}(F) \hookrightarrow \mathcal{G}$  induces a monomorphism

$$\phi: R\mathcal{LD}(F)/\mathcal{R}(RI, RII, RIII) \hookrightarrow R\mathcal{G}/\mathcal{R}(S_1, \dots, S_d).$$

A positive answer to this question would provide an obvious algorithm for distinguishing non-isotopic links in  $I$ -bundles over  $F$ . We conjecture that there does not exist a set of terminal and confluent reduction rules with these properties.

Confluence theory provides an immediate proof of the following statement.

**Theorem 9.2** *Let  $F$  be any surface and  $B \subset \partial F$  a finite set.*

- (1) *unoriented link diagrams ( $A_1$ -webs) in  $F$  are invariant in  $\mathbb{A}_1(F, B, R)$  under moves *RII*, *RIII*, and the first balanced Reidemeister move:*

$$R1b: \left. \begin{array}{c} \circlearrowleft \\ \circlearrowright \end{array} \right\} \rightarrow \left. \begin{array}{c} \circlearrowright \\ \circlearrowleft \end{array} \right\}.$$

- (2)  *$A_2$ -webs are invariant in  $\mathbb{A}_2(F, B, R)$  under oriented 2nd and 3rd Reidemeister moves as well as oriented 1st balanced Reidemeister move and under moving an arc over a vertex.*
- (3) *Similarly,  $B_2$ -webs and  $G_2$ -webs in  $F$  are invariant under all 2nd and 3rd Reidemeister moves (involving both single and double lines) and under moving an arc over a vertex.*

To prove invariance under any of the above relations it is enough to check that reduction rules applied to both sides of that relation yield identical linear diagrams.

**Corollary 9.3** *For orientable  $F$ ,  $\mathbb{A}_1(F, \emptyset, R)$ ,  $\mathbb{B}_2(F, \emptyset, R)$ ,  $\mathcal{G}_2(F, \emptyset, R)$ , provide invariants of framed unoriented links and  $\mathbb{A}_2(F, \emptyset, R)$  provides an invariant of framed oriented links under isotopy in  $F \times I$ .*

For  $F = \mathbb{R}^2$ ,  $B = \emptyset$ , and  $R = \mathbb{Z}[A^{\pm 1}]$ , the module  $\mathbb{A}_1(F, B, R)$  is free on one generator,  $\emptyset$ , and  $[L] \in \mathbb{A}_1(F, B, R) = R$  is the Kauffman bracket of  $L$ , (see Kauffman [34]). For other oriented surfaces,  $\mathbb{A}_1(F, \emptyset, R)$  is isomorphic to the Kauffman bracket skein module of  $F \times I$  (see Hoste–Przytycki [25], Przytycki [50; 52] and Przytycki–Sikora [56]). (For more on Kauffman bracket skein modules see Bullock [7; 8], Bullock–Frohman–Kania-Bartoszyńska [9; 10], Blanchet–Habegger–Masbaum–Vogel [5], Bullock–Przytycki [11], Frohman–Gelca [15], Frohman–Gelca–Lofaro [16], Frohman–Kania-Bartoszyńska [17], Gelca–Sain [20; 19], Gilmer–Harris [21], Hoste–Przytycki [26; 27], Lê [38], Sallenave [59; 58], Sikora [60; 61], and Turaev [64].) Consequently, Corollary 4.1 immediately implies the result [52, Theorem 3.1] of Przytycki:

**Theorem 9.4** *For any ring  $R$  with  $A^{\pm 1} \in R$  and any orientable surface  $F$ , the Kauffman bracket skein module of  $F \times [0, 1]$  is a free  $R$ -module with basis composed of links whose diagrams in  $F$  have no crossings and no contractible components.*

There are versions of this theorem for orientable  $I$ -bundles over non-orientable surfaces and for  $B \neq \emptyset$ . They can be easily proved by the method of confluence as well.

The module  $\mathbb{A}_2(F, \emptyset, R)$  is isomorphic to the  $SU_3$ -skein module of  $F \times I$  introduced in Frohman–Zhong [18] and Sikora [62] (see also Ohtsuki–Yamada [46]). Consequently, Corollary 5.2 implies

**Theorem 9.5** *The  $SU_3$ -skein module of  $F \times I$ ,  $\mathcal{S}_3(F \times I, R)$  (in notation of [62]) is a free  $R$ -module with a basis given by all  $A_2$ -webs in  $F$  with no 0-gons, no true bigons, and no true 4-gons.*

**Problem 9.6** A large number of skein modules is considered in the literature, other than those mentioned above, Andersen–Turaev [1; 2], Gilmer–Zhong [22], Hadji–Morton [23], Hoste–Przytycki [24], Kaiser [32; 30; 31], Kalfgianni–Lin [33], Lieberum [39], Przytycki [48; 47; 49; 51; 54; 53; 55], Przytycki–Tsukamoto [57], Zhong [67] and Zhong–Lu [68]. Can the method of confluence be applied to determine canonical bases of these modules for  $F \times I$ ?

The applications of confluence theory to knot theory discussed so far apply to links in  $I$ -bundles over surfaces. Unfortunately, reduction rules for links in arbitrary 3-manifolds are more difficult to handle. This is illustrated by the Kauffman bracket skein relations:

Let  $\mathcal{L}(M)$  be the set of all framed unoriented links in an orientable 3-manifold  $M$ . Let  $R$  be a ring with a distinguished element  $A^{\pm 1} \in R$  and let

$$S_1: \begin{array}{c} \diagup \\ \diagdown \end{array} \rightarrow A \left( + A^{-1} \begin{array}{c} \diagdown \\ \diagup \end{array} \right) \quad \text{and} \quad S_2: \bigcirc \rightarrow -(A^2 + A^{-2})\emptyset$$

be reduction rules taking place in  $D^3$ . The  $R$ -module  $R\mathcal{L}(M)/\mathcal{R}(S_1, S_2)$  is called *the Kauffman bracket skein module of  $M$* . We leave the proof of the following to the reader:

**Proposition 9.7**  *$S_1, S_2$  are confluent but not terminal.*

## References

- [1] **J E Andersen, V Turaev**, *Higher skein modules*, J. Knot Theory Ramifications 8 (1999) 963–984 MR1723433
- [2] **J E Andersen, V Turaev**, *Higher skein modules II*, from: “Topology, ergodic theory, real algebraic geometry”, Amer. Math. Soc. Transl. Ser. 2 202, Amer. Math. Soc., Providence, RI (2001) 21–30 MR1819178

- [3] **F Baader, T Nipkow**, *Term rewriting and all that*, Cambridge University Press, Cambridge (1998) MR1629216
- [4] **G M Bergman**, *The diamond lemma for ring theory*, Adv. in Math. 29 (1978) 178–218 MR506890
- [5] **C Blanchet, N Habegger, G Masbaum, P Vogel**, *Topological quantum field theories derived from the Kauffman bracket*, Topology 34 (1995) 883–927 MR1362791
- [6] **B Bollobás, O Riordan**, *A polynomial of graphs on surfaces*, Math. Ann. 323 (2002) 81–96 MR1906909
- [7] **D Bullock**, *The  $(2, \infty)$ -skein module of the complement of a  $(2, 2p + 1)$  torus knot*, J. Knot Theory Ramifications 4 (1995) 619–632 MR1361084
- [8] **D Bullock**, *Rings of  $SL_2(\mathbb{C})$ -characters and the Kauffman bracket skein module*, Comment. Math. Helv. 72 (1997) 521–542 MR1600138
- [9] **D Bullock, C Frohman, J Kania-Bartoszyńska**, *Understanding the Kauffman bracket skein module*, J. Knot Theory Ramifications 8 (1999) 265–277 MR1691437
- [10] **D Bullock, C Frohman, J Kania-Bartoszyńska**, *The Yang-Mills measure in the Kauffman bracket skein module*, Comment. Math. Helv. 78 (2003) 1–17 MR1966749
- [11] **D Bullock, J H Przytycki**, *Multiplicative structure of Kauffman bracket skein module quantizations*, Proc. Amer. Math. Soc. 128 (2000) 923–931 MR1625701
- [12] **P-L Curien**, *Categorical combinators, sequential algorithms, and functional programming*, second edition, Progress in Theoretical Computer Science, Birkhäuser, Boston (1993) MR1231971
- [13] **H Ehrig**, *Introduction to the algebraic theory of graph grammars (a survey)*, from: “Graph-grammars and their application to computer science and biology (Internat. Workshop, Bad Honnef, 1978)”, Lecture Notes in Comput. Sci. 73, Springer, Berlin (1979) 1–69 MR565034
- [14] **H Ehrig**, *Tutorial introduction to the algebraic approach of graph grammars*, from: “Graph-grammars and their application to computer science (Warrenton, VA, 1986)”, Lecture Notes in Comput. Sci. 291, Springer, Berlin (1987) 3–14 MR943167
- [15] **C Frohman, R Gelca**, *Skein modules and the noncommutative torus*, Trans. Amer. Math. Soc. 352 (2000) 4877–4888 MR1675190
- [16] **C Frohman, R Gelca, W Lofaro**, *The  $A$ -polynomial from the noncommutative viewpoint*, Trans. Amer. Math. Soc. 354 (2002) 735–747 MR1862565
- [17] **C Frohman, J Kania-Bartoszyńska**, *A quantum obstruction to embedding*, Math. Proc. Cambridge Philos. Soc. 131 (2001) 279–293 MR1857120
- [18] **C Frohman, J Zhong**, *The Yang–Mills measure in the  $SU_3$  skein module* preprint (2004)

- [19] **R Gelca, J Sain**, *The noncommutative  $A$ -ideal of a  $(2, 2p + 1)$ -torus knot determines its Jones polynomial*, *J. Knot Theory Ramifications* 12 (2003) 187–201 MR1967240
- [20] **R Gelca, J Sain**, *The computation of the non-commutative generalization of the  $A$ -polynomial of the figure-eight knot*, *J. Knot Theory Ramifications* 13 (2004) 785–808 MR2088746
- [21] **P M Gilmer, J M Harris**, *On the Kauffman bracket skein module of the quaternionic manifold* arXiv:math.GT/0406152
- [22] **P M Gilmer, J K Zhong**, *The Homflypt skein module of a connected sum of 3-manifolds*, *Algebr. Geom. Topol.* 1 (2001) 605–625 MR1875610
- [23] **R J Hadji, H R Morton**, *A basis for the full Homfly skein of the annulus*, *Math. Proc. Cambridge Philos. Soc.* 141 (2006) 81–100 MR2238644
- [24] **J Hoste, J H Przytycki**, *Homotopy skein modules of orientable 3-manifolds*, *Math. Proc. Cambridge Philos. Soc.* 108 (1990) 475–488 MR1068450
- [25] **J Hoste, J H Przytycki**, *A survey of skein modules of 3-manifolds*, from: “Knots 90 (Osaka, 1990)”, de Gruyter, Berlin (1992) 363–379 MR1177433
- [26] **J Hoste, J H Przytycki**, *The  $(2, \infty)$ -skein module of Whitehead manifolds*, *J. Knot Theory Ramifications* 4 (1995) 411–427 MR1347362
- [27] **J Hoste, J H Przytycki**, *The Kauffman bracket skein module of  $S^1 \times S^2$* , *Math. Z.* 220 (1995) 65–73 MR1347158
- [28] **F Jaeger**, *Confluent reductions of cubic plane maps* (1990), talk at Graph Theory and Combinatorics International Conference, Marseille
- [29] **V F R Jones**, *The Potts model and the symmetric group*, from: “Subfactors (Kyuzeso, 1993)”, World Sci. Publ., River Edge, NJ (1994) 259–267 MR1317365
- [30] **U Kaiser**, *Deformation of string topology into homotopy skein modules*, *Algebr. Geom. Topol.* 3 (2003) 1005–1035 MR2012962
- [31] **U Kaiser**, *Link bordism skein modules*, *Fund. Math.* 184 (2004) 113–134 MR2128047
- [32] **U Kaiser**, *Quantum deformations of fundamental groups of oriented 3-manifolds*, *Trans. Amer. Math. Soc.* 356 (2004) 3869–3880 MR2058509
- [33] **E Kalfagianni, X-S Lin**, *The HOMFLY polynomial for links in rational homology 3-spheres*, *Topology* 38 (1999) 95–115 MR1644083
- [34] **L H Kauffman**, *State models and the Jones polynomial*, *Topology* 26 (1987) 395–407 MR899057
- [35] **G Kuperberg**, *The quantum  $G_2$  link invariant*, *Internat. J. Math.* 5 (1994) 61–85 MR1265145
- [36] **G Kuperberg**, *Spiders for rank 2 Lie algebras*, *Comm. Math. Phys.* 180 (1996) 109–151 MR1403861

- [37] **R Lalement**, *Computation as logic*, Prentice Hall International Series in Computer Science, Prentice Hall International, Englewood Cliffs, NJ (1993) MR1232661 With a foreword by Michel Demazure, Translated from the 1990 French original by John Plaice
- [38] **T T Q Lê**, *The colored Jones polynomial and the  $A$ -polynomial of knots*, Adv. Math. 207 (2006) 782–804 MR2271986
- [39] **J Lieberum**, *Skein modules of links in cylinders over surfaces*, Int. J. Math. Math. Sci. 32 (2002) 515–554 MR1951085
- [40] **P Martin**, *Potts models and related problems in statistical mechanics*, Series on Advances in Statistical Mechanics 5, World Scientific Publishing Co., Teaneck, NJ (1991) MR1103994
- [41] **J C Mitchell**, *Foundations for Programming Languages*, MIT Press (1996)
- [42] **M Nagl**, *A tutorial and bibliographical survey on graph grammars*, from: “Graph-grammars and their application to computer science and biology (Internat. Workshop, Bad Honnef, 1978)”, Lecture Notes in Comput. Sci. 73, Springer, Berlin (1979) 70–126 MR565035
- [43] **M H A Newman**, *On theories with a combinatorial definition of “equivalence.”*, Ann. of Math. (2) 43 (1942) 223–243 MR0007372
- [44] **M J O’Donnell**, *Computing in systems described by equations*, Springer, Berlin (1977) MR0483644 Lecture Notes in Computer Science, Vol. 58
- [45] **E Ohlebusch**, *Advanced topics in term rewriting*, Springer, New York (2002) MR1934138
- [46] **T Ohtsuki, S Yamada**, *Quantum  $SU(3)$  invariant of 3-manifolds via linear skein theory*, J. Knot Theory Ramifications 6 (1997) 373–404 MR1457194
- [47] **J H Przytycki**, *Skein modules of 3-manifolds*, Bull. Polish Acad. Sci. Math. 39 (1991) 91–100 MR1194712
- [48] **J H Przytycki**, *Skein module of links in a handlebody*, from: “Topology ’90 (Columbus, OH, 1990)”, Ohio State Univ. Math. Res. Inst. Publ. 1, de Gruyter, Berlin (1992) 315–342 MR1184418
- [49] **J H Przytycki**, *Vassiliev–Gusarov skein modules of 3-manifolds and criteria for periodicity of knots*, from: “Low-dimensional topology (Knoxville, TN, 1992)”, Conf. Proc. Lecture Notes Geom. Topology, III, Int. Press, Cambridge, MA (1994) 143–162 MR1316179
- [50] **J H Przytycki**, *Algebraic topology based on knots: an introduction*, from: “KNOTS ’96 (Tokyo)”, World Sci. Publ., River Edge, NJ (1997) 279–297 MR1664968
- [51] **J H Przytycki**, *A  $q$ -analogue of the first homology group of a 3-manifold*, from: “Perspectives on quantization (South Hadley, MA, 1996)”, Contemp. Math. 214, Amer. Math. Soc., Providence, RI (1998) 135–144 MR1601241

- [52] **J H Przytycki**, *Fundamentals of Kauffman bracket skein modules*, Kobe J. Math. 16 (1999) 45–66 MR1723531
- [53] **J H Przytycki**, *Homotopy and  $q$ -homotopy skein modules of 3-manifolds: an example in algebra situs*, from: “Knots, braids, and mapping class groups—papers dedicated to Joan S. Birman (New York, 1998)”, AMS/IP Stud. Adv. Math. 24, Amer. Math. Soc., Providence, RI (2001) 143–170 MR1873115
- [54] **J H Przytycki**, *Skein module deformations of elementary moves on links*, from: “Invariants of knots and 3-manifolds (Kyoto, 2001)”, Geom. Topol. Monogr. 4, Geom. Topol. Publ., Coventry (2002) 313–335 MR2048107
- [55] **J H Przytycki**, *From 3-moves to Lagrangian tangles and cubic skein modules*, from: “Advances in topological quantum field theory”, NATO Sci. Ser. II Math. Phys. Chem. 179, Kluwer Acad. Publ., Dordrecht (2004) 71–125 MR2147417
- [56] **J H Przytycki**, **A S Sikora**, *On skein algebras and  $Sl_2(\mathbb{C})$ -character varieties*, Topology 39 (2000) 115–148 MR1710996
- [57] **J H Przytycki**, **T Tsukamoto**, *The fourth skein module and the Montesinos–Nakanishi conjecture for 3-algebraic links*, J. Knot Theory Ramifications 10 (2001) 959–982 MR1867103
- [58] **P Sallenave**, *Structure of the Kauffman bracket skein algebra of  $T^2 \times I$* , J. Knot Theory Ramifications 8 (1999) 367–372 MR1691417
- [59] **P Sallenave**, *On the Kauffman bracket skein algebra of parallelized surfaces*, Ann. Sci. École Norm. Sup. (4) 33 (2000) 593–610 MR1834496
- [60] **A S Sikora**, *Skein modules and TQFT*, from: “Knots in Hellas ’98 (Delphi)”, Ser. Knots Everything 24, World Sci. Publ., River Edge, NJ (2000) 436–439 MR1865721
- [61] **A S Sikora**, *Skein modules at the 4th roots of unity*, J. Knot Theory Ramifications 13 (2004) 571–585 MR2080123
- [62] **A S Sikora**, *Skein theory for  $SU(n)$ -quantum invariants*, Algebr. Geom. Topol. 5 (2005) 865–897 MR2171796
- [63] **C C Sims**, *Computation with finitely presented groups*, Encyclopedia of Mathematics and its Applications 48, Cambridge University Press, Cambridge (1994) MR1267733
- [64] **V G Turaev**, *The Conway and Kauffman modules of a solid torus*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 167 (1988) 79–89, 190 MR964255
- [65] **B W Westbury**, *Invariant tensors for the spin representation of  $\mathfrak{so}(7)$* , Math. Proc. Camb. Phil. Soc. (to appear)
- [66] **D N Yetter**, *On graph invariants given by linear recurrence relations*, J. Combin. Theory Ser. B 48 (1990) 6–18 MR1047550
- [67] **J K Zhong**, *The Kauffman skein module of the connected sum of 3-manifolds* arXiv: math.GT/0205131

- [68] **JK Zhong, B Lu**, *On the Kauffman skein modules*, Manuscripta Math. 109 (2002) 29–47 MR1931206

*Department of Mathematics, SUNY Buffalo  
Buffalo NY 14260, USA*

*Mathematics Institute, University of Warwick  
Coventry, CV4 7AL, UK*

asikora@buffalo.edu, Bruce.Westbury@warwick.ac.uk

<http://www.math.buffalo.edu/~asikora/>, <http://www.maths.warwick.ac.uk/~bww/>

Received: 9 October 2006