NEW ZEALAND JOURNAL OF MATHEMATICS
Volume 52 (2021), 1–107
https://doi.org/10.53733/172

PLANAR ALGEBRAS, I

V. F. R. Jones†

Abstract. We introduce a notion of planar algebra, the simplest example of which is a vector space of tensors, closed under planar contractions. A planar algebra with suitable positivity properties produces a finite index subfactor of a II_1 factor, and vice versa.

0. Introduction

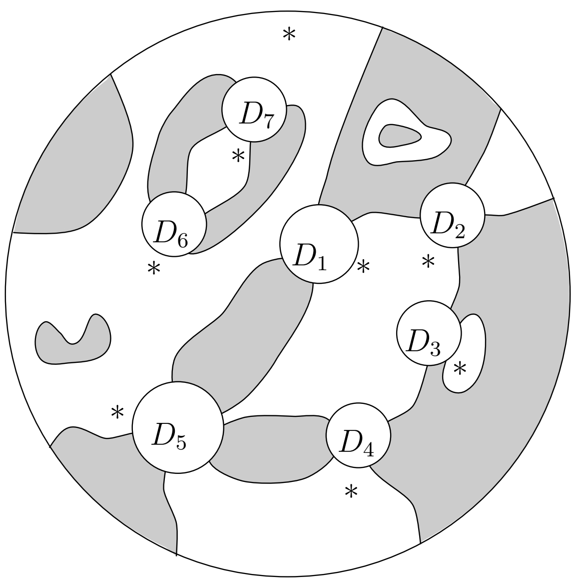
At first glance there is nothing planar about a subfactor. A factor M is a unital $*$-algebra of bounded linear operators on a Hilbert space, with trivial center and closed in the topology of pointwise convergence. The factor M is of type II_1 if it admits a (normalized) trace, a linear function tr: $M \to \mathbb{C}$ with $\mathrm{tr}(ab) = \mathrm{tr}(ba)$ and $\mathrm{tr}(1) = 1$. In [J1] we defined the notion of *index* $[M : N]$ for II_1 factors $N \subset M$. The most surprising result of [J1] was that $[M : N]$ is "quantized" — to be precise, if $[M : N] < 4$ there is an integer $n \geq 3$ with $[M : N] = 4\cos^2 \pi/n$. This led to a surge of interest in subfactors and the major theorems of Pimsner, Popa and Ocneanu ([PP],[Po1],[O1]). These results turn around a "standard invariant" for finite index subfactors, also known variously as the "tower of relative commutants", the "paragroup", or the "λ-lattice". In favorable cases the standard invariant allows one to reconstruct the subfactor, and both the paragroup and λ-lattice approaches give complete axiomatizations of the standard invariant. In this paper we give, among other things, yet another axiomatization which has the advantage of revealing an underlying *planar* structure not apparent in other approaches. It also places the standard invariant in a larger mathematical context. In particular we give a rigorous justification for pictorial proofs of subfactor theorems. Non-trivial results have already been obtained from such arguments in [BJ1]. The standard invariant is sufficiently rich to justify several axiomatizations — it has led to the discovery of invariants in knot theory ([J2]), 3-manifolds ([TV]) and combinatorics ([NJ]),

Research supported in part by NSF Grant DMS93–22675, the Marsden fund UOA520, and the Swiss National Science Foundation.
The editors thank Professor Martha "Wendy" Jones for agreeing to the publication of this paper which has been on arXiv at https://arxiv.org/abs/math/9909027 since 1999.
The editors also thank Dr Tsukasa Yashiro for locating and making available his copy of the manuscript, especially of the figures.

and is of considerable interest in conformal and algebraic quantum field theory ([Wa],[FRS],[Lo]).

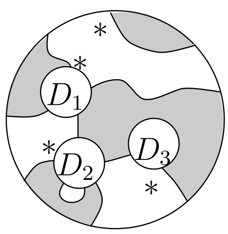
Let us now say exactly what we mean by a planar algebra. The best language to use is that of operads ([Ma]). We define the *planar operad*, each element of which determines a multilinear operation on the standard invariant.

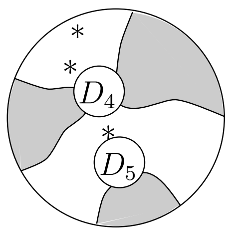
A planar k-tangle will consist of the unit disc D ($= D_0$) in $\mathbb{C}$ together with a finite (possibly empty) set of disjoint subdiscs $D_1, D_2, \ldots, D_n$ in the interior of D. Each disc D_i, $i \geq 0$, will have an even number $2k_i \geq 0$ of marked points on its boundary (with $k = k_0$). Inside D there is also a finite set of disjoint smoothly embedded curves called *strings* which are either closed curves or whose boundaries are marked points of the D_i's. Each marked point is the boundary point of some string, which meets the boundary of the corresponding disc transversally. The strings all lie in the complement of the interiors $\overset{\circ}{D}_i$ of the D_i, $i > 0$. The connected components of the complement of the strings in $\overset{\circ}{D} \setminus \bigcup_{i=1}^{n} D_i$ are called regions and are shaded black and white so that regions whose closures meet have different shadings. The shading is part of the data of the tangle, as is the choice, at every D_i, $i \geq 0$, of a white region whose closure meets that disc. The case $k = 0$ is exceptional - there are two kinds of 0-tangle, according to whether the region near the boundary is shaded black or white. An example of a planar 4-tangle, where the chosen white regions are marked with a $*$ close to their respective discs, is given below.

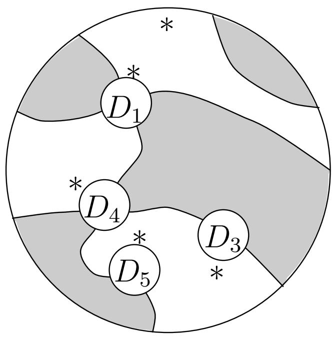


The *planar operad* $\mathbb{P}$ is the set of all orientation-preserving diffeomorphism classes of planar k tangles, k being arbitrary. The diffeomorphisms preserve the boundary of D but may move the D_i's, $i > 1$.

Given a planar k tangle T, a k'-tangle S, and a disk D_i of T with $k_i = k'$ we define the k tangle $T \circ_i S$ by isotoping S so that its boundary, together with the marked points, coincides with that of D_i, and the chosen white regions for D_i (in T) and S share a boundary segment. The strings may then be joined at the boundary of D_i and smoothed. The boundary of D_i is then removed to obtain the tangle $T \circ_i S$ whose diffeomorphism class clearly depends only on those of T and S. This gives $\mathbb{P}$ the structure of a coloured operad, where each D_i for $i > 0$ is assigned the colour k_i and composition is only allowed when the colours match. There are two distinct colours for $k = 0$ according to the shading near the boundary. The D_i's for $i \geq 1$ are to be thought of as inputs and D_0 is the output. (In the usual definition of an operad the inputs are labelled and the symmetric group S_n acts on them. Because of the colours, S_n is replaced by $S_{n_1} \times S_{n_2} \times \ldots \times S_{n_p}$ where n_j is the number of internal discs coloured j. Axioms for such a coloured operad could be given along the lines of [Ma] but we do not need them since we have a concrete example.) The picture below exhibits the composition

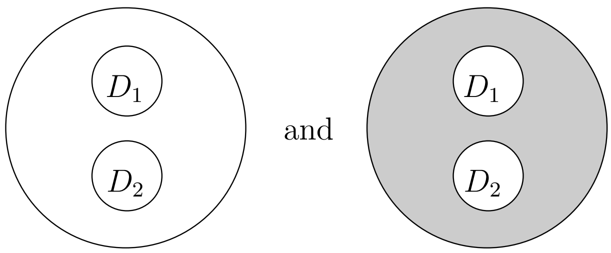
$T =$ 

$S =$ 

$T \circ_2 S$ 

The most general notion of a planar algebra that we will contemplate is that of an algebra over $\mathbb{P}$ in the sense of [Ma]. That is to say, first of all, a disjoint union V_k of vector spaces for $k > 0$ and two vector spaces V_0^{white} and V_0^{black} (which we will call P_0 and $P_{1,1}$ later on). Linear maps between tensor powers of these vector spaces form a coloured operad Hom in the obvious way under composition of maps and the planar algebra structure on the V's is given by a morphism of coloured operads from $\mathbb{P}$ to Hom. In practice this means that, to a k-tangle T in $\mathbb{P}$ there is a linear map $Z(T) : \bigotimes_{i=1}^{n} V_{k_i} \to V_k$ such that $Z(T \circ_i S) = Z(T) \circ_i Z(S)$ where the $\circ_i$ on the right-hand side is composition of linear maps in Hom .

Note that the vector spaces V_0^{white} and V_0^{black} may be different. This is the case for the "spin models" of §3. Both these V_0's become commutative associative algebras using the tangles



To handle tangles with no internal discs we decree that the tensor product over the empty set be the field K and identify $\mathrm{Hom}(K, V_k)$ with V_k so that each V_k will contain a privileged subset which is $Z(\{k\text{-tangles with no internal discs}\})$. This is the "unital" structure (see [Ma]).

One may want to impose various conditions such as $\dim(V_k) < \infty$ for all k. The condition $\dim(V_0^{\text{white}}) = 1 = (\dim(V_0^{\text{black}})$ is significant and we impose it in our formal definition of planar algebra (as opposed to general planar algebra) later on. It implies that there is a unique way to identify each V_0 with K as algebras, and $Z(\bigcirc) = 1 = Z(\bigcirc)$. There are thus also two scalars associated to a planar algebra, $\delta_1 = Z(\bigcirc)$ and $\delta_2 = Z(\bigcirc)$ (the inner circles are strings, not discs!). It follows that Z is multiplicative on connected components, i.e., if a part of a tangle T can be surrounded by a disc so that $T = T' \circ_i S$ for a tangle T' and a 0-tangle S, then $Z(T) = Z(S)Z(T')$ where $Z(S)$ is a multilinear map into the field K.

Two simple examples serve as the keys to understanding the notion of a planar algebra. The first is the Temperley-Lieb algebra TL, some vestige of which is present in every planar algebra. The vector spaces TL_k are:

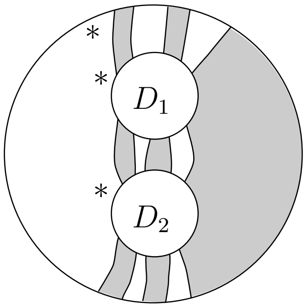
$$TL_0^{\text{black}} \simeq TL_0^{\text{white}} \simeq K$$

and TL_k is the vector space whose basis is the set of diffeomorphism classes of connected planar k-tangles with no internal discs, of which there are $\frac{1}{k+1}\binom{2k}{k}$. The action of a planar k-tangle on TL is almost obvious — when one fills the internal discs of a tangle with basis elements of TL one obtains another basis element, except for some simple closed curves. Each closed curve counts a multiplicative factor of δ and then is removed. It is easily verified that this defines an action of $\mathbb{P}$ on TL. As we have observed, any planar algebra contains elements corresponding to the TL basis. They are not necessarily linearly independent. See [GHJ] and §2.1.

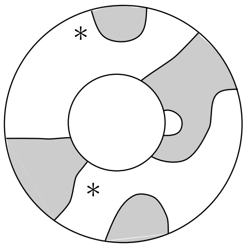
The second key example of a planar algebra is given by tensors. We think of a tensor as an object which yields a number each time its indices are specified. Let V_k be the vector space of tensors with $2k$ indices. An element of $\mathbb{P}$ gives a scheme for contracting tensors, once a tensor is assigned to each internal disc. The indices lie on the strings and are locally constant thereon. The boundary indices are fixed and are the indices of the output tensor. All indices on strings not touching D are summed over and one contracts by taking, for a given set of indices, the product of the values of the tensors in the internal discs. One recognizes the partition function of a statistical mechanical model ([Ba]), the boundary index values being the boundary conditions and the tensor values being the Boltzmann weights. This diagrammatic contraction calculus for tensors is well known ([Pe]) but here we are only considering *planar* contraction systems. If the whole planar algebra of all tensors were the only example this subject would be of no interest, but in fact there is a huge family of planar subalgebras — vector spaces of tensors closed under planar contractions — rich enough to contain the theory of finitely generated groups and their Cayley graphs. See §2.7.

The definition of planar algebra we give in §1 is not the operadic one. When the planar algebra structure first revealed itself, the V_k's already had an associative algebra structure coming from the von Neumann algebra context. Thus our definition is in terms of a universal planar algebra on some set of generators (labels) which can be combined in arbitrary planar fashion. The discs we have used above become boxes in section one, reflecting the specific algebra structure we began with. The equivalence of the two definitions is complete by Proposition 1.20. The main ingredient of the equivalence is that the planar operad $\mathbb{P}$ is generated by the Temperley-Lieb algebra and tangles of two kinds:

(1) Multiplication, which is the following tangle (illustrated for $k = 5$)



(2) Annular tangles; ones with only one internal disc, e.g.,



The universal planar algebra is useful for constructing planar algebras and restriction to the two generating tangles sometimes makes it shorter to check that a given structure is a planar algebra. The algebra structure we began with corresponds of course to the multiplication tangle given above.

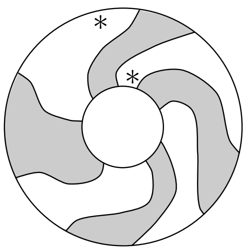
The original algebra structure has been studied in some detail (see §3.1) but it should be quite clear that the operad provides a vast family of algebra structures on a planar algebra which we have only just begun to appreciate. For instance, the annular tangles above form an algebra over which all the V_k's in a planar algebra are modules. This structure alone seems quite rich ([GL]) and we exploit it just a little to get information on principal graphs of subfactors in 4.2.11. We have obtained more sophisticated results in terms of generating functions which we will present in a future paper.

We present several examples of planar algebras in §2, but it is the connection with subfactors that has been our main motivation and guide for this work. The two leading theorems occur in §4. The first one shows how to obtain a planar algebra from a finite index subfactor $N \subset M$. The vector space V_k is the set of N-central vectors in the $N-N$ bimodule $M_{k-1} = M \otimes_N M \otimes_N \cdots \otimes_N M$ (k copies of M), which, unlike M_{k-1} itself, is finite dimensional. The planar algebra structure on these V_k's is obtained by a method reminiscent of topological quantum field theory.

Given a planar k-tangle T whose internal discs are labelled by elements of the V_j's, we have to show how to construct an element of V_k, associated with the boundary of T, in a natural way. One starts with a very small circle (the "bubble") in the distinguished white region of T, tangent to the boundary of D. We then allow this circle to bubble out until it gets to the boundary. On its way the bubble will have to cross strings of the tangle and envelop internal discs. As it does so it acquires shaded intervals which are its intersections with the shaded regions of T. Each time the bubble envelops an internal disc D_i, it acquires k_i such shaded intervals and, since an element of V_{k_i} is a tensor in $\otimes_N^{k_i} M$, we assign elements of M to the shaded region according to this tensor. There are also rules for assigning and contracting tensors as the bubble crosses strings of the tangle. At the end we have an element of $\otimes_N^k M$ assigned to the boundary. This is the action of the operad element on the vectors in V_{k_i}. Once the element of $\otimes_N^k M$ has been constructed and shown to be invariant under diffeomorphisms, the formal operadic properties are immediate.

One could try to carry out this procedure for an arbitrary inclusion $A \subset B$ of rings, but there are a few obstructions involved in showing that our bubbling process is well defined. Finite index (extremal) subfactors have all the special properties required, though there are surely other families of subrings for which the procedure is possible.

The following tangle:



defines a rotation of period k ($2k$ boundary points) so it is a consequence of the planar algebra structure that the rotation $x_1 \otimes x_2 \otimes \cdots \otimes x_k \mapsto x_2 \otimes x_3 \otimes \cdots \otimes x_k \otimes x_1$, which makes no sense on M_{k-1}, is well defined on N-central vectors and has period k. This result is in fact an essential technical ingredient of the proof of Theorem 4.2.1.

Note that we seem to have avoided the use of correspondences in the sense of Connes ([Co1]) by working in the *purely algebraic* tensor product. But the avoidance of L^2-analysis, though extremely convenient, is a little illusory since the proof of the existence and periodicity of the rotation uses L^2 methods. The

Ocneanu approach ([EK]) uses the L^2 definition and the vector spaces V_k are defined as $Hom_{N,N}(\otimes_N^j M)$ and $Hom_{N,M}(\otimes_N^j M)$ depending on the parity of k. It is no doubt possible to give a direct proof of Theorem 4.2.1 using this definition - this would be the "hom" or cohomology version, our method being the "$\otimes$" or homology method.

To identify the operad structure with the usual algebra structure on M_{k-1} coming from the "basic construction" of [J1], we show that the multiplication tangle above does indeed define the right formula. This, and a few similar details, is suprisingly involved and accounts for some unpleasant looking formulae in §4. Several other subfactor notions, e.g. tensor product, are shown to correspond to their planar algebra counterparts, already abstractly defined in §3. Planar algebras also inspired, in joint work with D. Bisch, a notion of free product. We give the definition here and will explore this notion in a forthcoming paper with Bisch.

The second theorem of §4 shows that one can construct a subfactor from a planar algebra with $*$-structure and suitable "reflection" positivity. It is truly remarkable that the axioms needed by Popa for his construction of subfactors in [Po2] follow so closely the axioms of planar algebra, at least as formulated using boxes and the universal planar algebra. For Popa's construction is quite different from the "usual" one of [J1], [F+], [We1], [We2]. Popa uses an amalgamated free product construction which introduces an unsatisfactory element in the correspondence between planar algebras and subfactors. For although it is true that the standard invariant of Popa's subfactor is indeed the planar algebra from which the subfactor was constructed, it is *not* true that, if one begins with a subfactor $N \subset M$, even hyperfinite, and applies Popa's procedure to the standard invariant, one obtains $N \subset M$ as a result. There are many difficult questions here, the main one of which is to decide when a given planar algebra arises from a subfactor of the Murray-von Neumann hyperfinite type II_1 factor ([MvN]).

There is a criticism that has and should be made of our definition of a planar algebra — that it is too restrictive. By enlarging the class of tangles in the planar operad, say so as to include oriented edges and boundary points, or discs with an odd number of boundary points, one would obtain a notion of planar algebra applicable to more examples. For instance, if the context were the study of group representations our definition would have us studying say, SU(n) by looking at tensor powers of the form $V \otimes \bar{V} \otimes V \otimes \bar{V} \dots$ (where V is the defining representation on $\mathbb{C}^n$) whereas a full categorical treatment would insist on arbitrary tensor products. In fact, more general notions already exist in the literature. Our planar algebras could be formulated as a rather special kind of "spider" in the sense of Kuperberg in [Ku], or one could place them in the context of pivotal and spherical categories ([FY],[BW]), and the theory of C^*-tensor categories even has the ever desirable positivity ([LR],[We3]). Also, in the semisimple case at least, the work in section 3.3 on cabling and reduction shows how to extend our planar diagrams to ones with labelled edges.

But it is the very restrictive nature of our definition of planar algebras that should be its great virtue. We have good reasons for limiting the generality. The most compelling is the equivalence with subfactors, which has been our guiding light. We have tried to introduce as little formalism as possible compatible with exhibiting quite clearly the planar nature of subfactor theory. Thus our intention has been to give pride of place to the pictures. But subfactors are not the only reason for our procedure. By restricting the scope of the theory one hopes to get to the most vital examples as quickly as possible. And we believe that we will see, in some form, *all* the examples in our restricted theory anyway. Thus the Fuss-Catalan algebras of [BJ2] (surely among the most basic planar algebras, whatever one's definition) first appeared with our strict axioms. Yet at the same time, as we show in section 2.5, the HOMFLY polynomial, for which one might have thought oriented strings essential, can be completely captured with our unoriented framework.

It is unlikely that any other restriction of some more general operad is as rich as the one we use here. To see why, note that in the operadic picture, the role of the identity is played by tangles without internal discs -see [Ma]. In our case we get the whole Temperley-Lieb algebra corresponding to the identity whereas any orientation restriction will reduce the size of this "identity". The beautiful structure of the Temperley-Lieb algebra is thus always at our disposal. This leads to the following rather telling reason for looking carefully at our special planar algebras among more general ones: if we introduce the generating function for the dimensions of a planar algebra, $\sum_{n=0}^{\infty} \dim(V_n) z^n$, we shall see that if the planar algebra satisfies reflection positivity, then this power series has non-zero radius of convergence. By contrast, if we take the natural oriented planar algebra structure given by the HOMFLY skein, it is a result of Ocneanu and Wenzl ([We1],[F+]) that there is a positive definite Markov trace on the whole algebra, even though the generating function has zero radius of convergence.

In spite of the previous polemic, it would be foolish to neglect the fact that our planar algebra formalism fits into a more general one. Subfactors can be constructed with arbitrary orientations by the procedure of [We1],[F+] and it should be possible to calculate their planar algebras by planar means.

We end this introduction by discussing three of our motivations for the introduction of planar algebras as we have defined them.

Motivation 1 Kauffman gave his now well-known pictures for the Temperley-Lieb algebra in [Ka1]. In the mid 1980's he asked the author if it was possible to give a pictorial representation of all elements in the tower of algebras of [J1]. We have only developed the planar algebra formalism for the sub-tower of relative commutants, as the all-important rotation is not defined on the whole tower. Otherwise this paper constitutes an answer to Kauffman's question.

Motivation 2 One of the most extraordinary developments in subfactors was the discovery by Haagerup in [Ha] of a subfactor of index $(5+\sqrt{13})/2$, along with the proof that this is the smallest index value, greater than 4, of a finite depth subfactor. As far as we know there is no way to obtain Haagerup's "sporadic" subfactor from the conformal field theory/quantum group methods of [Wa],[We3],[X],[EK]. It is our hope that the planar algebra context will put Haagerup's subfactor in at least one natural family, besides yielding tools for its study that are more general than those of [Ha]. For instance it follows from Haagerup's results that the planar algebra of his subfactor is generated by a single element in V_4 (a "4-box"). The small dimensionality of the planar algebra forces extremely strong conditions on this 4-box. The only two simpler such planar algebras (with reflection positivity) are those of the D_6 subfactor of index $4\cos^2\pi/10$ and the $\tilde{E}_7$ subfactor of index 4. There are analogous planar algebras generated by 2-boxes and 3-boxes. The simplest 2-box case comes from the D_4 subfactor (index 3) and the two simplest 3-box cases from E_6 and $\tilde{E}_6$ (indices $4\cos^2\pi/12$ and 4). Thus we believe there are a handful of planar algebras for each k, generated by a single k-box, satisfying extremely strong relations. Common features among these relations should yield a unified calculus for constructing and manipulating these planar algebras. In this direction we have classified with Bisch in [BJ1] *all* planar algebras generated by a 2-box and tightly restricted in dimension. A result of D.Thurston shows an analogous result for 3-boxes - see section 2.5. The 4-box case has yet to be attempted.

In general one would like to understand all systems of relations on planar algebras that cause the free planar algebra to collapse to finite dimensions. This is out of sight at the moment. Indeed it is know from [BH] that subfactors of index 6 are "wild" in some technical sense, but up to $3+\sqrt{3}$ they appear to be "tame". It would be significant to know for what index value subfactors first become wild.

Motivation 3 Since the earliest days of subfactors it has been known that they can be constructed from certain finite data known as a commuting square (see [GHJ]). A theorem of Ocneanu (see [JS] or [EK]) reduced the problem of calculating the planar algebra component V_k of such a subfactor to the solution of a finite system of linear equations in finitely many unknowns. Unfortunately the number of equations grows exponentially with k and it is unknown at present whether the most simple questions concerning these V_k are solvable in polynomial time or not. On the other hand the planar algebra gives interesting invariants of the original combinatorial data and it was a desire to exploit this information that led us to consider planar algebras. First it was noticed that there is a suggestive planar notation for the linear equations themselves. Then the invariance of the solution space under the action of planar tangles was observed. It then became clear that one should consider other ways of constructing planar algebras from combinatorial data, such as the planar algebra generated by a tensor in the tensor planar algebra.

These ideas were the original motivation for introducing planar algebras. We discuss these matters in more detail in section 2.11 which is no doubt the most important part of this work. The significance of Popa's result on $\lambda-$ lattices became apparent as the definition evolved. Unfortunately we have not yet been able to use planar algebras in a convincing way as a tool in the calculation of the planar algebra for specific commuting squares.

This paper has been written over a period of several years and many people have contributed. In particular I would like to thank Dietmar Bisch, Pierre de la Harpe, Roland Bacher, Sorin Popa, Dylan Thurston, Bina Bhattacharya, Zeph Landau, Adrian Ocneanu, Gib Bogle and Richard Borcherds. Deborah Craig for her patience and first-rate typing, and Tsukasa Yashiro for the pictures.

1. The Formalism

Definition 1.1. If k is a non-negative integer, the *standard k-box*, $\mathcal{B}_k$, is $\{(x,y)\in\mathbb{R}^2 \mid 0\le x\le k+1,\ 0\le y\le 1\}$, together with the $2k$ marked points, $1=(1,1)$, $2=(2,1)$, $3=(3,1),\dots,k=(k,1)$, $k+1=(k,0)$, $k+2=(k-1,0),\dots,2k=(1,0)$.

Definition 1.2. A *planar network* $\mathcal{N}$ will be a subset of $\mathbb{R}^2$ consisting of the union of a finite set of disjoint images of $\mathcal{B}_k$'s (with k varying) under smooth orientation-preserving diffeomorphisms of $\mathbb{R}^2$, and a finite number of oriented disjoint curves, smoothly embedded, which may be closed (i.e. isotopic to circles), but if not their endpoints coincide with marked points of the boxes. Otherwise the curves are disjoint from the boxes. All the marked points are endpoints of curves, which meet the boxes transversally. The orientations of the curves must satisfy the following two conditions.

a) A curve meeting a box at an odd marked point must exit the box at that point.
b) The connected components of $\mathbb{R}^2 \backslash \mathcal{N}$ may be oriented in such a way that the orientation of a curve coincides with the orientation induced as part of the boundary of a connected component.

Remark. Planar networks are of two kinds according to the orientation of the unbounded region.

Let L_i, $i = 1, 2, \ldots$ be sets and $L = \coprod_i L_i$ be their disjoint union. L will be called the set of "labels".

Definition 1.3. A *labelled planar network* (on L) will be a planar network together with a function from its k-boxes to L_k, for all k with $L_k \neq \emptyset$.

If the labelling set consists of asymmetric letters, we may represent the labelling function diagrammatically by placing the corresponding letter in its box, with the understanding that the first marked point is at the top left. This allows us to ignore the orientations on the edges and the specification of the marked points. In Fig. 1.4 we give an example of a labelled planar network with $L_1 = \{P\}$, $L_2 = \{R\}$, $L_3 = \{Q\}$. Here the unbounded region is positively oriented and, in order to make the conventions quite clear, we have explicitly oriented the edges and numbered the marked points of the one 3-boxed labelled Q.

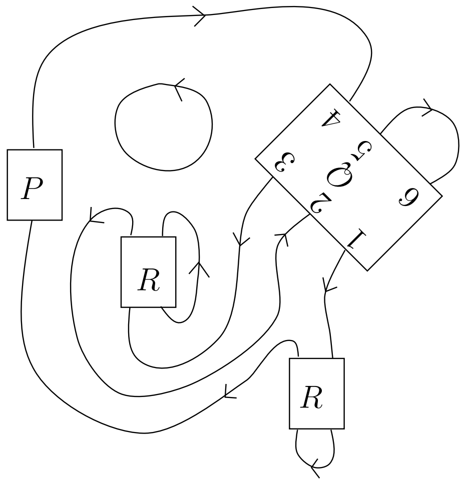


Figure 1.4

Note that the same picture as in Fig. 1.4, but with an R upside down, would be a different labelled planar tangle since the marked points would be different. With or without labels, it is only necessary to say which distinguished boundary point is first.

Remark 1.5. By shrinking each k-box to a point as in Fig. 1.6 one obtains from a planar network a system of immersed curves with transversal multiple point singularities.

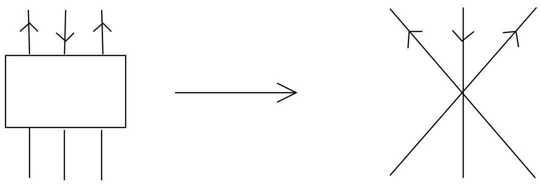


Figure 1.6

Cusps can also be handled by labelled 1-boxes. To reverse the procedure requires a choice of incoming curve at each multiple point but we see that our object is similar to that of Arnold in [A]. In particular, in what follows we will construct a huge supply of invariants for systems of immersed curves. It remains to be seen whether these invariants are of interest in singularity theory, and whether Arnold's invariants may be used to construct planar algebras with the special properties we shall describe.

Definition 1.7. A planar k-tangle T (for $k = 0, 1, 2, \ldots$) is the intersection of a planar network $\mathcal{N}$ with the standard k-box $\mathcal{B}_k$, with the condition that the boundary of $\mathcal{B}_k$ meets $\mathcal{N}$ transversally precisely in the set of marked points of $\mathcal{B}_k$, which are points on the curves of $\mathcal{N}$ other than endpoints. The orientation induced by $\mathcal{N}$ on a neighborhood of (0,0) is required to be positive. A labelled planar k-tangle is defined in the obvious way.

The connected curves in a tangle T will be called the *strings* of T.

The set of smooth isotopy classes of labelled planar k-tangles, with isotopies being the identity on the boundary of $\mathcal{B}_k$, is denoted $T_k(L)$.

Note. $T_0(L)$ is naturally identified with the set of planar isotopy classes of labelled networks with unbounded region positively oriented.

Definition 1.8. The associative algebra $\mathcal{P}_k(L)$ over the field K is the vector space having $T_k(L)$ as basis, with multiplication defined as follows. If $T_1, T_2 \in T_k(L)$, let $\tilde{T}_2$ be T_2 translated in the negative y direction by one unit. After isotopy if necessary we may suppose that the union of the curves in T_1 and $\tilde{T}_2$ define smooth curves. Remove $\{(x, 0) \mid 0 \leq x \leq k+1,\ x \notin \mathbb{Z}\}$ from $T_1 \cup \tilde{T}_2$ and finally rescale by multiplying the y-coordinates $\frac{1}{2}$, then adding $\frac{1}{2}$. The resulting isotopy class of labelled planar k-tangles is $T_1 T_2$. See Figure 1.9 for an example.

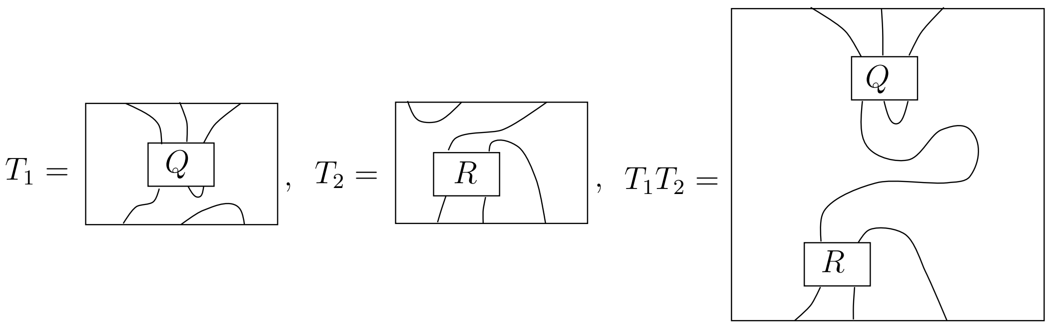


Figure 1.9

Remark. The algebra $\mathcal{P}_k(L)$ has an obvious unit and embeds unitally in $\mathcal{P}_{k+1}(L)$ by adding the line $\{(k+1, t) \mid 0 \leq t \leq 1\}$ to an element of $\mathcal{P}_k(L)$. Since isotopies are the identity on the boundary this gives an injection from the basis of $\mathcal{P}_k(L)$ to that of $\mathcal{P}_{k+1}(L)$.

If there is no source of confusion we will suppress the explicit dependence on L.

Definition 1.10. The *universal planar algebra* $\mathcal{P}(L)$ on L is the filtered algebra given by the union of all the $\mathcal{P}_k$'s ($k = 0, 1, 2, \dots$) with $\mathcal{P}_k$ included in $\mathcal{P}_{k+1}$ as in the preceding remark.

A planar algebra will be basically a filtered quotient of $\mathcal{P}(L)$ for some L, but in order to reflect the planar structure we need to impose a condition of *annular* invariance.

Definition 1.11. The $j-k$ *annulus* $A_{j,k}$ will be the complement of the interior of $\mathcal{B}_j$ in $(j+2)\mathcal{B}_k - (\frac{1}{2}, \frac{1}{2})$. So there are $2j$ marked points on the inner boundary of $\mathcal{A}_{j,k}$ and $2k$ marked points on the outer one. An *annular* $j-k$ *tangle* is the intersection of a planar network $\mathcal{N}$ with $A_{j,k}$ such that the boundary of $A_{j,k}$ meets $\mathcal{N}$ transversally precisely in the set of marked points of $A_{j,k}$, which are points on the curves of $\mathcal{N}$ other than endpoints. The orientation induced by $\mathcal{N}$ in neighborhoods of $(-\frac{1}{2}, -\frac{1}{2})$ and $(0,0)$ are required to be positive. Labeling is as usual.

Warning. The diagram in Fig. 1.11(a) is *not* an annular 2–1 tangle, whereas the diagram in Fig. 1.11 (b) is.

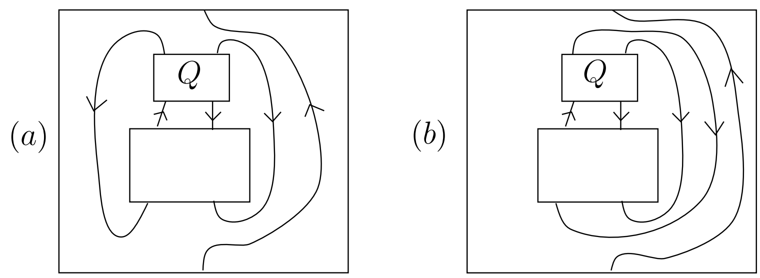


Figure 1.11

The set of all isotopy classes (isotopies being the identity on the boundary) of labelled annular $j-k$ tangles, $\mathcal{A}(L) = \bigcup_{j,k} \mathcal{A}_{j,k}(L)$ forms a category whose objects are the sets of $2j$-marked points of $\mathcal{B}_j$. To compose $A_1 \in \mathcal{A}_{j,k}$ and $A_2 \in \mathcal{A}_{k,\ell}$, rescale and move A_1 so that its outside boundary coincides with the inside boundary of A_2, and the $2k$ boundary points match up. Join the strings of A_1 to those of A_2 at their common boundary and smooth them. Remove that part of the common boundary that is not strings. Finally rescale the whole annulus so that it is the standard one. The result will depend only on the isotopy classes of A_1 and A_2 and defines an element A_2A_1 in $\mathcal{A}(L)$.

Similarly, an $A \in \mathcal{A}_{j,k}(L)$ determines a map $\pi_A : T_j(L) \to T_k(L)$ by surrounding $T \in T_j(L)$ with A and rescaling. Obviously $\pi_A\pi_B = \pi_{AB}$, and the action of $\mathcal{A}(L)$ extends to $\mathcal{P}(L)$ by linearity.

Definition 1.12 A *general planar algebra* will be a filtered algebra $P = \cup_k P_k$, together with a surjective homomorphism of filtered algebras, $\Phi : \mathcal{P}(L) \to P$, for some label set L, $\Phi(\mathcal{P}_k) = P_k$, with ker Φ invariant under $\mathcal{A}(L)$ in the sense that, if $\Phi(x) = 0$ for $x \in \mathcal{P}_j$, and $A \in \mathcal{A}_{j,k}$ then $\Phi(\pi_A(x)) = 0$, we say Φ presents P on L.

Note. Definition 1.12 ensures that $\mathcal{A}(L)$ acts on P via $\pi_A(\Phi(x)) \stackrel{\text{def}}{=} \Phi(\pi_A(x))$. In particular $\mathcal{A}(\emptyset)$ ($\emptyset$ = emptyset) acts on any planar algebra.

The next results show that this action extends multilinearly to planar surfaces with several boundary components.

If T is a planar k-tangle (unlabelled), number its boxes $b_1, b_2, \ldots, b_n$. Then given labelled tangles $T_1, \ldots, T_n$ with T_i having the same number of boundary points as b_i, we may form a labelled planar k-tangle $\pi_T(T_1, T_2, \ldots, T_n)$ by filling each b_i with T_i — by definition b_i is the image under a planar diffeomorphism θ of $\mathcal{B}_j$ (for some j), and T_i is in $\mathcal{B}_j$, so replace b_i with $\theta(T_i)$ and remove the boundary (apart from marked points, smoothing the curves at the marked points). None of this depends on isotopy so the isotopy class of T defines a multilinear map $\pi_T : \mathcal{P}_{j_1} \times \mathcal{P}_{j_2} \times \ldots \times \mathcal{P}_{j_n} \to \mathcal{P}_k$. Though easy, the following result is fundamental and its conclusion is the definition given in the introduction of planar algebras based on the operad defined by unlabeled planar tangles.

Proposition 1.13 *If P is a general planar algebra presented on L, by Φ, π_T defines a multilinear map $P_{j_1} \times P_{j_2} \times \ldots \times P_{j_n} \to P$.*

Proof. It suffices to show that, if all the T_i's but one, say i_0, are fixed in $\mathcal{P}_{j_i}$, then the linear map $\alpha : \mathcal{P}_{j_{i_0}} \to \mathcal{P}_k$, induced by π_T, is zero on ker Φ. By multilinearity the T_i's can be supposed to be isotopy classes of labelled tangles. So fill all the boxes other than the i_0'th box with the T_i. Then we may isotope the resulting picture so that b_{i_0} is the inside box of a $j_{i_0} - k$ annulus. The map α is then the map π_A for some annular tangle A so ker $\Phi \subseteq$ ker π_A by Definition 1.12. □

Proposition 1.14 *Let P be a general planar algebra presented on L by Φ. For each k let S_k be a set and $\alpha : S_k \to P_k$ be a function. Put $S = \coprod_k S_k$. Then there is a* *unique filtered algebra homomorphism* $\Theta_S : \mathcal{P}(S) \to P$ *with* $\ker \Theta_S$ *invariant under* $\mathcal{A}(S)$*, intertwining the* $\mathcal{A}(\emptyset)$ *actions and such that* $\Theta_S(\boxed{R}) = \alpha(R)$ *for* $R \in S$*.*

PROOF. Let T be a tangle in $P(S)$ with boxes $b_1, \ldots, b_n$ and let $f(b_i)$ be the label of b_i. We set $\Theta_S(T) = \pi_T(\alpha(f(b_1)), \alpha(f(b_2)), \ldots, \alpha(f(b_n))$ with π_T as in 1.13. For the homomorphism property, observe that $\pi_{T_1T_2}$ and $\pi_{T_1} \cdot \pi_{T_2}$ are both multilinear maps agreeing on a basis. For the annular invariance of $\ker \Theta_S$, note that Θ_S factors through $\mathcal{P}(L)$, say $\Theta_S = \Phi \circ \theta$, so that $\Theta_S(x) = 0 \Longleftrightarrow \theta(x) \in \ker \Phi$. Moreover, if $A \in \mathcal{A}(S)$, $\theta \circ \pi_A$ is a linear combination of $\pi_{A'}$'s for A' in $\mathcal{A}(L)$. Hence $\Phi(\theta(\pi_A(x)) = 0$ if $\theta(x) \in \ker \Phi$.

Finally we must show that Θ_S is unique. Suppose we are given a tangle $T \in \mathcal{P}(S)$. Then we may isotope T so that all its boxes occur in a vertical stack, as in Figure 1.15.

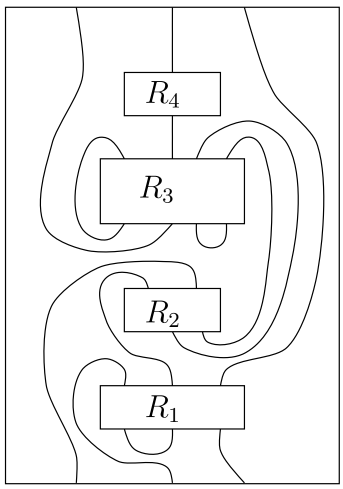


Figure 1.15

In between each box cut horizontally along a level for which there are no critical points for the height function along the curves. Then the tangle becomes a product of single labelled boxes surrounded by $\mathcal{A}(\emptyset)$ elements. By introducing kinks as necessary, as depicted in Figure 1.16, all the surrounded boxes may be taken in $P_k(S)$ for some large fixed k.

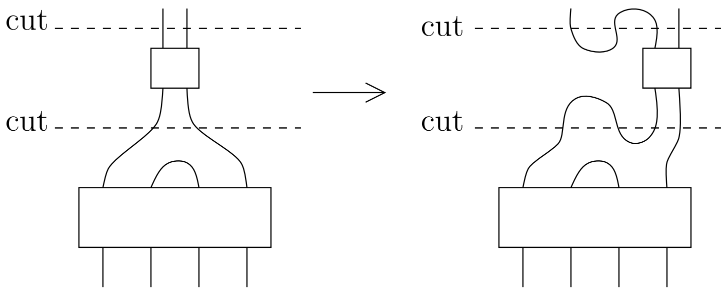


Figure 1.16

Since Θ_S is required to intertwine the $\mathcal{A}(\emptyset)$ action and is an algebra homomorphism, it is determined on all the surrounded boxes by its value on $\{\boxed{R} : R \in S\}$, and their products. The beginning and end of T may involve a change in the value of k, but they are represented by an element of $\mathcal{A}(\emptyset)$ applied to the product of the surrounded boxes. So Θ_S is completely determined on T. $\square$

Definition 1.17. Let P^1, P^2 be general planar algebras presented by Φ_1, Φ_2 on L^1, L^2 respectively. If $\alpha : L^1_k \to P^2_k$, as in 1.14, is such that $\ker\, \Theta_\alpha \supseteq \ker\, \Phi_1$, then the resulting homomorphism of filtered algebras $\Gamma_\alpha : P^1 \to P^2$ is called *planar algebra homomorphism.* A *planar subalgebra* of a general planar algebra is the image of a planar algebra homomorphism. A planar algebra homomorphism that is bijective is called a *planar algebra isomorphism.* Two presentations Φ_1 and Φ_2 of a planar algebra will be considered to define the same planar algebra structure if the identity map is a planar algebra homomorphism.

Remarks. (i) It is obvious that planar algebra homomorphisms intertwine the $\mathcal{A}(\emptyset)$ actions.

(ii) By 1.14, any presentation of a general planar algebra P can be altered to one whose labelling set is the whole algebra itself, defining the same planar algebra structure and such that $\Phi(\boxed{R}) = R$ for all $R \in P$. Thus there is a canonical, if somewhat unexciting, labelling set. We will abuse notation by using the same letter Φ for the extension of a labelling set to all of P. Two presentations defining the same planar algebra structure will define the same extensions to all of P as labelling set.

Proposition 1.18 *Let P be a general planar algebra, and let $C_n \subseteq P_n$ be unital subalgebras invariant under $\mathcal{A}(\emptyset)$ (i.e., $\pi_A(C_j) \subseteq C_k$ for $A \in \mathcal{A}_{j,k}(\emptyset)$). Then $C = \cup C_n$ is a planar subalgebra of P.*

Proof. As a labelling set for C we choose C itself. We have to show that $\Theta_C(\mathcal{P}(\mathcal{C})) \subseteq \mathcal{C}$. But this follows immediately from the argument for the uniqueness of Θ_S in 1.14. (Note that $C_n \subseteq C_{n+1}$ as subalgebras of P_{n+1} is automatic from invariance under $\mathcal{A}(\emptyset)$.) $\square$

The definition of isomorphism was asymmetric. The next result shows that the notion **is** symmetric.

Proposition 1.19 *If $\Gamma_\alpha : P^1 \to P^2$ is an isomorphism of planar algebras, so is $(\Gamma_\alpha)^{-1}$.*

Proof. Define $\alpha^{-1} : L_k^2 \to P_k^1$ by $\alpha^{-1}(R) = (\Gamma_\alpha)^{-1}(\Phi_2(\boxed{R})$. Then $\Gamma_\alpha \circ \Theta_{\alpha^{-1}}$ is a filtered algebra homomorphism intertwining the $\mathcal{A}(\emptyset)$ actions so it equals Φ_1 by 1.14. Thus ker $\Phi_2 \subseteq$ ker $\Theta_{\alpha^{-1}}$ and $\Gamma_{\alpha^{-1}} = (\Gamma_\alpha)^{-1}$. $\square$

The definitions of planar algebra homomorphisms, etc., as above are a little clumsy. The meaning of the following result is that this operadic definition of the introduction would give the same notion as the one we have defined.

Proposition 1.20 *If P_i, Φ_i, L_i for $i = 1, 2$ are as in Definition 1.17, then linear maps $\Gamma : P_k^1 \to P_k^2$ define a planar algebra homomorphism iff*

$$\pi_T(\Gamma(x_1), \Gamma(x_2), \ldots, \Gamma(x_n)) = \Gamma(\pi_T(x_1, x_2, \ldots, x_n))$$

for every unlabelled tangle T as in 1.13.

Proof. Given Γ, define $\alpha : L_1 \to P^2$ by $\alpha(R) = \Gamma(\Phi_1(\boxed{R}))$. Then $\Theta_\alpha = \Gamma \circ \Phi_1$ by the uniqueness criterion of 1.14 (by choosing T appropriately it is clear that Γ is a homomorphism of filtered algebras intertwining $\mathcal{A}(\emptyset)$-actions). On the other hand, a planar algebra isomorphism provides linear maps Γ which satisfy the intertwining condition with π_T. $\square$

Definition 1.21. For each $j, k = 0, 1, 2, \ldots$ with $j \leq k$, $\mathcal{P}_{j,k}(L)$ will be the subalgebra of $\mathcal{P}_k(L)$ spanned by tangles for which all marked points are connected by vertical straight lines except those having x coordinates $j+1$ through k. (Thus $\mathcal{P}_{0,k} = \mathcal{P}_k$.) If B is a general planar algebra, put $P_{j,k} = \Phi(\mathcal{P}_{j,k})$ for some, hence any, presenting map Φ.

Definition 1.22. A *planar algebra* will be a general planar algebra P with dim $P_0 = 1 =$ dim $P_{1,1}$ and $\Phi(\circlearrowright), \Phi(\circlearrowleft)$ both nonzero.

A planar algebra P, with presenting map $\Phi : \mathcal{P}(L) \to P$, defines a planar isotopy invariant of labelled planar networks, $\mathcal{N} \mapsto Z_\Phi(\mathcal{N})$ by $Z_\Phi(\mathcal{N})\mathrm{id} = \Phi(\boxed{\mathcal{N}}) \in P_0$ if the unbounded region of $\mathcal{N}$ is positively oriented (and $\mathcal{N}$ is moved inside $\mathcal{B}_0$ by an isotopy), and $Z_\Phi(\mathcal{N})\mathrm{id} = \Phi(\boxed{\uparrow \mathcal{N}}) \in P_{1,1}$ in the other case ($\mathcal{N}$ has been isotoped into the right half of $\mathcal{B}_1$). The invariant Z is called the *partition function.* It is multiplicative in the following sense.

Proposition 1.23 *Let P be a planar algebra with partition function Z. If T is a labelled tangle containing a planar network $\mathcal{N}$ as a connected component, then*

$$\Phi(T) = Z(\mathcal{N})\Phi(T')$$

where T' is the tangle T from which $\mathcal{N}$ has been removed.

Proof. If we surround $\mathcal{N}$ by a 0-box (after isotopy if necessary) we see that T is just $\boxed{\mathcal{N}}$ to which a $0 - k$ annular tangle has been applied. But $\Phi(\boxed{\mathcal{N}}) = Z(\mathcal{N})\Phi(\boxed{})$, so by annular invariance, $\Phi(T) = Z(\mathcal{N})T$ $\square$

A planar algebra has two scalar parameters, $\delta_1 = Z(\circlearrowleft)$ and $\delta_2 = Z(\circlearrowright)$ which we have supposed to be non-zero.

We present two useful procedures to construct planar algebras. The first is from an invariant and is analogous to the GNS method in operator algebras.

Let Z' be a planar isotopy invariant of labelled planar networks for some labelling set L. Extend Z' to $\mathcal{P}_0(L)$ by linearity. Assume Z' is multiplicative on connected components and that $Z'(\circlearrowright) \neq 0$, $Z'(\circlearrowleft) \neq 0$. For each k let $\mathcal{J}_k = \{x \in \mathcal{P}_k(L) \mid Z'(A(T)) = 0 \ \forall A \in \mathcal{A}_{k,0}\}$. Note that Z' (empty network) $=1$.

Proposition 1.24 (i) *$\mathcal{J}_k$ is a 2-sided ideal of $\mathcal{P}_k(L)$ and $\mathcal{J}_{k+1} \cap \mathcal{P}_k(L) = \mathcal{J}_k$.*

(ii) *Let $P_k = \mathcal{P}_k(L)/\mathcal{J}_k$ and let Φ be the quotient map. Then $P = \cup P_k$ becomes a planar algebra presented by Φ with partition function $Z_\Phi = Z'$.*

(iii) *If $x \in P_k$ then $x = 0$ iff $Z_\Phi(A(x)) = 0 \quad \forall\ A \in \mathcal{A}_{k,0}$.*

Proof. (i) If T_1 and T_2 are tangles in $\mathcal{P}_k(L)$, the map $x \mapsto T_1 x T_2$ is given by an element T of $\mathcal{A}_{k,k}(L)$, and if $A \in \mathcal{A}_{k,0}$ then $Z'(A(Tx)) = Z'(AT)(x)) = 0$ if $x \in \mathcal{J}_k$. Hence $\mathcal{J}_k$ is an ideal.

It is obvious that $\mathcal{J}_k \subset \mathcal{J}_{k+1} \cap \mathcal{P}_k(L)$. So suppose $x \in \mathcal{J}_{k+1} \cap \mathcal{P}_k$. Then for some $y \in \mathcal{P}_k$, $x = \boxed{\boxed{y}\,\updownarrow}$, the orientation of the last straight line depending on the parity of k. We want to show that $y \in \mathcal{J}_k$. Take an $A \in \mathcal{A}_{k,0}$ and form the element $\tilde{A}$ in $\mathcal{A}_{k+1,0}$ which joins the rightmost two points, inside the annulus, close to the inner boundary. Then $\tilde{A}(x)$ will be $A(y)$ with a circle inserted close to the right extremity of $\boxed{y}$. So by multiplicativity, $Z'(\tilde{A}(x)) = Z'(\bigcirc)Z'(A(y))$. Since $Z'(\bigcirc) \neq 0$, $Z'(A(y)) = 0$ and $y \in \mathcal{J}_k$.

(ii) By (i) we have a natural inclusion of P_k in P_{k+1}. Invariance of the $\mathcal{J}_k$'s under $\mathcal{A}$ is immediate. To show that $\dim P_0 = 1 = \dim P_{1,1}$, define maps $U : \mathcal{P}_0 \to K$ (K = the field) and $V : \mathcal{P}_{1,1} \to K$ by linear extensions of $U(\boxed{\mathcal{N}}) = Z'(\mathcal{N})$ and $V(\boxed{\mathcal{N} \uparrow \mathcal{M}}) = Z'(\mathcal{N})Z'(\mathcal{M})$. Observe that $U(\mathcal{J}_0) = 0$ and if $\mathcal{N}_i, \mathcal{M}_i, \lambda_i (\in K)$ satisfy $Z'(\sum_i \lambda_i A(\boxed{\mathcal{N}_i \uparrow \mathcal{M}_i})) = 0$ for all $A \in \mathcal{A}_{1,0}$, then by multiplicativity, $Z'(\bigcirc)(\sum \lambda_i Z(\mathcal{N}_i)Z(\mathcal{M}_i)) = 0$, so that U and V define maps from P_0 and $P_{1,1}$ to K, respectively. In particular both U and V are surjective since $U(\boxed{}) = 1$, $V(\boxed{\uparrow}) = 1$. We need only show injectivity. So take a linear combination $\sum \lambda_i \boxed{\mathcal{N}_i}$ with $\sum \lambda_i Z'(\mathcal{N}_i) = 0$. Then if $A \in \mathcal{A}_{0,0}$, $Z'(\sum \lambda_i \cdot A(\boxed{\mathcal{N}_i}))=0$ by multiplicativity so $\sum \lambda_i \cdot \boxed{\mathcal{N}_i} \in \mathcal{J}_0$. Similarly for $\sum_i \lambda_i A(\boxed{\mathcal{N}_i \uparrow \mathcal{M}_i}) \in P_{1,1}$.

Thus $\dim P_0 = 1 = \dim P_{1,1}$ and by construction, $Z = Z'$.

(iii) This is the definition of $\mathcal{J}_k$ (and $Z_\Phi = Z'$). $\square$

Remark. If one tried to make the construction of 1.24 for an invariant that was not multiplicative, one would rapidly conclude that the resulting algebras all have dimension zero.

Definition 1.25. A planar algebra satisfying condition (iii) of 1.24 will be called non-degenerate.

The second construction procedure is by generators and relations. Given a label set L and a subset $R \subseteq \mathcal{P}(L)$, let $\mathcal{J}_j(R)$ be the linear span of $\bigcup_{\substack{T \in R \\ T \in \mathcal{P}_k(L)}} \mathcal{A}_{k,j}(L)(T)$. It is immediate that $\mathcal{J}_{j+1}(R) \cap \mathcal{P}_j(L) = \mathcal{J}_j(R)$ (just apply an element of $\mathcal{A}(\emptyset)$ to kill off the last string), and $\mathcal{J}_j(R)$ is invariant under $\mathcal{A}(L)$ by construction.

Definition 1.26. With notation as above, set $P_n(L,R) = \dfrac{\mathcal{P}_n(L)}{\mathcal{J}_n(R)}$. Then $P(L,R) = \cup_n P_n(L,R)$ will be called the planar algebra with generators L and relations R.

This method of constructing planar algebras suffers the same drawbacks as constructing groups by generators and relations. It is not clear how big $\mathcal{J}_n(R)$ is inside $\mathcal{P}_n(L)$. It is a very interesting problem to find relation sets R for which $0 < \dim P_n(L,R) < \infty$ for each n. Knot theory provides some examples as we shall see.

Definition 1.27. A planar algebra is called *spherical* if its partition function Z is an invariant of networks on the two-sphere S^2 (obtained from $\mathbb{R}^2$ by adding a point at infinity).

The definition of non-degeneracy of a planar algebra involves all ways of closing a tangle. For a spherical algebra these closures can be arranged in a more familiar way as follows.

Definition 1.28. Let P be a planar algebra with partition function Z. Define two traces tr_L and tr_R on P_k by

Note. For a spherical planar algebra P, $\delta_1 = \delta_2$ and we shall use δ for this quantity. Similarly $\mathrm{Tr}_L = \mathrm{Tr}_R$ and we shall use Tr. If we define $\mathrm{tr}(x) = \frac{1}{\delta^n}\mathrm{Tr}(x)$ for $x \in P_n$ then tr is compatible with the inclusions $P_n \subseteq P_{n+1}$ (and $\mathrm{tr}(1) = 1$), so defines a trace on P itself.

Proposition 1.29 *A spherical planar algebra is nondegenerate iff* Tr *defines a nondegenerate bilinear form on* P_k *for each* k.

PROOF. ($\Leftarrow$) The picture defining Tr is the application of a particular element A of $\mathcal{A}_{k,0}$ to $x \in P_k$.

($\Rightarrow$) It suffices to show that, for any $A \in \mathcal{A}_{k,0}(L)$ there is a $y \in P_k$ such that $\mathrm{Tr}(xy) = Z(A(x))$. By spherical invariance one may arrange $A(x)$ so that the box containing x has no strings to its left. The part of $A(x)$ outside that box can then be isotoped into a k-box which contains the element y. □

Remark 1.30. One of the significant consequences of 1.29 is that, for nondegenerate P, if one can find a finite set of tangles which linearly span P_k, the calculation of dim P_k is reduced to the *finite* problem of calculating the rank of the bilinear form defined on P_k by Tr. Of course this may not be easy!

Corollary 1.31 *A nondegenerate planar algebra is semisimple.*

Positivity

For the rest of this section suppose the field is $\mathbb{R}$ or $\mathbb{C}$.

Suppose we are given an involution $R \to R^*$ on the set of labels L. Then $\mathcal{P}(L)$ becomes a $*$-algebra as follows. If T is a tangle in $T_k(L)$ we reflect the underlying unlabelled tangle in the line $y = \frac{1}{2}$ and reverse all the orientations of the strings. The first boundary point for a box in the reflected unlabelled tangle is the one that was the last boundary point for that box in the original unlabelled tangle.The new tangle T^* is then obtained by assigning the label R^* to a box that was labelled R. This operation is extended sesquilinearly to all of $\mathcal{P}_k(L)$. If Φ presents a general planar algebra, $*$ preserves ker Φ and defines a $*$-algebra structure on $\Phi(\mathcal{P}_k(L))$. The operation $*$ on $T_0(L)$ also gives a well-defined map on isotopy classes of planar networks and we say an invariant Z is sesquilinear if $Z(\mathcal{N}^*) = \overline{Z(\mathcal{N})}$.

Definition 1.32. A $*$-algebra P is called a (general) planar $*$-algebra if it is presented by a Φ on $\mathcal{P}(L)$, L with involution $*$, such that Φ is a $*$-homomorphism.

Note that if P is planar, Z is sesquilinear. Moreover if Z is a sesquilinear multiplicative invariant, the construction of 1.24 yields a planar $*$-algebra. The partition function on a planar algebra will be called *positive* if $\mathrm{tr}_L(x^*x) \geq 0$ for $x \in P_k$, k arbitrary.

Proposition 1.33 *Let P be a planar $*$-algebra with positive partition function Z. The following are equivalent:*

(i) *P is non-degenerate* (Def. 1.24).

(ii) $\mathrm{tr}_R(x^*x) > 0$ *for* $x \neq 0$.

(iii) $\mathrm{tr}_L(x^*x) > 0$ *for* $x \neq 0$.

PROOF. For (ii)$\Leftrightarrow$(iii), argue first that $\delta_1 = \mathrm{tr}_R(\boxed{\uparrow}) > 0$ and $\delta_2 = Z(\circlearrowleft) = \frac{1}{\delta_1}\mathrm{tr}_R(\boxed{\uparrow\downarrow}) > 0$ and then define antiautomorphisms j of P_{2n} by $j(\boxed{R}) = \boxed{R}$, so that $j(x^*) = j(x)^*$ and $\mathrm{tr}_L(j(x)) = \mathrm{tr}_R(x)$.

(ii)⇒(i) is immediate since $\mathrm{tr}_R(x^*x) = \sum \lambda_i A_i(x^*)$ where $A_i \in \mathcal{A}_{k,0}$ is the annular tangle of Figure 1.34,

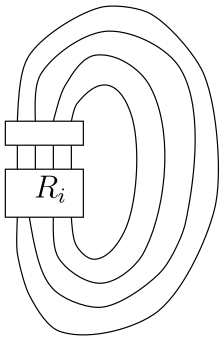


Figure 1.34

writing $x = \Phi(\sum_i \lambda_i \boxed{R_i}) \in B_k$.

(i)⇒(ii) Suppose $x \in P_k$ satisfies $\mathrm{tr}_R(x^*x) = 0$. Then if $A \in \mathcal{A}_{k,0}$, we may isotope $A(x)$ so it looks like Figure 1.35

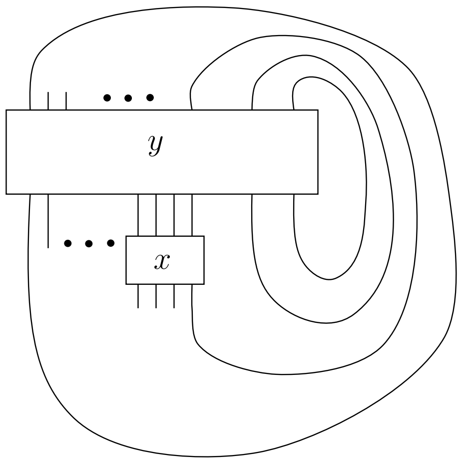


Figure 1.35

where $y \in P_n$, $n \geq k$. Thus $Z(A(x)) = \mathrm{tr}_R(\tilde{x}y)$ where $\tilde{x}$ denotes x with $n-k$ vertical straight lines to the right and left of it. By the Cauchy-Schwartz inequality, $|\mathrm{tr}_R(\tilde{x}y)| \leq \sqrt{\mathrm{tr}_R(\tilde{x}^*\tilde{x})}\sqrt{\mathrm{tr}_R(y^*y)}$, so if $\mathrm{tr}_R(x^*x) = 0$, $Z(A(x)) = 0$. □

We will call a general planar algebra P *finite-dimensional* if dim $P_k < \infty$ for all k.

Corollary 1.36 *If P is a non-degenerate finite-dimensional planar $*$-algebra with positive partition function then P_k is semisimple for all k, so there is a unique norm $\| \ \|$ on P_k making it into a C^*-algebra.*

Proof. Each P_k is semisimple since $\mathrm{tr}(x^*x) > 0$ means there are no nilpotent ideals. The rest is standard. $\square$

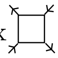
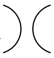
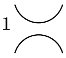
Definition 1.37. We call a planar algebra (over $\mathbb{R}$ or $\mathbb{C}$) a C^*-planar algebra if it satisfies the conditions of corollary 1.34.

2. Examples

Example 2.1: Temperley-Lieb algebra. If δ_1 and δ_2 are two non-zero scalars, one defines $TL(n, \delta_1, \delta_2)$ as being the subspace of $\mathcal{P}_n(\emptyset)$ spanned by the tangles with no closed loops. Defining multiplication on $TL(n, \delta_1, \delta_2)$ by multiplication as in $\mathcal{P}_n(\emptyset)$ except that one multiplies by a factor δ_1 for each loop and δ_2 for each loop , then discarding the loop. Clearly the map from $\mathcal{P}_n(\emptyset)$ to $TL(n, \delta_1, \delta_2)$ given by multiplying by δ_1's or δ_2's then discarding loops, gives a Φ exhibiting $TL(n, \delta_1, \delta_2)$ as a planar algebra. For general values of δ_1 and δ_2, $TL(n)$ is not non-degenerate. An extreme case is $\delta_1 = \delta_2 = 1$ where $(Z(c(T_1 - T_2)) = 0$ for all relevant tangles c, T_1, T_2. In fact the structure of the algebras $TL(n, \delta_1, \delta_2)$ (forgetting Φ), depends only on $\delta_1\delta_2$. To see this, show as in [GHJ] that $TL(n, \delta_1, \delta_2)$ is presented as an algebra by E_i, $i \leq 1, \ldots, n-1$ with $E_i^2 = \delta_1 E_i$ for i odd, $E_i^2 = \delta_2 E_i$ for i even, and $E_i E_{i\pm1} E_i = E_i$ and $E_i E_j = E_j E_i$ for $|i-j| \geq 2$. Then setting $e_i = \frac{1}{\delta_1} E_i$ (i odd), $e_i = \frac{1}{\delta_2} E_i$ (i even), the relations become $e_i^2 = e_i$, $e_i e_{i\pm1} e_i = \frac{1}{\delta_1\delta_2} e_i$, $e_i e_j = e_j e_i$ for $|i-j| \geq 2$. If $\delta_1 = \delta_2 = \delta$, we write $TL(\delta_1, \delta_2) = TL(\delta)$.

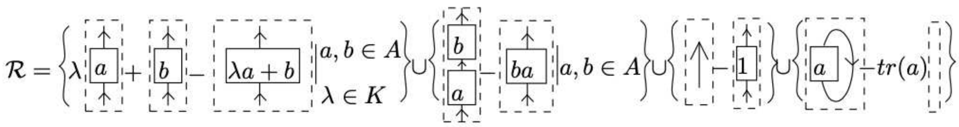
One may also obtain $TL(n)$ via invariants, as a planar algebra on one box, in several ways.

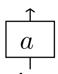
(i) The chromatic polynomial. A planar network $\mathcal{N}$ on $L = L_2$ with $\#(L_2) = 1$ determines a planar graph $G(\mathcal{N})$ by choosing as vertices the positively oriented regions of $\mathbb{R}^2 \backslash \mathcal{N}$ and replacing the 2-boxes by edges joining the corresponding vertices (thus $\rightsquigarrow$ •—•). Fix $Q \in \mathbb{C} - \{0\}$ and let $Z(\mathcal{N})$ = (chromatic polynomial of $G(\mathcal{N})$ as evaluated at Q) $\times f$, where $f = 1$ if the outside region is negatively oriented and $f = Q^{-1}$ if the outside region is positively oriented. To see that P_Z is Temperley-Lieb, define the map $\alpha : \mathcal{P}(L) \to TL(1, Q)$ by $\alpha(\ ) = \ - \ $ (extended by multilinearity to $\mathcal{P}(L)$). It is easy to check that α makes $TL(1, Q)$ a planar algebra on L and the corresponding partition function is Z as above. Thus P_Z is the non-degenerate quotient of $TL(1, Q)$.

(ii) The knot polynomial of [J2]. Given a planar network $\mathcal{N}$ on one 2-box, replace the 2-box  by  to get an unoriented link diagram. Define $Z(\mathcal{N})$ to be the Kauffman bracket ([Ka1]) of this diagram. Sending  to A  $+ A^{-1}$  we see that this defines a map from $\mathcal{P}(L)$ to $TL(-A^2 - A^2)$ with the Temperley-Lieb partition function.

Both (i) and (ii) are generalized by the dichromatic polynomial (see [Tut]).

Example 2.2: Planar algebras on 1-boxes. If A is an associative algebra with identity and a trace functional tr: $A \to K$, $\mathrm{tr}(ab) = \mathrm{tr}(ba)$, $\mathrm{tr}(1) = \delta$, we may form a kind of "wreath product" of A with $TL(n, \delta)$. In terms of generators and relations, we put $L = L_1 = A$ and



One may give a direct construction of this planar algebra using a basis as follows. Choose a basis $\{a_i \mid i \in I\}$ of A with $a_i a_j = \sum c_{ij}^k a_k$ for scalars c_{ij}^k. (Assume $1 \in \{a_i\}$ for convenience.) Let P_n^A be the vector space whose basis is the set of all Temperley-Lieb basis n-tangles together with a function from the strings of the tangle to $\{a_i\}$. Multiply these basis elements as for Temperley-Lieb except that, when a string labelled a_i is joined with one labelled a_j, the result gives a sum over j of c_{ij}^k times the same underlying Temperley-Lieb tangle with the joined string labelled a_k. In the resulting sum of at most $\#(I)^n$ terms, if a closed loop is labelled a_k, remove it and multiply by a factor of $\mathrm{tr}(a_k)$. This gives an associative algebra structure on each P_n^A. It becomes a planar algebra on A in the obvious way with Φ mapping  to a linear combination of strings labelled a_j, the coefficients being those of a in the basis $\{a_i\}$.

If a string in $\mathcal{P}(A)$ has no 1-box on it, it is sent to the same string labelled with 1. One may check that the kernel of Φ is precisely the ideal generated by our relations R, so $P^A = \mathcal{P}(A)/\mathcal{J}(\mathcal{R})$.

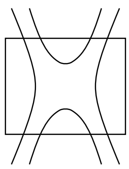
Observe how P_n^A is a sum over Temperley-Lieb basis tangles of tensor powers of A. When $n = 2$ this gives an associative algebra structure on $A \otimes A \oplus A \otimes A$.

Explicitly, write $a \otimes b \oplus 0$ as $a \otimes b$ and $0 \oplus x \otimes y$ as $x \otimes y$. Multiplication is then determined by the rules:

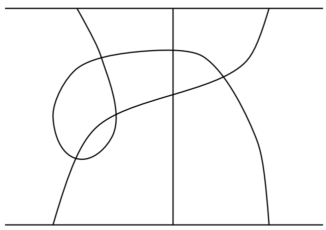
$$\begin{aligned}
(a_1 \otimes b_1)(a_2 \otimes b_2) \quad &= a_1 a_2 \otimes b_1 b_2 \\
(x_1 \otimes y_1)(x_2 \otimes y_2) \quad &= \mathrm{tr}(y_1 x_2) x_1 \otimes y_2 \\
(a \otimes b)(x \otimes y) \quad &= 0 \oplus axb \otimes y \\
(x \otimes y)(a \otimes b) \quad &= 0 \oplus x \otimes bya
\end{aligned}$$

The planar algebra P^A may be degenerate, even when tr on A is non-degenerate and δ is such that $TL(\delta)$ is non-degenerate. We will give more details on the structure of P^A in §3.1.

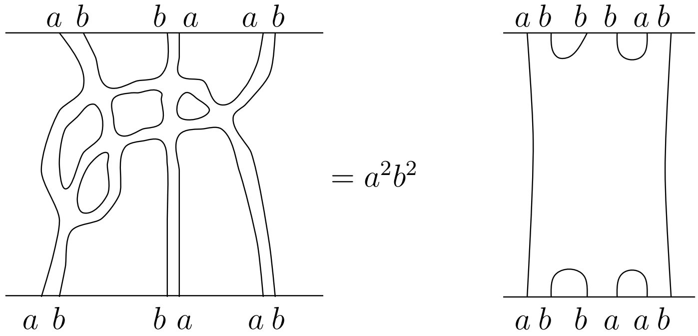
Example 2.3: The Fuss-Catalan algebras (see [BJ2]). If $a_1, a_2, \ldots, a_k \in K - \{0\}$, $FC(n, a_1, \ldots, a_k)$ is the algebra having as basis the Temperley-Lieb diagrams in $TL(nk)$ for which, for each $p = 1, 2, \ldots, k$, the set of all boundary points (counting from the left) indexed by $\{jk + (-1)^j p + (\sin^2 \frac{j\pi}{2})(k + j) \mid j = 0, 1, 2, \ldots (n-1)\}$ are connected among themselves. Assign a colour to each $p = 1, 2, \ldots, k$ so we think of the Temperley-Lieb strings as being coloured. Then multiplication preserves colours so that closed loops will have colours. Removing a closed loop coloured m contributes a multiplicative factor a_m. To see that $FC(n, a_1, \ldots, a_k)$ is a planar algebra, begin with the case $k = 2$. We claim $FC(n, a, b)$ is planar on one 2-box. We draw the box symbolically as



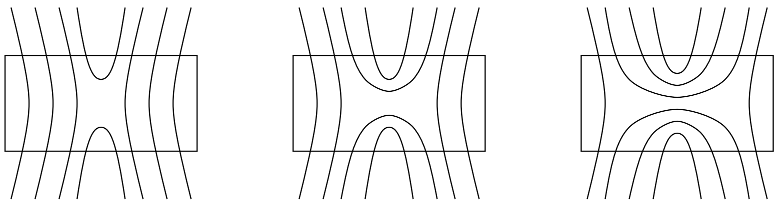
This shows in fact how to define the corresponding $\Phi : \mathcal{P}_n \to FC(n, a_1, a_0)$: double all the strings and replace all the 2-boxes according to the diagram. Thus for instance the network $\mathcal{N}$ below (with boxes shrunk to points, there being only one 2-box),



is sent to $\Phi(\mathcal{N})$ below



It is clear that Φ defines an algebra homomorphism and surjectivity follows from [BJ2]. That ker Φ is annular invariant is straightforward. The general case of $FC(n, a_1, \ldots, a_k)$ is similar. One considers the $k-1$ 2-boxes drawn symbolically as

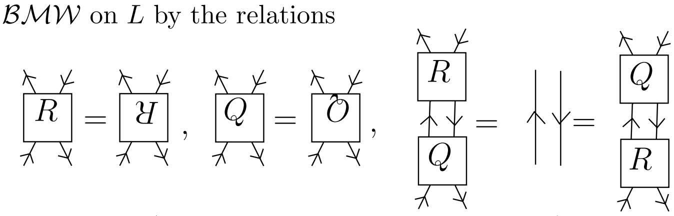


One proceeds as above, replacing the single strings in an $\mathcal{N}$ by k coloured strings. Surjectivity follows from [La].

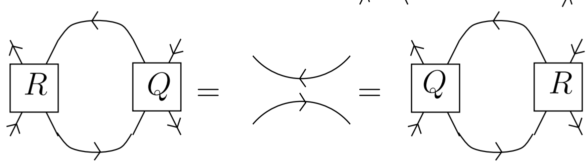
Note that these planar algebras give invariants of systems of immersed curves with generic singularities, and/or planar graphs. The most general such invariant may be obtained by introducing a single 2-box which is a linear combination of the $k-1$ 2-boxes described above. This will generalize the dichromatic polynomial.

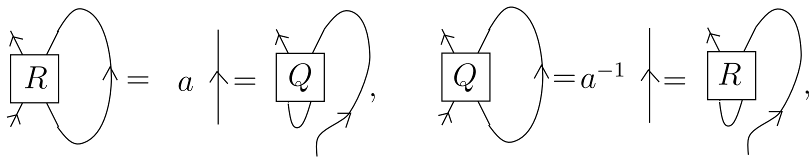
Example 2.4: The BMW algebra. Let $L = L_2 = \{R, Q\}$ and define the planar algebra $\mathcal{BMW}$ on L by the relations

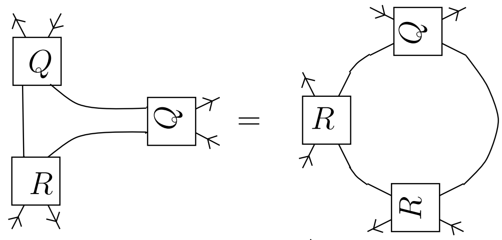
(i)

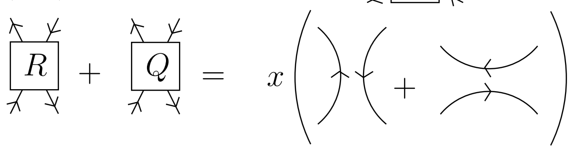


(ii)



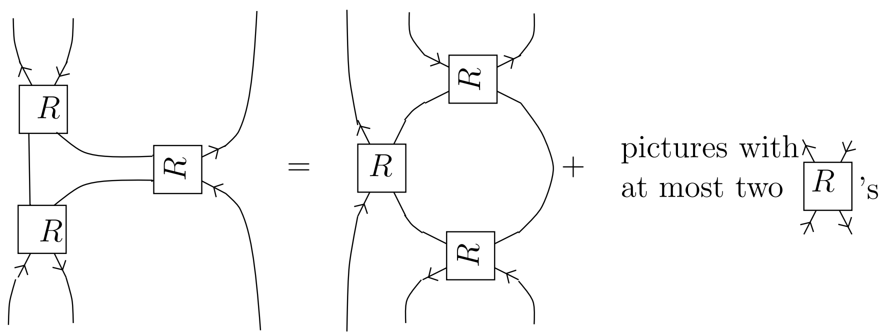
(iii)  $(a \in C - \{0\})$

(iv) 

(v) 

Note that we could use relation (v) to express $\mathcal{BMW}$ using only the one label R, but the relatives would then be more complicated. At this stage $\mathcal{BMW}$ could be zero or infinite dimensional, but we may define a homomorphism from $\mathcal{BMW}$ to the algebra BMW of [BiW],[Mu] by sending $\boxed{R}$ to  and $\boxed{Q}$ to . This homomorphism is obviously surjective and one may use the dimension count of [BiW] to show also that dim $\mathcal{BMW}(n) \leq 1.3.5.\ldots(2n-1)$ so that $\mathcal{BMW} \cong BMW$ as algebras. Thus BMW is planar. It is also connected and the invariant of planar networks is the Kauffman regular-isotopy two-variable polynomial of [Ka2].

Remark. Had we presented BMW on the single 2-box R, the Reidemeister type III move (number (iv) above) would have been



This leads us to consider the general planar algebra B_n with the following three conditions:

(1) B_n is planar on one 2-box.

(2) dim $B_2 = 3$.

(3) dim $B_3 \leq 15$.

If one lists 16 tangles in B_3 then generically any one of them will have to be a linear combination of the other 15. Looking at the 15th and 16th tangles in a listing according to the number of 2-boxes occurring in the tangle, we will generically obtain a type III Reidemeister move, or Yang-Baxter equation, modulo terms with less 2-boxes, as above. It is not hard to show that these conditions force dim $B_n \leq 1.3.5 \cdot \cdots \cdot (2n-1)$ since there are necessarily Reidemeister-like moves of types I and II. Note that $FC(n,a,b)$ satisfies these conditions as well as BMW! For C^*-planar algebras, we have shown with Bisch ([BJ1]) that the only B's with (1) and (2) as above, and dim $B_3 \leq 12$ are the Fuss-Catalan algebras (with one exception, when dim $B_3 = 9$).

Example 2.5: a Hecke-algebra related example. The HOMFLY polynomial ([F+]) is highly sensitive to the orientation of a link and we may not proceed to use it to define a planar algebra as in Example 2.4. In particular, a crossing in the HOMFLY theory is necessarily oriented as . Thus it does *not* yield a 2-box in our planar algebra context. Nevertheless it is possible to use the HOMFLY skein theory to define a planar algebra. We let P_k^H be the usual HOMFLY skein algebra of linear combinations of (3-dimensional) isotopy classes of oriented tangles in the product of the k-box with an interval, with orientations alternating out-in, modulo the HOMFLY skein relation $t \quad - t^{-1} \quad = x\,)\,($ where $t \neq 0$ and x are scalars. Projected onto the k-box, such a tangle could look as in Figure 2.5.1,

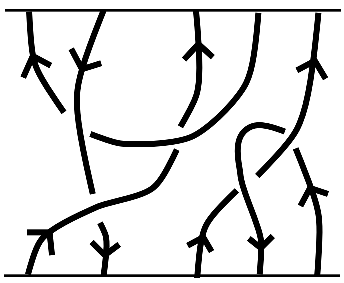
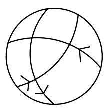


Figure 2.5.1

If we take the labeling set $L_k = P_k^H$, then P^H is a general planar algebra since planar isotopy of projections implies 3-dimensional isotopy (invariance under the annular category is easy). Standard HOMFLY arguments show dim $(P_k^H) \leq k!$ and a specialization could be used to obtain equality. Thus the algebra is planar and the invariant is clearly the HOMFLY polynomial of the oriented link diagram given by a labeled network in P_0^H. If we used the invariant to define the algebra as in §1, we would only obtain the same algebra for generic values of t and x. Note that P_k^H is *not* isomorphic to the Hecke algebra for $k \geq 4$, e.g. P_4^H has an irreducible 4-dimensional representation. In fact P_k^H is, for generic (t, x) and large n, isomorphic to $\mathrm{End}_{SU(n)} \underbrace{(V \otimes \bar{V} \otimes V \otimes \bar{V} \dots)}_{k \text{ vector spaces}}$, where $V = \mathbb{C}^n$, the obvious $SU(n)$-module. This isomorphism is only an algebra isomorphism, not a planar algebra isomorphism.

It is clear that the labeling set for P^H could be reduced to a set of $k!$ isotopy classes of tangles for P_k^H. But in fact a single 3-label suffices as we now show.

Theorem 2.5.2 *Any tangle in the knot-theoretic sense with alternating in and out boundary orientations is isotopic to a tangle with a diagram where all crossings occur in disjoint discs which contain the pattern*



with some non-alternating choice of crossings.

PROOF. We begin with a tangle without boundary, i.e. an oriented link L. Choose a diagram for L and add a parallel double L' of L to the left of L and oppositely oriented, with crossings chosen so that

a) L' is always under L
b) L' itself is an unlink.

An example of the resulting diagram (for the Whitehead Link) is given in Figure 2.5.3.

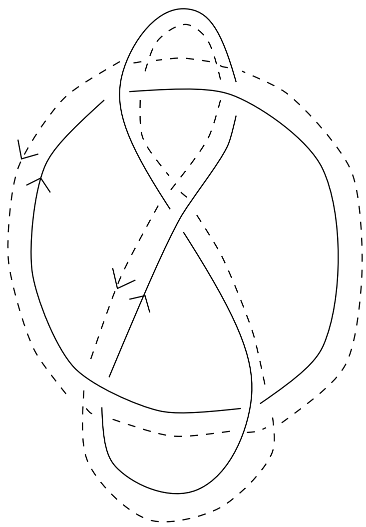
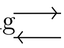
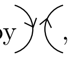
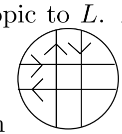
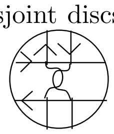


Figure 2.5.3

Now join L to L', component by component, by replacing  by , at some point well away from any crossings. Since L' is an unlink below L, the resulting link is isotopic to L. All the crossings in $L \cup L'$ occur in disjoint discs containing the pattern  which can be isotoped to the pattern  containing two discs of the required form. Alternating patterns can be avoided by keeping the top string on top in this isotopy.

For a tangle with boundary we make the doubling curve follow the boundary, turning right just before it would hit it and right again as it nears the point where the next string exits the tangle, as in Figure 2.5.4.

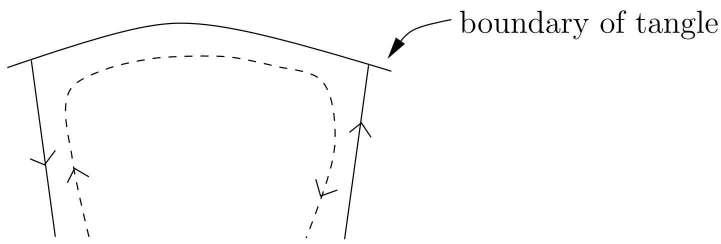
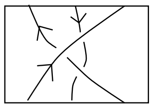


Figure 2.5.4

Join the original tangle to L' one string at a time and proceed as before. $\square$

Corollary 2.5.5 *The planar algebra P^H is generated by the single 3-box* 

PROOF. The HOMFLY relations can be used to go between the various possible choices of crossings in the 3-box of Theorem 2.5.2. □

This corollary was first proved by W. B. R. Lickorish using an argument adapted to the HOMFLY skein. His argument is much more efficient in producing a skein element involving only the above 3-box. The tangles may be chosen *alternating* in Theorem 2.5.2.

Remarks. 1) Another way of stating Theorem 2.5.2 is to say that any tangle can be projected with only simple triple point singularities. One may ask if there are a set of "Reidemeister moves" for such non-generic projections.

2) A related question would be to find a presentation of P^H on the above 3-box.

Discussion 2.5.6 In the remark of Example 2.4 we introduced relations on the planar algebra generated by a 2-box, which force finite dimensionality of all the P_n's. One should explore the possibilities for the planar algebra generated by a single 3-box. The dimension restrictions analogous to the 1,3 $\leq$ 15 values of Example 2.4 are 1,2,6,$\leq$ 24 and we conjecture, somewhat weakly, the following

Conjecture 2.5.7 Let (P, Φ) be a planar algebra with labelling set $L = L_3$, $\#(L_3) = 1$. Suppose $\dim P_n \leq n!$ for $n \leq 4$. Let V be the subspace of $\mathcal{P}_4(L)$ spanned by tangles with at most two labeled 3-boxes, and let $R = V \cap \ker \Phi$ be relations. Then

$$\dim \left(\frac{\mathcal{P}_n(L)}{\mathcal{J}_n(R)} \right) \leq n! \quad \text{forall } n.$$

There is some evidence for the conjecture. It would imply in particular the $n = 0$ case which implies the following result, proved by D. Thurston, about hexagons:

"Consider all graphs with hexagonal faces that may be drawn on S^2 with non-intersecting edges. Let M be the move of Figure 2.5.8 on the set of all such graphs (where the 8 external vertices are connected in an arbitrary way to the rest of the graph).

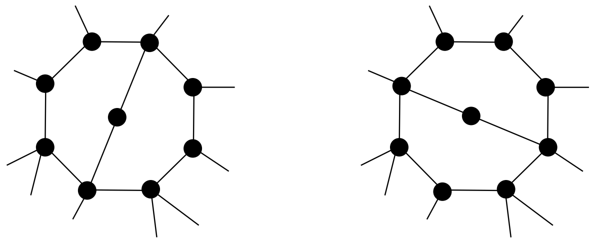


Figure 2.5.8: The move M

Then one may find a finite number of applications of the move M leading to a graph with two adjacent 2-valent vertices."

Example 2.6: Tensors. Let V be a finite dimensional vector space with dual $\hat{V}$. We will define a planar algebra $P^{\otimes} = \cup_k P_k^{\otimes}$ with dim $P_0^{\otimes} = 0$ and $P_k^{\otimes} = \mathrm{End}(V \otimes \hat{V} \otimes V \otimes \hat{V} \otimes \dots)$ where there are k vector spaces in the tensor product. The planar structure on $P^{\otimes}$ can be defined invariantly using the canonical maps $V \otimes \hat{V} \to K$ and $\hat{V} \otimes V \to K$, which are applied to any pair of vector spaces connected by an internal edge in a planar tangle, where V and $\hat{V}$ are associated with the marked points of a k-box in an alternating fashion with V associated to $*$. One could also think of $\mathbb{C} \subset \mathrm{End}(V)$ as a finite factor and use the method of Theorem 4.2.1. It is perhaps easier to understand this structure using a basis $(v_1, v_2, \dots, v_n)$ of V, with corresponding dual basis. An element of $P_k^{\otimes}$ is then the same as a tensor $X_{i_1 i_2 \dots i_k}^{j_1 j_2 \dots j_k}$. The labelling set L_k is $P_k^{\otimes}$ itself and the presenting map $\Phi : \mathcal{P}_k(L) \to P_k^{\otimes}$ is defined by summing ("contracting") over all the internal indices in a labelled planar tangle. The first marked point in a box corresponds to the "j_1" above. To be more precise, one defines a *state* σ of a planar tangle T to be the function from the connected components of the set $S(T)$ of curves in T, $\sigma : S(T) \to \{1, 2, \dots, n\}$, to the basis elements of V. A state defines a set of indices around every box B in T, and since the label associated to B is a tensor, with the appropriate number of indices, to each labelled box, the state σ associates a number, $\sigma(B)$. A state also induces a function $\partial\sigma$ from the marked points on the boundary of T to $\{1, 2, \dots, n\}$. Now we associate a tensor $\Phi(T)$ with T as follows: let $f : \{\text{marked points } (T)\} \to \{1, 2, \dots, n\}$ denote the indices $\begin{pmatrix} j_1 \dots j_k \\ i_1 \dots i_k \end{pmatrix}$

of a tensor in $P_k^{\otimes}$ $(f(p,0) = i_p, f(p,1) = j_p)$. Then

$$\Phi(T)^{j_1 \dots j_k}_{i_1 \dots i_k} = \sum_{\sigma:\partial\sigma = f} \left(\prod_{\substack{B \in \{\text{labelled} \\ \text{boxes of } T\}}} \sigma(B) \right) .$$

An empty sum is zero and an empty product is 1. One easily checks that Φ defines an algebra homomorphism and that ker Φ is invariant under the annular category.

This planar algebra has an obvious $*$-structure. The invariant Z is recognizable as the partition function for the "vertex model" defined by the labelled network, the labels supplying the Boltzmann weights (see [Ba]).

For further discussion we introduce the following notation — consider the indices as a (finite) set Δ. Given a function $\begin{pmatrix} \gamma_1 \dots \gamma_k \\ \delta_1 \dots \delta_k \end{pmatrix}$ from the marked points of a k-box to Δ, we define the corresponding *basic tensor* to be

$$T^{j_1 \dots j_k}_{i_1 \dots i_k} = \begin{cases} 1 & \text{if } \ i_1 = \delta_1, \ i_2 = \delta_2 \ \text{ etc.} \\ & \text{and } \ j_1 = \gamma_1, \ j_2 = \gamma_2 \ \text{ etc.} \\ 0 & \text{otherwise.} \end{cases}$$

If $\mathfrak{S}(\Delta)$ is the free semigroup on Δ, ∂T is then the word $\gamma_1\gamma_2\gamma_3 \dots \gamma_k\delta_k\delta_{k-1} \dots \delta_1$, and we will use the notation $\boxed{\begin{matrix} \gamma_1 \ \gamma_2 \dots \gamma_k \\ \delta_1 \ \delta_2 \dots \delta_k \end{matrix}}$ for this basic tensor.

The planar algebra $P^{\otimes}$ is not terribly interesting by itself (and there seems to be no reason to limit the contractions allowed to planar ones). But one may look for planar subalgebras. One way is to take a set $\{A_i \in P^{\otimes}_{k_i}\}$ and look at the planar subalgebra $P_k(A_i)$ they generate. The calculation of $P_k(A_i)$ as a function of the A_i's can be extremely difficult. While it is easy enough to decide if the A_i's are in the TL subalgebra, we will see in the next example that the question of whether $P_k(A_i) \neq P_k^{\otimes}$ is undecidable, even for $k = 1$!

Note that if the tensors A_i have only 0–1 entries, the partition function will be simply the number of "edge colourings" of the network by n colours with colourings allowed only if they correspond to a non-zero entry of the tensor label at each box.

The next example gives a situation where we can say $P_k(A_i) \neq P_k^{\otimes}$.

Example 2.7: Finitely generated groups. As in 2.6, if Δ is a set, $\mathfrak{S}(\Delta)$ will denote the free semigroup on Δ, and $F(\Delta)$ will denote the free group on Δ. We define the map alt:$\mathfrak{S}(\Delta) \to F(\Delta)$ by alt$(\gamma_1 \dots \gamma_m) = \gamma_1\gamma_2^{-1}\gamma_3\gamma_4^{-1} \dots \gamma_m^{\pm 1}$, (where the + sign occurs only if m is odd, – if m is even). Note that alt is only a homomorphism from the subsemigroup of words of even length.

Now let Γ be a discrete group and Δ a finite set, together with a function $\delta \mapsto \tilde{\delta}$ from Δ to Γ. There is then a natural map $\phi : F(\Delta) \to \Gamma$ defined by $\phi(\delta) = \tilde{\delta}$. Let V be the vector space with basis Δ. Use V and the basis Δ to form the planar algebra $P^{\otimes}$ of §2.6. Recall that, for a basic tensor $T \in P_k^{\otimes}$, ∂T is the element of $\mathfrak{S}(\Delta)$ obtained by reading around the boundary of T. Let $*$ denote the involution on $\mathfrak{S}(\Delta)$ given by writing words backwards.

Definition. $P^{\Gamma,\Delta} = \cup_k P_k^{\Gamma,\Delta}$ is the linear span of all basic tensors T such that $\phi(\mathrm{alt}(\partial T)) = 1$ in Γ.

Proposition 2.7.1 *$P^{\Gamma,\Delta}$ is a planar $*$-subalgebra of $P^{\otimes}$.*

Proof. By 1.18, it suffices to show $P^{\Gamma,\Delta}$ is a unital subalgebra invariant under the annular category $\mathcal{A}(\emptyset)$. If T is a basic tensor in $P_k^{\Gamma,\Delta}$, let $\partial^+ T$ (resp. $\partial^- T$) be the element of $\mathfrak{S}(\Delta)$ obtained by reading along the top of T (resp. the bottom), so $\partial T = \partial^+ T(\partial^- T)^*$. Then for the product $T_1 T_2$ to be non-zero, $\partial^- T_1 = \partial^+ T_2$. In the product $\mathrm{alt}(T_1)\mathrm{alt}(T_2)$, the last letter of $(\partial^- T_1)^*$ is then the same as the first letter of ∂T_2, but with opposite sign. Thus the contribution of $\partial^- T_1$ cancels with that of $\partial^+ T_2$ and $\phi(\mathrm{alt}(T_1)\mathrm{alt}(T_2)) = 1$. So $P^{\Gamma,\Delta}$ is a subalgebra, clearly unital and self-adjoint.

Now consider a typical $\mathcal{A}(\emptyset)$ element C applied to a basic tensor T as in Figure 2.7.2.

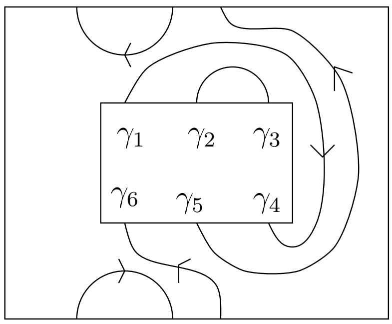


Figure 2.7.2

This tensor is a sum of basic tensors, the sum ranging over all functions from the curves in the diagram to Δ. If R is a particular basic tensor in the sum, notice that the non through-strings in C contribute to $\mathrm{alt}(\partial R)$ in two ways — either they occur in cancelling pairs or, if their beginning and end are separated by the left-hand side of the diagram, they change $\mathrm{alt}(\partial R)$ by conjugation (eliminate all the cancelling pairs first to see this). Thus the conjugacy class in $F(\Delta)$ of $\mathrm{alt}(R)$ (and $\mathrm{alt}(T)$) is not changed by removing all non through-strings. Once this is done, however, the word w around the outer boundary is just an even cyclic permutation of the word w' around the inner boundary, so $\phi(\text{alt } w) = 1 \iff \phi(\text{alt}(w')) = 1$. Thus $\partial^{\Gamma,\Delta}$ is invariant under $\mathcal{A}(\emptyset)$. $\square$

Note that the basic tensors in $P^{\Gamma,\Delta}$ are unchanged if we change $\sim: \Delta \to \Gamma$ by right multiplication by an element of Γ. So we may suppose there is an element e of Δ with $\tilde{e} = 1 \in \Gamma$. To denote this situation we will say simply "$e \in \Delta$".

Let $G \subseteq \Gamma$ (resp. G') $= \{\phi(\text{alt}(\partial^+ T)) \mid T$ a basic tensor in $P_{2k}^{\Gamma,\Delta}$ (resp. $P_k^{\Gamma,\Delta}$), $k \in \mathbb{N}\}$.

Lemma 2.7.3 *G is the subgroup $\langle \tilde{\Delta}\tilde{\Delta}^{-1} \rangle$ of Γ generated by $\tilde{\Delta}\tilde{\Delta}^{-1}$, and if $e \in \Delta$, $G = G'$.*

Proof. The definition of alt implies immediately that $G \subseteq \langle \tilde{\Delta}\tilde{\Delta}^{-1} \rangle$. That $G = G^{-1}$ follows from Figure 2.7.4

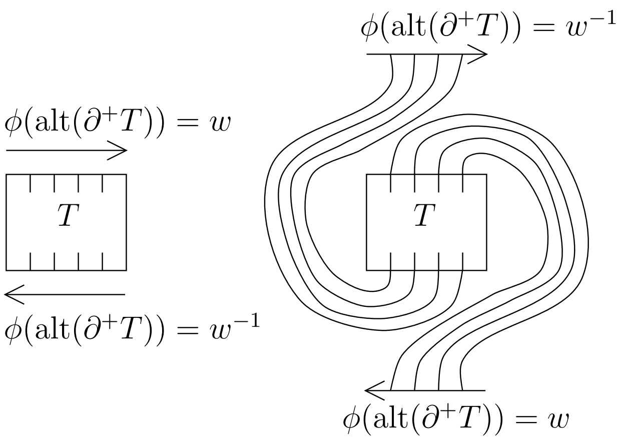


Figure 2.7.4

That G is a group follows from Figure 2.7.5

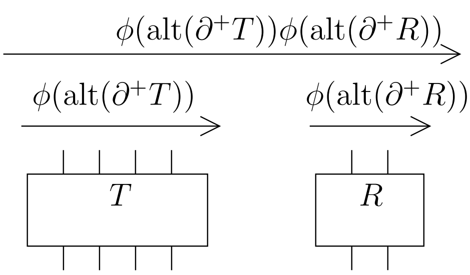


Figure 2.7.5

Also G contains $\tilde{\Delta}\tilde{\Delta}^{-1}$ since $\boxed{\boxed{\gamma}\ \boxed{\delta}}$ is in $P_2^{\Gamma,\Delta}$ where for $\gamma \in \Delta$, $\boxed{\gamma}$ is the "diagonal" tensor $\boxed{\begin{matrix}\gamma\\ \gamma\end{matrix}}$. That $G = G'$ if $e \in \Delta$ is easily seen by attaching $\boxed{e}$ to the right of basic tensors in $P_k^{\Gamma,\Delta}$ when k is even. □

We see that, if $e \in \Delta$, a basis for $P^{\Gamma,\Delta}$ is formed by all random walks on G, starting and ending at $1 \in \Gamma$, where the odd transitions correspond to multiplying by a $\tilde{\delta}$ for each $\delta \in \Delta$, and the even ones by $\tilde{\delta}^{-1}$ for $\delta \in \Delta$. If $\tilde{\Delta} = \tilde{\Delta}^{-1}$ and $\tilde{}$ is injective, these are just random walks on the Cayley graph of G.

If $e \in \Delta$, each basic tensor $T \in P^{\Gamma,\Delta}$ gives the *relation* $\mathrm{alt}(\partial T)$ in G, thinking of G as being presented on $\Delta\backslash\{e\}$.

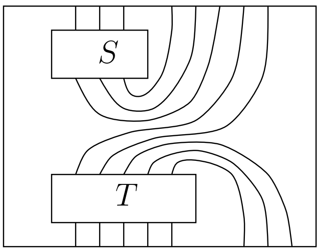
Suppose $G = \langle \Delta\backslash\{e\} \mid r_1, r_2, \ldots\rangle$ is a presentation of G, i.e. the kernel of the map induced by $\sim$ from $F(\Delta\backslash\{e\})$ to G is the normal closure of the r_i's. Then each r_i may be represented by a k-box, written $\boxed{r_i}$, with $\mu(\mathrm{alt}(\partial(\boxed{r_i}))) = r_i$, for some k with $2k \geq \ell(r)$. (We use $\ell(w)$ to denote the length of a word w.) To do this one may have to use $e \in \Delta$ so that the word r_i conforms with the alternating condition. For instance to represent $\gamma\delta^2\gamma^{-1}\delta$ one might use the basic tensor $\boxed{\begin{matrix}\gamma\ e\ \delta\ e\\ e\ \delta\ \gamma\ \delta\end{matrix}}$.

Let $\mu : F(\Delta) \to F(\Delta\backslash\{e\})$ be the homomorphism defined by $\mu(e) = 1 \in F(\Delta\backslash\{e\})$, $\mu|_{\Delta\backslash\{e\}} = \mathrm{id}$.

Definition. Let $R = \bigcup_{k=0}^{\infty} R_k$ be the planar subalgebra of $P^{\Gamma,\Delta}$ generated by $\{\boxed{\delta} : \delta \in \Delta\} \cup \{\boxed{r}\} \cup \{\boxed{r_i^{-1}}\}$. Let $H = \{\mu(\mathrm{alt}(\partial T)) \mid T \text{ is a basic tensor in } R\}$.

Theorem 2.7.6 *The set H is a subgroup of $F(\Delta\backslash\{e\})$ equal to the normal closure N of $\{r_i\}$ in $F(\Delta\backslash\{e\})$. Moreover, $P^{G,\Delta} = R$.*

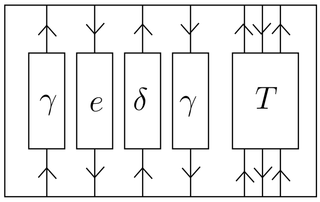
Proof. That H is multiplicatively closed follows from Figure 2.7.7.

$Q =$ 

$\mathrm{alt}(\partial Q) = \mathrm{alt}(\partial S)\mathrm{alt}(\partial T)$

Figure 2.7.7

To see that $H = H^{-1}$, note that the transpose of a basic tensor gives the inverse boundary word, and a planar algebra generated by a $*$-closed set of boxes is $*$-closed (note that the box $\boxed{\delta}$ is self-adjoint). Figure 2.7.8 exhibits conjugation of $\alpha = \mathrm{alt}(\partial T)$ by $\gamma\delta\gamma^{-1}$ for $\gamma, \delta \in \Delta$, which shows how to prove that H is normal.

$Q =$ 

$$\mu(\mathrm{alt}(\partial T) = \gamma\delta\gamma^{-1}\mathrm{alt}(\partial Q)\gamma\delta^{-1}\gamma^{-1}$$

Figure 2.7.8

Thus H contains the normal closure N.

Now the tangle picture of an arbitrary basic tensor in R can be isotoped so that it is as in Figure 2.7.9.

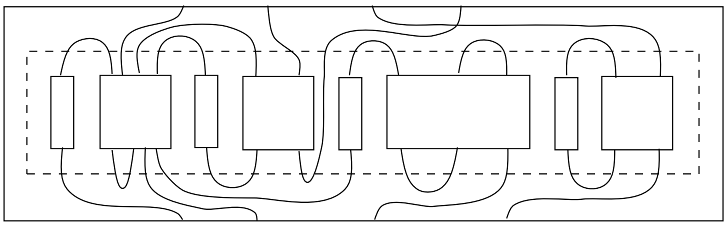


Figure 2.7.9

This is a tangle T all of whose curves are vertical straight lines surrounded by an element of the category $\mathcal{A}(\emptyset)$. But it is easy to see that applying an $\mathcal{A}(\emptyset)$ element changes $\mathrm{alt}(\partial T)$ at most by a conjugation.

The most difficult part of Theorem 2.7.6 is to show that $P^{\Gamma,\Delta} = R$. We must show that if w is a word of even length on Δ with $\phi(\mathrm{alt}(w)) = 1$ then there is a basic tensor $T \in R$ with $\partial T = w$.

As a first step, observe that if μ is a homomorphism from $F(\Delta)$ to $F(\Delta - \{e\})$ sending e to the identity and with $\mu(\delta) = \delta$ for $\delta \neq e$, then if $w_1, w_2 \in \mathfrak{S}(\Delta)$ are of even length and $\mu(\mathrm{alt}(w_1)) = \mu(\mathrm{alt}(w_2))$ then $\mathrm{alt}(w_1) = \mathrm{alt}(w_2)$. This is because $w_1 w_2^*$ (w^* is w written backwards) satisfies $\mathrm{alt}(w_1 w_2^*) = \mathrm{alt}(w_1)\mathrm{alt}(w_2)^{-1}$, thus $\mathrm{alt}(w_1 w_2^*) \in \ker \mu$ which is the normal closure of e. The length of $w_1 w_2^*$ can be reduced (if necessary) by eliminating consecutive letters two at a time to obtain another word w, of even length, with $\ell(w) = \ell(\mathrm{alt}\ w)$ (ℓ = length). By the uniqueness of reduced words in a free group, w must be a product of words of the form $x\ e\ y$, which map to conjugates of $e^{\pm 1}$ in $F(\Delta)$. But the last letter of x and the first letter of y must then be the same, and alt will send both these letters to the same free group element. Thus in the process of reducing $w_1 w_2^*$, all occurrences of e must disappear and $\mathrm{alt}(w_1) = \mathrm{alt}(w_2)$.

A consequence of this observation is that, if T is a basic tensor with $\phi(\mathrm{alt}(\partial T)) = 1$ in Γ so that $\mu(\mathrm{alt}(\partial T))$ is in the normal closure of $\{r_i\}$ in $F(\Delta\backslash\{e\})$, then $\mathrm{alt}(\partial T)$ is the normal closure of $\{\mathrm{alt}(\partial(\boxed{r_i}))\}$ in $F(\Delta)$. Thus it suffices to show that, if T_1 and T_2 are basic tensors with $\mathrm{alt}(\partial T_1) = \mathrm{alt}(\partial T_2) \in F(\Delta)$, then $T_1 = cT_2$ for some c in $\mathcal{A}(\emptyset)$ (since for any $x \in \mathfrak{S}(\Delta)$ with $\phi|\mu(\mathrm{alt}(x)) = 1$ we have shown there is a T in the planar algebra R with $\mathrm{alt}(\partial T) = \mathrm{alt}(x)$). But this is rather easy — we may suppose without loss of generality that no cancellation happens going from ∂T_1 to $\mathrm{alt}(\partial T_1)$ and then use induction on $\ell(\partial T_2)$. If $\ell(\partial T_2) = \ell(\partial T_1)$ then $\partial T_2 = \partial T_1$. Otherwise there must be a sequence $\ldots\delta\delta\ldots$ in ∂T_2 for some $\delta \in \Delta$. Connecting δ to δ in the tangle reduces the length of ∂T_2 by 2, and the remaining region is a disc. □

A.Casson has pointed out the connection between $P^{\Gamma,\Delta}$ and van Kampen diagrams.

If $e \in \Delta$, the dimension of $P_n^{\Gamma,\Delta}$ is the number of ways of writing $1 \in \Gamma$ as a product of elements $\tilde{\delta}$, $\delta \in \Delta$, with alternating signs. In particular, $\dim(P_1^{\Gamma,\Delta}) = |\Delta|^2$ iff Γ is trivial. Since the problem of the triviality of a group with given presentation is undecidable, we conclude the following.

Corollary 2.7.10 *The calculation of the dimension of a planar subalgebra of $P^{\otimes}$ is undecidable.*

Since there are groups which are finitely generated but not finitely presented we have

Corollary 2.7.11 *There are finite dimensional planar $*$-algebras which are not finitely generated as planar algebras.*

Proof. If finitely many linear combinations of basic tensors generated a planar algebra, then certainly the basic tensors involved would also. But by 2.7.6, the group would then be finitely presented. □

Example 2.8 Spin models. We give a general planar algebra that is not planar, although it is the planar algebras associated with it that will be of most interest. In some sense it is a "square root" of the planar algebra $P^{\otimes}$ of §2.6.

Let V be a vector space of dimension Q with a basis indexed by "spin states" $\{1, 2, \ldots Q\}$. For each odd n let P_n^σ be the subalgebra of $\mathrm{End}(V^{\otimes \frac{n+1}{2}})$ given by $\mathrm{End}(V^{\otimes \frac{n-1}{2}}) \otimes \Delta$ where Δ is the subalgebra of $\mathrm{End}(V)$ consisting of linear maps diagonal with respect to the basis. For $n = 0$, P_0 is the field K and for n even, $P_n^\sigma = \mathrm{End}(V^{\otimes \frac{n}{2}})$. Elements of P_n^σ will be identified with functions from $\{1, 2, \ldots, Q\}^n$ to K, the value of the function on $(i_1, i_2, \ldots, i_n)$ being the coefficient basic tensor $\boxed{\begin{matrix} i_1 & i_2 & \ldots & i_{m-1} & i_m \\ i_n & i_{n-1} & \ldots & i_{m+2} & i_{m+1} \end{matrix}}$ for $n = 2m$, and the coefficient of $\boxed{\begin{matrix} i_1 & i_2 & \ldots & i_{m-1} & i_m \\ i_n & i_{n-1} & \ldots & i_{m+2} & i_m \end{matrix}}$ for $n = 2m - 1$. (See §2.6 for notation.) We shall make P^σ into a planar algebra in two slightly different ways. In both cases the labelling set will be P^σ itself.

First planar structure on P^σ. Take a tangle T in $\mathcal{P}_k(L)$. We will define $\Phi_0(T) \in \mathcal{P}_k^\sigma$ as follows.

First, shade the connected components of $\mathcal{B}_k \backslash T$ (called regions) black and white so that the region containing a neighborhood of (0,0) is white, and so that regions whose closures intersect (i.e. which share an edge) have different colours. In other words, regions whose boundary induces the positive orientation of $\mathbb{R}^2$ are coloured white and negatively oriented ones are black. Observe that the top and bottom of $\mathcal{B}_k$ consists of segments of length one forming parts of the boundaries of regions alternately coloured white and black. If k is odd, the right-hand boundary of $\mathcal{B}_k$ can be joined with the rightmost top and bottom segments to form part of the boundary of a black region. This way the boundary of $\mathcal{B}_k$ always has k segments attached to black regions whose closure meets the boundary. Number these segments cyclically $1, 2, \ldots, k$ starting from the top left and going clockwise. To define an element of P_k^σ from T we must give a function $\Phi_0(T) : \{1, 2, \ldots Q\}^k \to K$. It is

$$\Phi_0(T)(i_1, i_2, \ldots i_k) = \sum_\sigma \prod_{B \in \left\{ \substack{\text{labelled boxes} \\ \text{of } T} \right\}} \sigma(B)$$

where σ runs over all functions from the black regions of T to $\{1, 2, \ldots, Q\}$ which take the value i_j on the black region whose closure contains the j^{th} boundary segment, for all $j = 1, 2, \ldots, k$. Given a labeled box B of T, and such a σ, the boundary segments of B which meet closures of black regions are numbered 1 to k_B so σ defines an element of $\{1, 2, \ldots, Q\}^{k_B}$, and thus the label of B gives a scalar $\sigma(B)$ in K. As usual empty sums are zero and empty products are 1. This completes the definition of Φ_0 and it is easily checked that Φ_0 presents P^σ as a planar algebra. The induced representation of $\mathcal{P}(\phi)$ gives a representation of TL with $\delta_1 = Q$, $\delta_2 = 1$. It is precisely the representation associated with the Potts model used by Temperley and Lieb in [TL].

To be sure of relevance to subfactors, we now show how to adjust these parameters so that $\delta_1 = \delta_2 = \sqrt{Q}$.

Second planar structure on P^σ. If T is a labeled tangle in $\mathcal{P}_k(L)$, we define a tangle $\tilde{T}$ in $\mathcal{P}_k(\phi)$ by "smoothing" all the boxes of $\tilde{T}$, i.e. replacing $\boxed{R}$ by $\uparrow\downarrow\uparrow$, and shrinking all non-through-strings to semicircles near the top or bottom of $\mathcal{B}_k$. Put $f(T) = Q^{\frac{1}{2}(n_+ - n_-) + \frac{1}{4}(n_+^\partial - n_-^\partial)}$ where n_+ and n_- are the numbers of positively and negatively oriented circles in $\tilde{T}$ respectively, and $n_\pm^\partial$ are similarly the numbers of positively and negatively oriented semicircles near the top and bottom. Thus defined, $f(T)$ is clearly an isotopy invariant, so we could redefine it by assuming all the boxes are parallel to the x-axis. Assuming all maxima and minima of the y-coordinate restricted to the strings of T are nondegenerate, $2(n_+ - n_-) - (n_+^\partial - n_-^\partial)$ is just $p_+ + q_+ - p_- - q_-$ where p_+, p_- are the numbers of local maxima of y oriented to the left and right respectively and similarly q_+ and q_- count minima to the right and left respectively. It follows that $T \to f(T)$ is multiplicative and indeed that if A is in the annular category one may define $f(A)$ so that $f(AT) = f(A)f(T)$.

The normalisation constant $n_+ - n_- + \frac{1}{2}(n_+^\partial - n_-^\partial)$ may seem mysterious. What is actually being calculated is the isotopy invariant $\int d\theta$ where the intgral is taken over the strings of the tangle and $d\theta$ is the change of angle or curvature 1-form, normalised so that integrating over a positively oriented circle counts one. The above factor is then this integral when all strings meet all boxes at right angles. Thus, by following shaded regions at every internal box, another formula for this normalisation factor k-tangle is

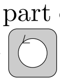
$$\left[\frac{k+1}{2}\right] - b - \sum_{i\geq 1}\left[\frac{n_i}{2}\right].$$

If all boxes are 2- or 3-boxes we get

$$\frac{1}{2}(\#(\text{black regions}) - \#(\text{boxes})) + \frac{1}{4}(n_+^\partial - n_-^\partial),$$

where now n_+^∂ and n_-^∂ are calculated by eliminating the boxes by following the black regions rather than going straight through the box.

Proposition 2.8.1 *The map $\Phi^\sigma : \mathcal{P}(L) \to P^\sigma$, $\Phi(T) = f(T)\Phi_0(T)$ (linearly extended) presents P^σ as a planar algebra, $\Phi|_{\mathcal{P}(\phi)}$ presents TL with $\delta_1 = \delta_2 = \sqrt{Q}$.*

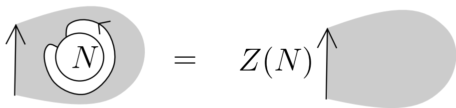
PROOF. Annular invariance follows from the relation $f(AT) = f(A)f(T)$ and the annular invariance of Φ_0. If  is a part of a tangle T then $\Phi(T) = QQ^{-\frac{1}{2}}\Phi(\tilde{T})$ where  has been removed from T. If  is part of T, $\Phi(T) = Q^{\frac{1}{2}}\Phi(\tilde{T})$. □

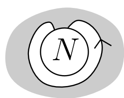
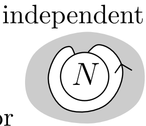
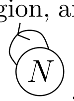
When we refer to P^σ, we will mean P^σ, together with Φ.

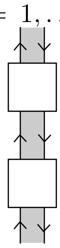
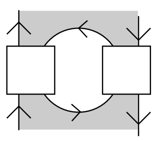
Although dim $P_0^\sigma = 1$, so that P^σ gives an invariant of labeled planar networks with unbounded region positively oriented, $\dim(P_{1,1}^\sigma) = Q$ so P^σ is not planar. However, P^σ does have the obvious $*$ structure and tr_R is defined and positive definite, so that any connected self-adjoint planar subalgebra of P^σ will be a C^*-planar algebra.

Proposition 2.8.2 *A planar subalgebra P of P^σ is spherical.*

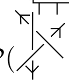
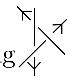
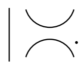
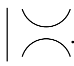
PROOF. Given a planar network N in P with positively oriented unbounded region, we need only show that



But since P is planar, the sum over all internal spins in  is independent of the spin value in the unbounded region, and each term in the sum for  is Q times the corresponding term for . Taking the sum over all Q spin states in the shaded region we are done. □

There are ways to obtain connected planar subalgebras of P^σ. An obvious place to look is association schemes where one is given a family of (0,1) $Q \times Q$ matrices A_i, $i = 1, \ldots, d$, whose linear span is closed under the operations given by the tangles  (matrix multiplication) and  (Hadamard product). The requirement that this linear span (the "Bose-Mesner algebra") be closed under all planar contractions is presumably much more stringent. If the requirement is satisfied, and the row and column sums of each A_i do not depend on the row or column, we will have a planar subalgebra of P^σ. A particularly simple example of this comes from transitive actions of a finite group G on a set S. Then the orbits of G on $S \times S$ define an association scheme whose Bose-Mesner algebra is the fixed points for the action on $M_{|S|}(\mathbb{C})$ by conjugation. We get a planar subalgebra of P^σ either by taking the fixed points for the G-action on P^σ or the planar subalgebra generated by the assocation scheme. They are different in general. A case where they are the same is for the dihedral group on a set with five elements (see [J4]). They are different for Jaeger's Higman-Sims model ([Ja],[dlH]) — although the dimensions of the two planar algebras agree for a while, they have different asymptotic growth rates, one being that of the commutant of Sp(4) on $(\mathbb{C}^4)^{\otimes k}$ and the other being 100^k.

Here is an interesting example for a doubly transitive group. It connects with Example 2.5 and gives a new kind of "spin model" for link invariants from links projected with only triple point singularities.

The alternating group A_4 is doubly transitive on the set $\{1, 2, 3, 4\}$ but there are two orbits on the set of ordered triples (a, b, c) of distinct elements according to whether $1 \mapsto a$, $2 \mapsto b$, $3 \mapsto c$, $4 \mapsto d$ (with $\{a, b, c, d\} = \{1, 2, 3, 4\}$) is an even or odd permutation. Let $\boxed{e} \in \mathcal{P}_3(L)$ be such that $\Phi^\sigma(\boxed{e})$ is the characteristic function of the even orbit. Defining a mapping from $\mathcal{P}($$)$ (the universal planar algebra on a single 3-box) to P^σ by sending  to $\boxed{e} - \frac{1}{2}$ . It is possible to prove that this map passes to the quotient P^H (the planar algebra of 2.5) with parameters $t = i = x$. This is equivalent to showing that twice the value of the HOMFLY polynomial of a link obtained by connecting 3-boxes  (at 1,–1 in $\ell - m$ variables) in an oriented way is the partition function in P^σ (with $Q = 4$) given by filling the same three boxes with $\boxed{e} - \frac{1}{2}$ . We give a sample calculation below which illustrates all the considerations. Note that, for $t = i = x$, the value of a single circle in the HOMFLY skein is 2.

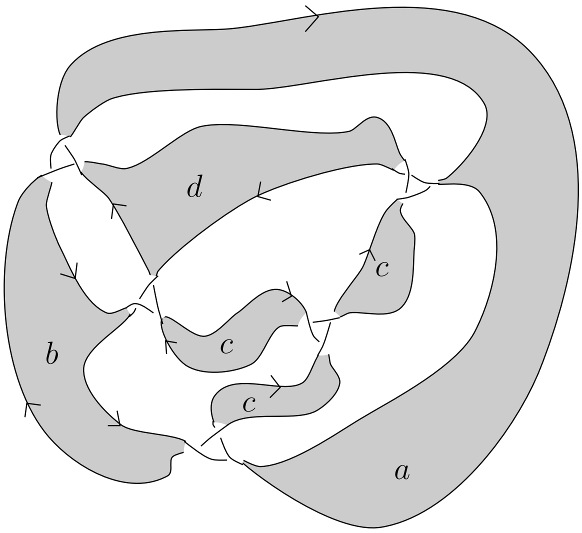
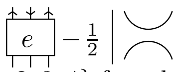
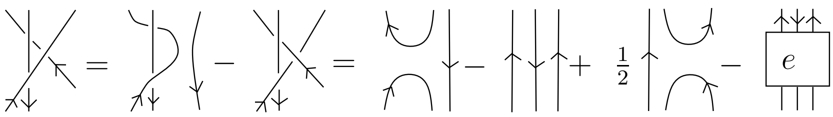


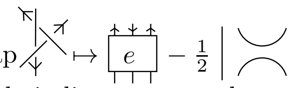
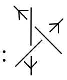
Figure 2.8.3

Smoothing all the 3-boxes leads to a single negatively oriented circle so we must divide the final partition function by 2. Replacing the 3-boxes by 

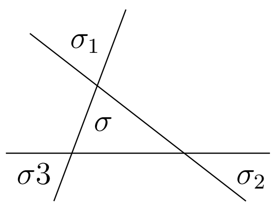
we look for spin states, i.e. functions from the shaded regions to $\{1,2,3,4\}$ for which each 3-box yields a non-zero contribution to the partition function. Around each 3-box this means that either the three spin values are in the even orbit under A_4, or they are all the same. The first case contributes $+1$ to the product over boxes, the second case contributes -1 (**not** $-\frac{1}{2}$ because of the maxima and minima in the box). If the box labeled (†) is surrounded by the same spin value, all the spin states must be the same for a nonzero contribution to Z. This gives a factor $4 \times (-1)^5$. On the other hand, if the spins at (†) are as in Figure 2.8.3 with (a,b,c) in the even orbit, the other spin choices are forced (where $\{a,b,c,d\} = \{1,2,3,4\}$), for a contribution of -1. The orbit is of size 12 so the partition function is $\frac{1}{2}(-12-4) = -8$. For this link the value of the HOMFLY polynomial $P_L(1,-1)$ is -4. The factor of 2 is accounted for by the fact that our partition function is 2 on the unknot. Thus our answer is correct. Note how few spin patterns actually contributed to Z!

If we wanted to use non-alternating 3-boxes we could simply use the HOMFLY skein relation to modify the 3-box. For instance



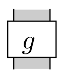
In general by [LM], $P_L(1,-1)$ is $(-1)^{c-1}(-2)^{\frac{1}{2}d}$ where c is the number of components of L and d is the dimension of the first homology group (with $\mathbb{Z}/2\mathbb{Z}$ coefficients) of the triple branched cover of S^3, branched over L. It would be reassuring to be able to see directly why our formula gives this value. This would also prove directly that the map  passes to the HOMFLY quotient. Our proof of this is a little indirect — one shows that the planar subalgebras P^H and $(P^\sigma)^{A_4}$ are the same by showing they arise as centralizer towers from the same subfactor (constructed in [GHJ]). Thus there must be a 3-box corresponding to: 
and we obtained the explicit expression for it by solving an obvious set of equations.

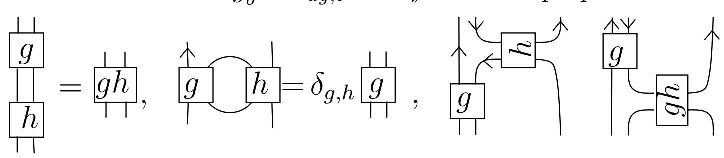
As far as we know, this is the first genuine "3-spin interaction" statistical mechanical model for a link invariant. Of course one may produce 3-spin interaction models by taking a 2-spin one and summing over the internal spin σ in the picture



but that is of little interest. One may check quite easily that the above model does not factorize in this way.

Example 2.9. Finite groups

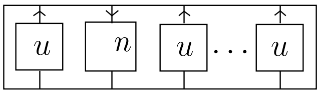
A special case of the subalgebra of P^σ of Example 2.8 where a group acts on a set S is where $S = G$ itself, the action being left multiplication. The algebra P_2 is then the group algebra $\mathbb{C}G$, linearly spanned by elements  for g in G, which are by definition the matrices $g^a_b = \delta_{ag,b}$. They have the properties



which can be used to present P as a planar algebra (see [La]). In particular $P_{1,3} = \ell^\infty(G)$ and $P_2 = M_{|G|}(\mathbb{C})$. It is obvious that G can be recovered from the abstract planar algebra P by using the minimal idempotents of $P_{1,3}$.

A noncommutative, finite-dimensional Hopf algebra gives a planar algebra but it cannot be a subalgebra of P^σ.

Example 2.10 Invariant planar algebras. Given an invertible element $u \in P_1$ in a general planar algebra P we define $u^{\otimes k}$ to be the element of P_k defined by the following k-tangle

$u^{\otimes k} =$ 

(k is odd in the picture). The u's and upside down u^{-1}'s alternate.

Proposition 2.10.1 *If P is a general planar algebra and S is a set of invertible elements of P_1, set*

$$P_k^S = \{x \in P_k \mid u^{\otimes k} x = x u^{\otimes k} \ \forall \ u \in S\} \ .$$

Then P^S is a general planar subalgebra of P (planar if P is) and a $*$-planar algebra if P is, and $S = S^*$.

Proof. That P^S is a unital filtered subalgebra is obvious. In the $*$ case note that $(u^{\otimes k})^* = (u^*)^{\otimes k}$. So by Lemma 1.18 we only have to check invariance under $\mathcal{A}(\phi)$. Given an $A \in \mathcal{A}_{k,n}(\phi)$ consider the tangle representing $(\otimes^n u)\pi_A(x)(\otimes^n u)^{-1}$ in Figure 2.10.2.

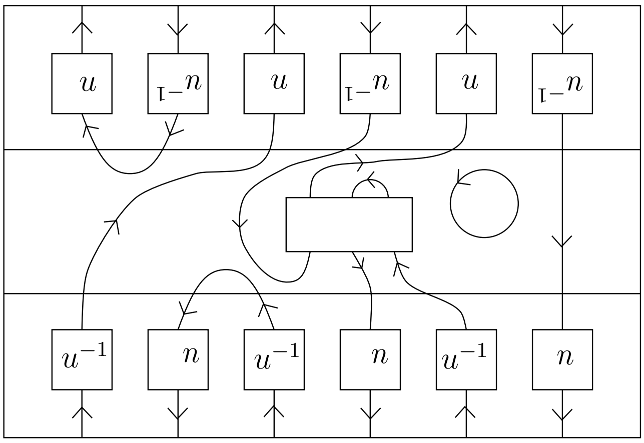
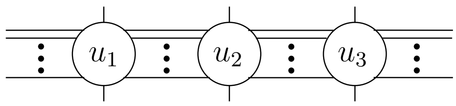


Figure 2.10.2

Each string of $\otimes^k u$, at the top and bottom, either connects to another external boundary point, in which case the u cancels with u^{-1}, or it connects to an internal boundary point of the annulus. Isotoping each such u and u^{-1} close to the internal boundary and inserting cancelling pairs of u and u^{-1} on strings connecting internal boundary points, we see that the tangle of Figure 2.10.2 gives the same element of P as the one where the only instances of u and u^{-1} surround the $\boxed{x}$ in an alternating fashion. Since $x \in P^S$, these u's may be eliminated and we are left with $\pi_A(x)$. □

This gives a useful way of constructing planar subalgebras. In particular P_1^S is the commutant of S in P_1 which may be much smaller than P_1. Of special interest is the case where $P = P^{\otimes}$ and S is a subgroup G of the unitary group. In this case $P_k^G = \mathrm{End}_G(V \otimes V^* \otimes V \otimes V^* \otimes \dots)$ (k copies of V or V^*) where V^* represents the contragredient representation of G. Other cases of interest can be constructed by cabling as in §3 and then picking some set of invertible elements in the original P_n.

To obtain a more general construction one may replace $\otimes^n u$ with tangles of the form

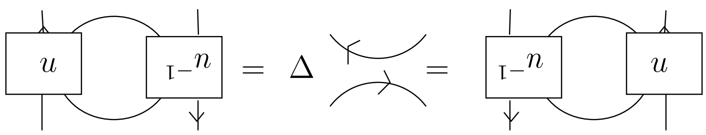


for u's satisfying appropriate equations. We will use this approach with $n = 1$ to pick up some important cases of commuting squares in Example 2.11.

Example 2.11. Binunitaries

Commuting squares have been used to construct subfactors (see [GHJ],[Ha]) and the general theory of calculating the subfactor planar algebra of §4.2 from a commuting square will be dealt with in a future paper. The treatment uses the language of statistical mechanical models with some attention paid to critical points as in chapter 4 of [JS]. Here we give a different approach which seems more natural from a planar point of view and will allow us to capture, as special cases, spin model commuting squares and some vertex model ones with no extension of the planar algebra formalism. The main concept is that of a bi-invertible element in a planar algebra which is the next step in the hierarchy discussed at the end of Example 2.10.

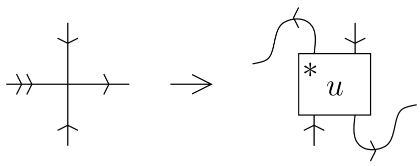
Definition 2.11.1. Let P be a general planar algebra. An invertible element $u \in P_2$ will be called *bi-invertible* if



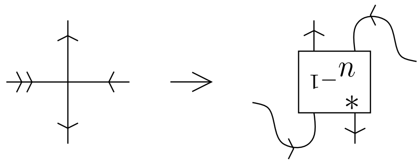
for some non-zero scalar Δ. (If P is planar Δ is necessarily δ_1 (δ_2). If P is a planar $*$-algebra, a bi-invertible u is called biunitary if $u^* = u^{-1}$. Bi-invertible elements define planar subalgebras as we now describe. It will be convenient to consider labelled tangles containing certain distinguished curves joining boundary points, which intersect with only the other strings in a tangle and do not meet any internal boxes. From such a tangle, and a bi-invertible element u, we construct an honest labelled tangle, in the sense of §1, in two steps.

(i) Orient the distinguished curves. The global orientation will be denoted $\twoheadrightarrow\!-$ and the tangle orientation by $\rightarrow\!-$.

(ii) At a point of intersection between the distinguished curves and the strings of the tangle, insert 2-boxes containing u or u^{-1} according to the following conventions



(global orientation)
(=tangle orientation going out)

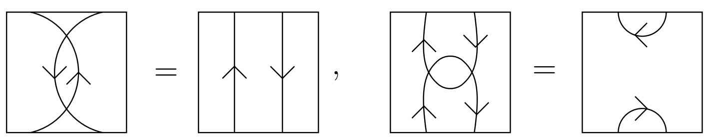


(global orientation)
(=tangel orientation going in)

Thus along a distinguished curve one alternately meets u or u^{-1}.

From now on we suppose for convenience that $\Delta = 1$.

Lemma 2.11.2 *The Reidemeister type II moves are satisfied, i.e.,*

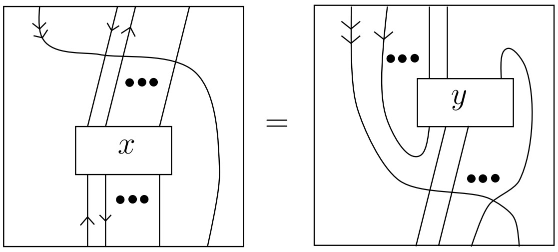


where either of the two curves is distinguished and its global orientation is arbitrary.

Proof. This is just a re-expression of bi-invertibility. □

Theorem 2.11.3 *If P is a general planar algebra and u is bi-invertible, let P_k^u be*

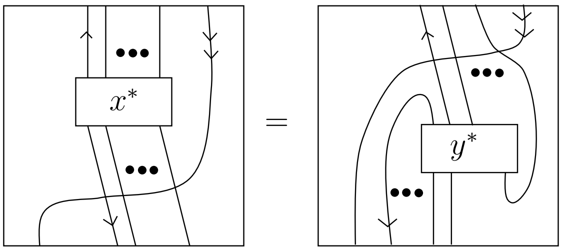
$\{x \in P_k$ s.t.



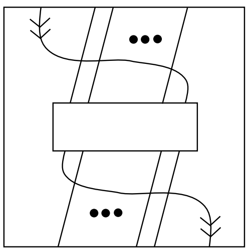
for some $y \in P_k\}$.

Then P^u is a general planar subalgebra, P is a (general) $*$-planar subalgebra if P is a planar algebra and u is bi-unitary. The properties of being planar, C^* and spherical are inherited from P.

Proof. That P^u is a subalgebra is obvious. The $*$-property is more interesting. Applying $*$ to the pictures we obtain (using $u^* = u^{-1}$)

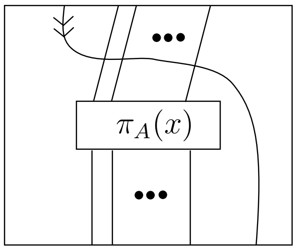


but we surround these pictures with the annular tangle



and then apply type II Reidemeister moves. We see that $x^* \in P^u$ if x does (though note that the "y" for x^* is y^*, but rotated).

To show that P^u is a general planar subalgebra, we only have to show by 1.18 that it is invariant under $\mathcal{A}(\phi)$. But if we arrange the annular tangle A so that all critical points of the height function on strings are local maxima and minima, the distinguished line



can be moved through $\pi_A(x)$, close to $\boxed{x}$ using only planar isotopy and type II Reidemeister moves. It can then go past x, producing a y, and down to the bottom via type II Reidemeister moves. Thus $\pi_A(x)$ is in P^u.

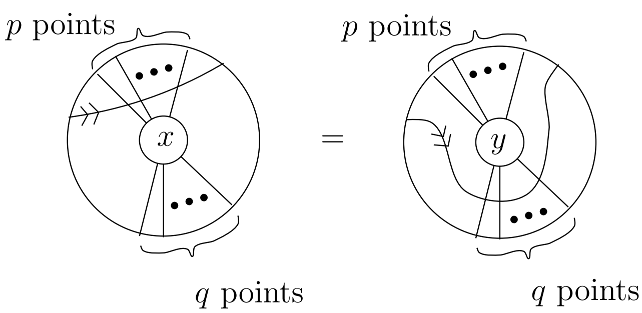
Planarity, positivity and sphericity are all inherited. □

Notes. (i) We will see that P^u may be planar even when P is only general planar.

(ii) When $k = 2$ the equation of Theorem 2.11.3 is an abstract version of the Yang-Baxter equation ([Ba]).

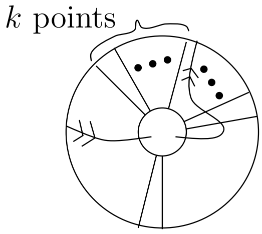
We now recast the equation of Theorem 2.11.3 in some equivalent forms which reveal some of its structure. The idea of the equation, as seen clearly in the proof of 2.11.3 is just that the distinguished lines can move freely past the boxes. As stated this requires a special configuration as the distinguished line approaches a box, but by Reidemeister type II invariance any approach will do. We record this below, keeping the notation of 2.11.3. (Note that it is somewhat cumbersome to force all the pictures to fit appropriately into the standard k-box. We use a disk with $2k$ boundary points as in the introduction.)

Proposition 2.11.4 *An x in P_k is in P_k^u iff there is a y in P_k^u with*



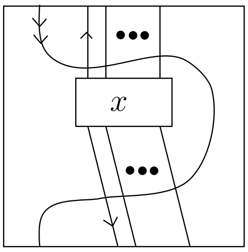
where $p + q = 2k$, where the pictures may be isotoped in any way so that the annular region becomes the standard annular region (with $*$ anywhere allowed by the orientations).

Proof. If $p > q$, surround the two pictures with the annular tangle



and use isotopy, rotation (if necessary to get $*$ in the right place), and type II Reidemeister moves to obtain the same picture as in 2.11.3. □

Definition 2.11.5. Given P and u as above, define $\sigma_u : P_k \to P_{k+1}$ by the tangle below:

$\sigma_u(x) =$ 

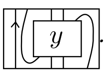
Proposition 2.11.6 *The map σ_u is a unital endomorphism of the filtered algebra P (a $*$-endomorphism if u is unitary) and $P_k^u = \{x \mid \sigma_u(x) \in P_{1,k+1}\}$.*

Proof. That σ_u preserves multiplication follows for type II Reidemeister moves. The alternative definition of P^u is just the case $q = 0$, with $*$ appropriately placed, in 2.11.4. □

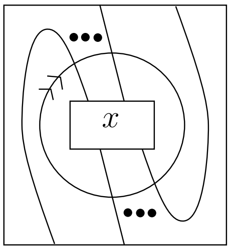
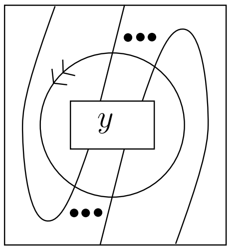
Note that the endomorphism σ_u is the obvious "shift de un" when restricted to the Temperley-Lieb subalgebra.

The condition of 2.11.5 involved a pair (x, y). In fact x is determined by y and vice versa as we now record.

Proposition 2.11.7 *Suppose P is a general planar algebra. If $x \in P^u$ with $\sigma_u(x) =$*



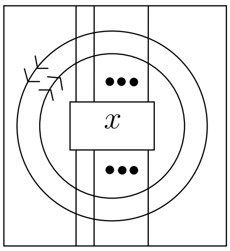
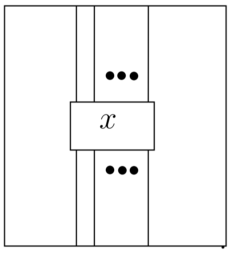
. Then

$y = \frac{1}{\delta_2}$  and $x\frac{1}{\delta_1} =$ 

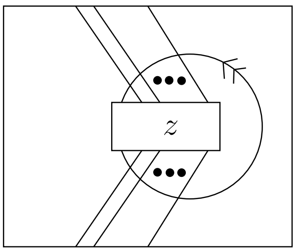
Proof. Just apply the appropriate annular tangles and use Reidemeister moves. □

Thus we could rewrite equations for $x \in P^u$ entirely in terms of x. The least obvious reformulation of these equations involves less boundary points than above and requires positivity in our proof.

Theorem 2.11.8 *Let P be a spherical finite-dimensional C^*-planar algebra and $u \in P_2$ be bi-unitary. Then $x \in P^u$ iff*

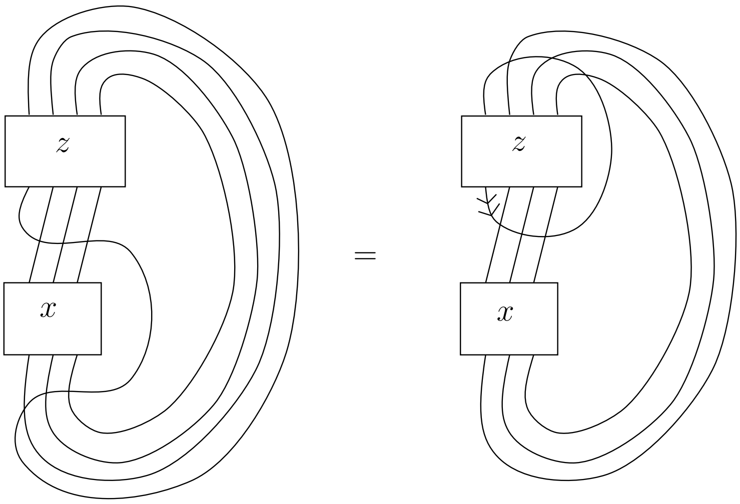
 $= \quad \delta^2$ 

Proof. (⇒) This is easy and requires no positivity.
(⇔) We begin by observing that $\delta\sigma_u^*(z)$ is given by the tangle below

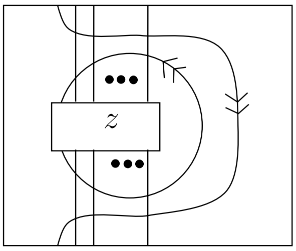


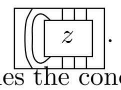
Here the adjoint of σ_u is as a map between the finite-dimensional Hilbert spaces P_k and P_{k+1} with inner products given by the normalized traces. This formula for σ_u^*

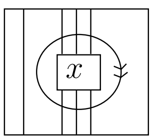
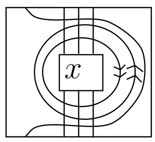
thus follows from the equality by isotopy of the following two networks, the first of which is, up to a power of δ, $\langle \sigma_u(x), z^* \rangle$



Thus orthogonal projection E onto $\sigma_u(P_k)$ is $\sigma_u\sigma_u^*$ which is given on $z \in P_k$ by $\frac{1}{\delta}$ times the following picture



Orthogonal projection F onto $P_{1,k+1}$ is given by $\delta F(z) =$ . An element w of $\sigma_u(P_k)$ is thus in $P_{1,k+1}$ iff $EF(w) = w$. But if x satisfies the condition of the theorem we have

$EF(\sigma_u(x)) = \frac{1}{\delta}($  $) = \frac{1}{\delta^2}$  $= \sigma_u(x).$

Hence $x \in P^u$. □

Remark. We did not use the full force of the hypotheses. The result will hold in a finite-dimensional general C^*-planar algebra provided isolated circles can be removed with a multiplicative factor δ, that the tangle formula for orthogonal projection onto $P_{1,k+1}$ is correct, and dim $P_1 = 1$.

We will see in the case of spin models that P^u may be planar although P is not. (But the conditions of the above remark are satisfied by P^σ.)

It is easy to check that a bi-invertible $u \in P_2$ may be altered by four invertible elements A, B, C, D in P_1 as in Figure 2.11.9.

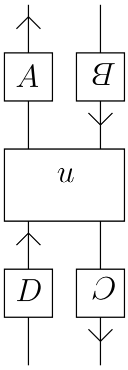


Figure 2.11.9

Definition 2.11.10. Two bi-invertibles are said to differ by a *gauge transformation* if one is obtained from the other as in Figure 2.11.9.

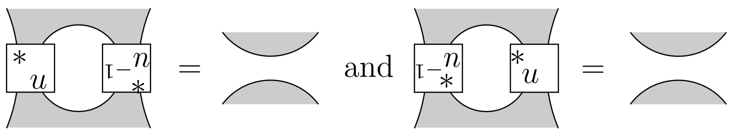
Gauge transformations have an inessential effect on P^u; A and C change absolutely nothing, B and D change P^u by a planar algebra isomorphism (induced by one on P-conjugation by $\boxed{d}\ \boxed{d}\ \boxed{d} \dots$).

In the $*$ case, gauge transformations on bi-unitary matrices are ones with A, B, C, D unitary.

A significant observation about the equations defining P^u above is that they are *linear* so the calculation of P_k^u, given P_k and u, is a *finite* problem, unlike the calculation of the planar subalgebra generated by some set, which requires consideration of infinitely many tangles. In practice, however, the brute force calculation, even just of dim P_k^u, runs into a serious problem. For the dimension of P_k grows exponentially with k. For $k = 2$ the calculation is usually easy enough (indeed we give an entirely satisfactory general solution for $k = 2$ when $P = P^\sigma$, below) and somewhat harder for $k = 3$. For $k = 4$ it tends to be very demanding even for relatively "small" P's. On the other hand, we are dealing with objects with a lot of structure. For instance, once we have calculated P_2^u by brute force or otherwise, the fact that P^u is a planar algebra means that every unlabelled 2-tangle gives a *nonlinear constraint.* For if a 2-tangle is given labels with elements in P, in order for the corresponding element of P_2 to be in P^u, it must lie in the linear subspace of P_2 already calculated. It was the desire to systematically exploit these highly interesting nonlinear constraints that led to the theory of planar algebras — their generality was only appreciated afterwards.

There are good reasons for wanting to calculate P^u. In general the calculation is greatly facilitated by the presence of group symmetry but there are many cases of bi-invertible elements with no apparent symmetry. We hope that the planar algebra P^u plays the role of "higher", non-group-like symmetries which reveal structural properties of the combinatorial object u.

We turn now to a special case where this program has had some partial success, namely in Hadamard matrices. The theory is no different for generalized Hadamard matrices, which occur as biunitaries for spin models. Consider the spin model P^σ with Q spins and its spherical planar algebra structure (§2.8). A bi-invertible element u of P_2^σ is an invertible $Q \times Q$ matrix u_a^b such that



So if $(u^{-1})_a^b = v_a^b$ we have $u_{a,b}v_{b,a} = 1/Q$ where the factor $1/Q$ comes from counting oriented circles after smoothing. If u is biunitary, $v_{a,b} = \overline{u_{b,a}}$, so the condition is precisely

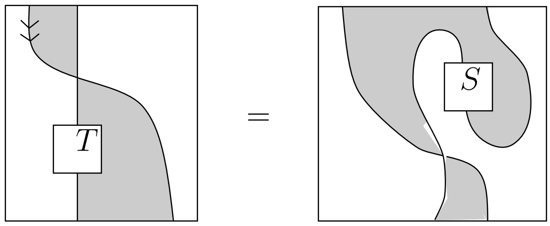
$$(2.11.11) \qquad |u_{a,b}| = \frac{1}{\sqrt{Q}}$$

We call a unitary matrix satisfying 2.11.11 a *generalized Hadamard matrix.* A Hadamard matrix is just $\sqrt{Q}$ times a real generalized Hadamard matrix.

Gauge transformations alter a generalized Hadamard matrix by multiplying rows and columns by scalars of modulus one (± 1 in the Hadamard case). This, together with permutations of the rows and columns, gives what is called *Hadamard equivalence* of (generalized) Hadamard matrices. Row and column permutations are easily seen by 2.11.6 to produce equivalent P^u's so any information about u obtained from P^u alone will be *invariant under Hadamard equivalence.* (The endomorphism σ_u of 2.11.5 itself is more information than just P^u.)

Proposition 2.11.12 *If u is a generalized Hadamard matrix, P^u is planar, hence a spherical C^*-planar algebra. Moreover,* $\dim P_1^u = 1$*, and P_2^u and $P_{1,3}^u$ are abelian.*

Proof. Obviously dim $P_1^u = 1$ implies planarity, so consider a tangle T representing an element of P_1. It consists of a vertical straight line with a 1-box on it, and networks to the left and right. The networks to the left have exterior shaded white so only contribute scalars. The picture below is the condition for such an element to be in P_1^u (for some tangle S)



If the bottom shaded region is assigned a spin a, and the top region a spin b, the left-hand side gives $u_a^b T_a$ and the right-hand side gives $u_a^b S_b$, so T_a is independent of a, and dim $P_1^u = 1$. $P_{1,3}^u$ is abelian because $P_{1,3}$ is P_2^u is abelian since it is $\sigma_u^{-1}(\sigma_u(P_2) \cap P_{1,3})$ by 2.11.6. □

So by §4.3, a generalized Hadamard matrix u yields a subfactor whose planar algebra invariant is P^u. In fact such a subfactor was the starting point of the theory of planar algebras, as the equations for P^u are those for the relative commutants of a spin model commuting square given in [JS]. Note that the original subfactor is hyperfinite whereas the one obtained from 4.3 is not! We now determine P_2^u for a generalized Hadamard matrix u.

Definition 2.11.13. Given a $Q \times Q$ generalized Hadamard matrix u_a^b we define the $Q^2 \times Q^2$ *profile matrix* Prof(u) by

$$\mathrm{Prof}(u)_{a,b}^{c,d} = \sum_x u_a^x \overline{u_b^x}\ \overline{u_c^x}\ u_d^x\ .$$

The profile matrix is used in the theory of Hadamard matrices. We will see that it determines P^u.

Definition 2.11.14. Given the $Q^2 \times Q^2$ matrix Prof(u), define the directed graph Γ_u on Q^2 vertices by $(a,b) \rightarrow\!\!- (c,d)$ iff $\mathrm{Prof}(u)_{a,b}^{c,d} \neq 0$.

The isomorphism class of Γ_u is an invariant of Hadamard equivalence.

Theorem 2.11.15 *If u is a $Q \times Q$ generalized Hadamard matrix thought of as a biunitary for the spin model P^σ, then the minimal projections of the abelian C^*-algebra P_2^u are in bijection with the connected components of the graph Γ_u.*

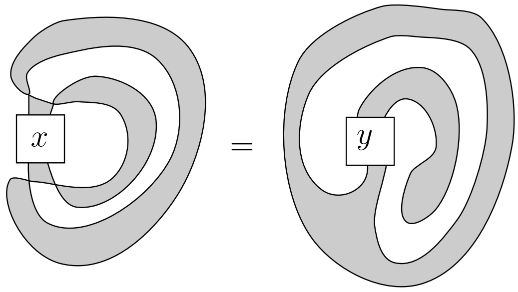
Moreover the (normalized) trace of such a projection is n/Q^2 where n is the size of the connected component, which is necessarily a multiple of Q.

PROOF. For matrices x_a^b, y_a^b, the equations of Theorem 2.11.3 are the "star-triangle" equations

$$\sum_d u_a^d\, \overline{u_b^d}\, x_d^c = u_a^c\, \overline{u_b^c}\, y_a^b$$

which amount to saying that, for each (a,b), the vector $v_{(a,b)}$ whose d^{th} component is the $u_a^d\overline{u_b^d}$ is an eigenvector of the matrix x_d^c with eigenvalue y_a^b. The profile matrix is just the matrix of inner products $\langle v_{(a,b)}, v_{(c,d)}\rangle$ so the orthogonal projection onto the linear span of $v_{(a,b)}$'s in a connected component is in P_2^u and is necessarily minimal since eigenvectors for distinct minimal projections are orthogonal.

If the matrix x is an orthogonal projection, y_a^b is either 1 or 0 depending on whether $v_{(a,b)}$ is in the connected component or not. Consider the picture



Applying Reidemeister type II moves and summing we obtain the assertion about the trace. (It is a multiple of $1/Q$ since x is a $Q \times Q$ matrix.) □

If G is a finite abelian group and $g \mapsto \hat{g}$ is an isomorphism of G with its dual $\hat{G}$ (=Hom$(G, \mathbb{C}^*)$), we obtain a generalized Hadamard matrix u, with $Q = |G|$, by setting $u_g^h = \frac{1}{\sqrt{Q}}\, \hat{h}(g)$. We call this a *standard* generalized Hadamard matrix. It is Hadamard if $G = (\mathbb{Z}/2\mathbb{Z})^n$ for some n. We leave it to the reader to check that if u is standard P^u is exactly the planar algebra of §2.9 for the group G. In particular, $\dim(P_k^u) = Q^k$. It is well known in subfactor theory that any subfactor with $N' \cap M_1 = \mathbb{C}^{[M:N]}$ comes from a group. It can also be seen directly from association schemes that if $\dim(P_2^u) = Q$ then u is standard up to gauge equivalence (recall that P^u is always an assocation scheme as remarked in §2.8).

We have, together with R. Bacher, P. de la Harpe, and M.G.V. Bogle performed many computer calculations. So far we have not found a generalized Hadamard matrix u for which $\dim(P_2^u) = 2$ but $\dim(P_3^u) > 5$. Such an example would be a confirmation of our non-group symmetry program as group-like symmetries tend to show up in P_2. In particular the five 16×16 Hadamard matrices have $\dim P_2^u =$ 16,8,5,3 and 3, and are completely distinguished by the trace. There are group-like symmetries in all cases corresponding to the presence of normalizer in the subfactor picture.

Haagerup has shown how to construct many interesting examples and given a complete classification for $Q = 5$. In the circulant case he has shown there are only finitely many examples for fixed prime Q (see [ ]).

Perhaps somewhat surprisingly, the presence of a lot of symmetry in u can cause P_2^u to be small! The kind of biunitary described in the following result is quite common — the Paley type Hadamard matrices give an example.

Proposition 2.11.17 *Suppose $Q-1$ is prime and let u be a $Q \times Q$ generalized Hadamard matrix with the following two properties (the first of which is always true up to gauge equivalence):*

(i) There is an index $*$ with $u_*^a = u_a^* = 1$ for all a.

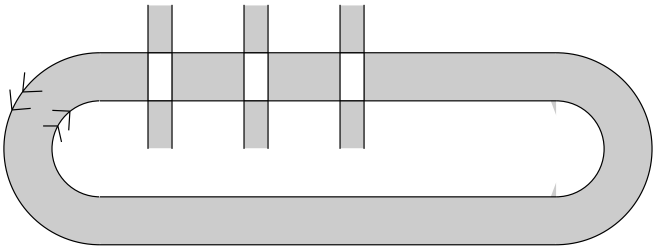
(ii) The group $\mathbb{Z}/(Q-1)\mathbb{Z}$ acts transitively on the spins other than $*$, and $u_{gb}^{ga} = u_b^a$ for all $g \in \mathbb{Z}/(Q-1)\mathbb{Z}$.

Then $\dim(P_2^u) = 2$ or u is gauge equivalent to a standard matrix.

Proof. The nature of the equations 2.11.15 makes it clear that $\mathbb{Z}/(Q-1)\mathbb{Z}$ acts by automorphisms on P_2^u, obviously fixing the projection e_1 which is the matrix $x_a^b = 1/Q$. Thus the action preserves $(1-e_1)P_2^u(1-e_1)$. Since $(Q-1)$ is prime there are only two possibilities: either the action is non-trivial and $\dim(P_2^u) = Q$ so P^u is standard, or every solution of 2.11.15 is fixed by $\mathbb{Z}/(Q-1)\mathbb{Z}$. In the latter case let x_a^b, y_a^b be a solution of 2.11.15. Then putting $c = *$ we obtain $\sum_d u_a^d \, \overline{u_b^d} x_d^* = y_a^b$, so y_a^b is determined by the two numbers x_*^* and x_d^*, $d \neq *$. So by 2.11.7 we are done. □

Note that the standard case in the above result can occur. The 8×8 Hadamard matrix is of the required form, but it is Hadamard equivalent to a standard matrix. For $Q = 12$ and 24 this cannot be the case and $\dim P_2^u = 2$.

We have very few general results on P_k^u for $k > 2$. We only record the observation that P_k^u is the δ^2 eigenspace for the $Q^k \times Q^k$ matrix given by the "transfer matrix with periodic horizontal boundary conditions" for the Q-spin vertex model having the profile matrix as Boltzmann weights. The transfer matrix is given by the picture:



where of course the internal spins have been summed over. This is an immediate consequence of 2.11.8.

We would like to make the following two open problems about matrices quite explicit. Both concern a generalized Hadamard matrix u.

(i) Is the calculation of dim P_k^u feasible in the polynomial time as a function of k?

(ii) Is there a u for which dim $P_k^u = \frac{1}{k+1}\binom{2k}{k}$? (i.e., P_k^u is just the Temperley-Lieb algebra).

Finally we make some comments on vertex models. There are many formal connections with Hopf algebras here which is not surprising since quantum groups arose from vertex models in statistical mechanical models ([Dr]). Banica has done some interesting work from this point of view — see [Ban].

A vertex model, in the above context, is simply a biunitary (or biinvertible) in the planar algebra $P^{\otimes}$. The equations of Theorem 2.11.3 are then just the equations for the higher relative commutants of a subfactor coming from a certain commuting square (see [JS]). Perhaps the most interesting examples not coming from the quantum group machinery are the Krishan-Sunder "bipermutation matrices" where u is a permutation matrix with respect to some basis of the underlying vector space (see [KS]). B.Bhattacharya has exhibited a planar algebra which is bigger than that of example 2.3. (Fuss Catalan) and which is necessarily a planar subalgebra of P^u if u is a bipermutation matrix.

3. General Structure Theory

3.1. Algebra structure, Markov trace

The proof of Theorem 3.1.3 below is routine for those conversant with [J1] or [GHJ]. We include it since, as stated, it can be useful in determining principal graphs. Recall that in a planar algebra P, e_k denotes the idempotent in P_k equal to $\frac{1}{\delta}(||\dots\ \ )$.

Lemma 3.1.1 *Let P be a finite-dimensional spherical nondegenerate planar algebra over an algebraically closed field. Then for each k, $P_{k-1}e_{k-1}P_{k-1}$ is a 2-sided ideal, denoted I_k, in P_k and if $\mathcal{M}_k$ is a set of minimal idempotents in P_k generating all the distinct minimal ideals in P_k/I_k, we have*

(i) $pP_ke_{k-1} = 0$ *for p in $\mathcal{M}_k$*

(ii) $pP_kq = 0$ *for $p \neq q$ in $\mathcal{M}_k$*

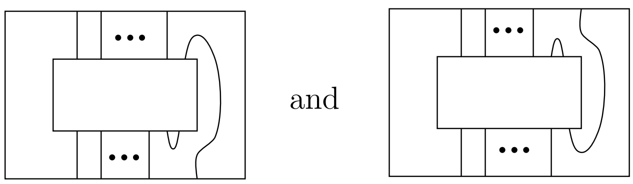
(iii) For each x in $P_k \backslash I_k$ there is a $p \in \mathcal{M}_k$ with $xP_kp \neq 0$.

(iv) $\mathrm{tr}(p) \neq 0$ *for all $p \in \mathcal{M}_k$.*

(v) $I_{k+2} = \bigoplus_{p \in \mathcal{M}_k} P_{k+2}pe_{k+1}P_{k+2}$, *$pe_{k+1}$ being a minimal idempotent in P_{k+2}.*

Moreover, if, for each k, $\mathcal{N}_k$ is a set of minimal idempotents of P_k satisfying (i) … (iv) *(with $\mathcal{M}_k$ replaced by $\mathcal{N}_k$), then there is an invertible u_k in P_k with $u_k\mathcal{N}_ku_k^{-1} = \mathcal{M}_k$ (so in particular (v) is true for $\mathcal{N}_k$).*

Proof. To see that $P_{k-1}e_{k-1}P_{k-1}$ is an ideal, consider the maps $\alpha, \beta : P_k \to P_{k-1}$ given by the annular tangles



respectively. A diagram shows that $xe_{k-1}y = \alpha(x)e_{k-1}\beta(y)$ for $x, y \in P_k$.

By Corollary 1.30, P_k is semisimple and multimatrix since K is algebraically closed. Thus properties (i), (ii), (iii) and (v) are obvious for $\mathcal{M}_k$. If $p \neq 0$ satisfied $\mathrm{tr}(p) = 0$, then tr would vanish on the whole matrix algebra containing p which would then be orthogonal to P_k.

Finally, suppose we are given $\mathcal{N}_k$ satisfying (i)–(iv). Then since P_k is multimatrix, each p in $\mathcal{N}_k$ belongs to a unique matrix algebra summand in which there is an invertible u_p with $u_ppu_p^{-1} \in \mathcal{M}_k$. Putting together the u_p's, and the identity of I_k, we get u_k. Property (v) for $\mathcal{N}_k$ then follows from (iii). $\square$

Definition 3.1.2. With P and $\mathcal{N}_k$ as in 3.1.1, we define the *principal graph* Γ_P of P to be the (bipartite) graph whose vertices are $\bigcup_{k \geq 0} \mathcal{N}_k$ with distinguished vertex $*$ so that $\mathcal{N}_0 = \{*\}$, and dim $(pP_{k+1}q)$ edges between $p \in \mathcal{N}_k$ and $q \in \mathcal{N}_{k+1}$. Let d_p denote the distance from p to $*$ on Γ_P.

Theorem 3.1.3 *As an algebra, P_k is isomorphic to the algebra whose basis is random walks of length $2k$ on Γ_P beginning and ending at $*$ with multiplication rule $w_1 w_2 = w_3$ if the first half of the walk w_2 is equal to the second half of w_1, and w_3 is the first half of w_1 followed by the second half of w_2; 0 otherwise. Moreover, if $\vec{t}$ is the function from the vertices of Γ_P to K, $\vec{t_p} = \delta^{d_p}\mathrm{tr}(p)$, $\vec{t}$ is an eigenvector for the adjacency matrix of Γ, eigenvalue δ.*

Proof. The first assertion is easily equivalent to showing that the Bratteli diagram (see [GHJ]) of the multimatrix algebra P_k in P_{k+1} is the bipartite graph consisting of those vertices p with $d_p \equiv k \pmod 2$ and $d_p \leq k$, connected to those with $d_p \equiv (k+1) \pmod 2$ and $d_p \leq k+1$ with appropriate multiplicities. Observe first that $P_0 = \mathbb{C}$ and $\mathcal{M}_1$ is a set of minimal projections, one for each matrix algebra summand of P_1, so the Bratelli diagram is correct for $P_0 \subset P_1$. Now proceed by induction on k. The trace on P_k is nondegenerate, as is its restriction to P_{k-1} so one may perform the abstract "basic construction" of [J1] to obtain the algebra $\langle P_k, e_{P_{k-1}}\rangle$ which is multimatrix and isomorphic to $P_k \otimes_{P_{k-1}} P_k$ as a $P_k - P_k$ bimodule via the map $x \otimes y \mapsto x e_{P_{k-1}} y$. Moreover the matrix algebra summands of $\langle P_k, e_{P_{k-1}}\rangle$ are indexed by those of P_{k-1}, which by induction are the vertices of Γ_P with $d_p \leq k-1$, $d_P \equiv (k+1) \bmod 2$. If one defines the trace tr on $\langle P_k, e_{P_{k-1}}\rangle$ by $\widetilde{\mathrm{tr}}(x e_{P_{k-1}} y) = \frac{1}{\delta^2}\mathrm{tr}(xy)$ then the traces of minimal projections in $\langle P_k, e_{P_{k-1}}\rangle$ are $\frac{1}{\delta^2}$ times those in P_{k-1}. Moreover, setting $\gamma(x e_{P_{k-1}} y) = x e_k y$ defines an algebra homomorphism from $\langle P_k, e_{P_{k-1}}\rangle$ which is injective by property (iv) and onto I_{k+1}. And $\mathrm{Tr}(x e_k y) = \frac{1}{\delta}\ \mathrm{Tr}(xy)$ so $\mathrm{tr} = \widetilde{\mathrm{tr}} \circ \gamma^{-1}$ on I_{k+1}. Properties (i), (ii) and (iii) ensure that the other vertices of the Bratteli diagram for $P_k \subset P_{k+1}$ are labelled by vertices p of Γ_P with $d_P = k+1$. And the number of edges on Γ_P connecting a p in $\mathcal{M}_k$ to a q in $\mathcal{M}_{k+1}$ is by definition the number of edges in the Bratteli diagram. That there are no edges between $\mathcal{M}_{k+1}$ and $\mathcal{M}_j$, $j < k$, follows from

$$(xe_{k-1}y)p = \frac{1}{\delta^2}\ x(e_{k-1}e_k e_{k-1})yp = \frac{1}{\delta^2}\ x(e_{k-1}e_k p e_{k-1})y = 0$$

by (i) for $x, y \in P_{k-1}$ and $p \in \mathcal{M}_{k+1}$.

Finally, the (normalized) trace of a minimal projection p in I_k is $\delta^{d_P - k}\vec{t_p}$ so the assertion about the trace follows as usual (see [J1]). □

Remarks. (1) Similarly, the algebras $P_{1,k}$ have a principal graph Γ'_P with the trace vector $\vec{s}$. We call Γ'_P the dual principal graph. Ocneanu has shown, in the C^* case, how to associate numerical data encoding the ensuing embedding of the random walk algebra of Γ'_P into that of Γ_P. This completely captures the planar algebra structure and is analogous to choosing local coordinates on a manifold. The same could be done under the hypotheses of Lemma 3.1.1. The principal graphs alone do not determine the planar algebra — for instance the algebras $(P^\sigma)^{\mathbb{Z}/4\mathbb{Z}}$ and $(P^\sigma)^{\mathbb{Z}/2\mathbb{Z}\oplus\mathbb{Z}/2\mathbb{Z}}$ of 2.8 have the same principal graph but are readily distinguished by counting fixed points under the rotation.

(2) In fact, the assumption of nondegeneracy on P in 3.1.1 and 3.1.3 could be replaced by the hypothesis P_k/I_k semisimple. Then conditions (i)–(iv) could be used to inductively guarantee nondegeneracy.

(3) If P had been a C^*-planar algebra, we would have a theorem $(3.1.3)^*$ with all $\mathrm{tr}(p)$ positive, all p's projections, and the obvious adjoint * on random walks.

Theorem 3.1.3 can be used to compute the principal graphs for Temperley Lieb and the Fuss Catalan algebra (it is the "middle pattern" method of [BJ2]). We now illustrate its use by calculating the principal graph of the nondegenerate planar algebras coming from Example 2.2. We work over $\mathbb{C}$ for convenience, and in the C^* case to simplify life.

Let (A, TR) be a finite-dimensional unital C^*-algebra with normalized faithful (positive) trace TR. The labelling set L is $L_1 = A$. We choose a number $\delta > 0$ and let τ_p be $\mathrm{TR}(p)$ for projections $p \in A$. A labelled network is then a disjoint union of smoothly embedded circles, each one containing a (possibly empty) sequence of 1-boxes labelled by elements in A. We define the partition function Z of such a collection of circles to be $\delta^{\#\text{ (circles)}} \prod_{\text{(circles)}} \mathrm{TR}(a_1 a_2 \dots a_n)$, where $a_1 a_2 \dots a_n$ are the labels on the given circle, numbered in order around the circle. The partition function Z is obviously multiplicative so we define $P^{(A,\mathrm{TR})}$ to be the nondegenerate planar algebra, with obvious $*$-structure, defined by 1.23. It is linearly spanned by Temperley-Lieb diagrams with a single labelled box on each string. The relations of Example 2.2 hold, noting that $Z(\boxed{p}) = \delta\tau_p$. For certain values of δ and traces TR we will compute the principal (and dual principal) graphs of $P^{(A,\mathrm{TR})}$ and the Markov trace, and show it to be a planar C^*-algebra. Let us first describe the graphs. Let $\mathcal{M} = \{p\}$ be a set of minimal projections in A, one for each matrix algebra direct summand and let $n_p = \dim(pA)$. Let $\mathfrak{S}(\mathcal{M})$ be the free semigroup with identity on $\mathcal{M}$. Let $\lambda : \mathrm{Proj} \to \mathbb{N} \cup \{\infty\}$ be a function, and W_λ be the set of words in $\mathfrak{S}(\mathcal{M})$ which contain no consecutive string of p's longer than $\lambda(p)$, for each p.

Definition 3.1.4. The graph $\Gamma_{A,\lambda}$ is the rooted tree having vertices W_λ, with n_p edges between w and wp for every $p \in \mathcal{M}$ with $\{w \cup wp\} \subset W_f$. The root $*$ is the identity of $\mathfrak{S}(\mathcal{M})$.

Thus if $\lambda(p) = 1$ for all p and A is abelian, $\Gamma_{A,\lambda}$ is the regular tree of valence $|\mathcal{M}|$. If $\lambda(p) = \infty$ for all p and A is abelian, the root $*$ of the tree $\Gamma_{A,\lambda}$ has valence $|\mathcal{M}|$ and all other vertices have valence $|\mathcal{M}| + 1$. If $A = \mathbb{C}p + \mathbb{C}q$ and $\lambda(p) = 1$, $\lambda(q) = 2$, the tree $\Gamma_{A,\lambda}$ is as in Figure 3.1.5.

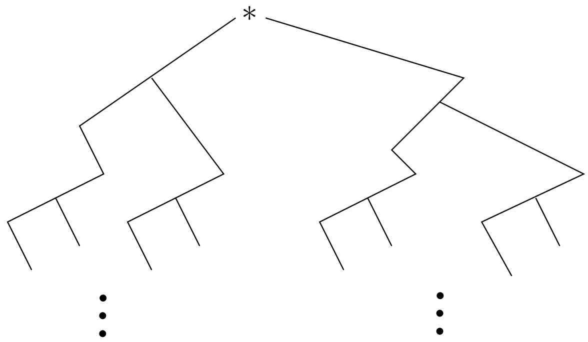


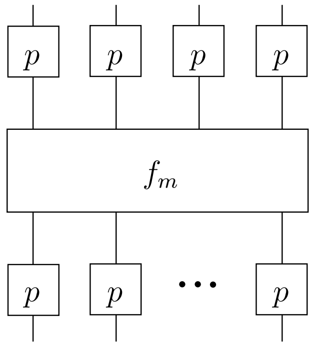
Figure 3.1.5

Recall the polynomials $T_n(x)$ of [J1], $T_1 = 1$, $T_2 = 1$, $T_{n+1} = T_n - xT_{n-1}$ and the "Jones-Wenzl" projections $f_k \in TL(k)$ with $f_k^* = f_k^2 = f_k$, $\;f_k e_i = 0$ for $i = 1, 2, \ldots, k-1$ so that f_k are elements of any (spherical) planar algebra, and $\mathrm{tr}(f_k) = T_{k-2}(\frac{1}{\delta^2})$ if $T_j(\frac{1}{\delta^2}) \neq 0$ for $j < k+2$.

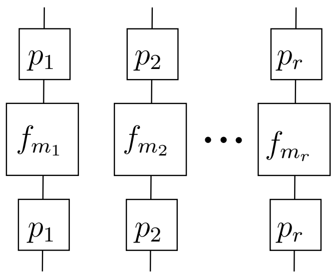
Theorem 3.1.6 *With notation as above, suppose $\tau_p\delta = 2\cos\pi/(\lambda(p)+2)$ ($\tau_p\delta \geq 2$ if $\lambda(p) = \infty$). Then $P^{(A,\mathrm{TR})}$ is a (spherical) C^*-planar algebra with principal and dual principal graphs equal to $\Gamma_{A,\lambda}$. The (normalized) trace of the minimal projection in $P_k^{(A,\mathrm{TR})}$ corresponding to the word $w = p_1^{m_1} p_2^{m_2} \ldots p_r^{m_r}$ ($\sum m_i = k$ and $p_i \neq p_{i+1}$) is $\prod_{i=1}^r \tau_{p_i} T_{m_i}(\tau_{p_i}\delta)$.*

Proof. We shall give explicit projections satisfying conditions (i)–(iv) of 3.1.1. The key observation is that tangles with a fixed p labelling each string form a Temperley-Lieb subalgebra B_p with parameter $\delta\tau_p$ (and identity $\boxed{p}\,\boxed{p}\ldots\boxed{p}$).

So if $m < \lambda(p)$ we consider the projection

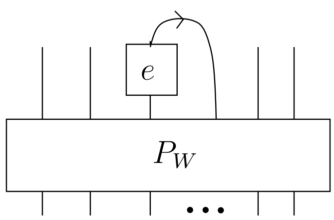


where f_m is calculated in Temperley-Lieb with j strings and $Z(\bigcirc) = \tau_p \delta$. Now set



where we have combined the m_j strings at the top (and bottom) of f_{m_j} into one. Orientations are completely forgotten and may be inserted, if required, so as to satisfy Definition 1.7.

Condition (1) of 3.1.1 is easy: P_k is linearly spanned by Temperley Lieb diagrams with matrix units (with the p's among the diagonal ones) from the simple summands of A in a single box on each string. If x is such an element, then if xe_{k-1} is non-zero, the product $p_w x e_{k-1}$ contains

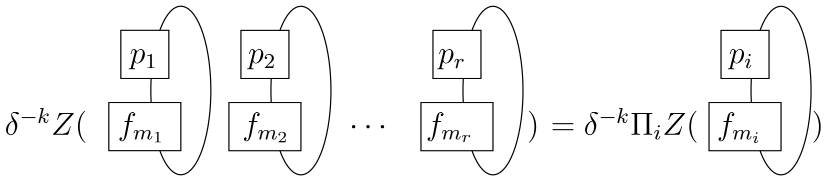


with e being a matrix unit. The string containing e enters the box for p_w either connecting two distinct f's or two strings of the same one. In the first case the result is zero since e belongs to precisely one of the direct summands. In the second case it is zero because of the properties of the f's.

Condition (ii) follows similarly, noting that a picture like the above will occur unless all the strings of x are through strings.

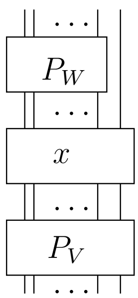
For condition (iii) we observe that the ideal I_k is linearly spanned by x's as above with less than k through-strings, so we may suppose x is composed of through-strings, each with a matrix unit in its box. Note that any relation true in B_p is true in P_k. Hence if $j \geq \lambda(p)$, $f_j = 0$ and the identity of B_p is a linear combination of tangles with less than $\lambda(p)$ through-strings. So we can suppose that in x there is no sequence of $\lambda(p)$ strings in a row whose matrix unit labels are in the same simple summand as p. Thus by multiplying x to the left and right by tangles with the appropriate matrix unit labels, we get $axb = p_w$ for some word w in $\mathfrak{S}(\mathcal{M})$ of length k. Thus (iii) will follow provided (iv) holds.

We calculate the normalized trace of p_w. It is

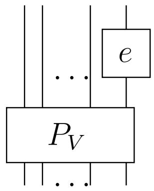


Now the partition function on P_m, restricted to B_p, gives a Markov trace which will be normalized after division by $Z(\boxed{p})^m = (\delta\tau_p)^m$. So $Z(\boxed{f_m}) = \delta^m \tau_p^m T_m(\tau_p\delta)$. Hence $\operatorname{tr}(p_w) = \prod_{i=1}^{r} \tau_{p_i} T_{m_i}(\tau_{p_i}\delta)$.

Finally we must calculate the multiplicities $\dim(p_v P_k p_w)$ for v of length k and w of length $k-1$. We must consider diagrams of the form



where x is a Temperley Lieb diagram decorated with matrix units as before. Arguing on p_v, x has only through-strings, $v = w_p$ for some $p \in \mathcal{M}$, and the first $k-1$ strings of x are labelled by the elements of $\mathcal{M}$ in w. Thus the diagram is in fact equal to

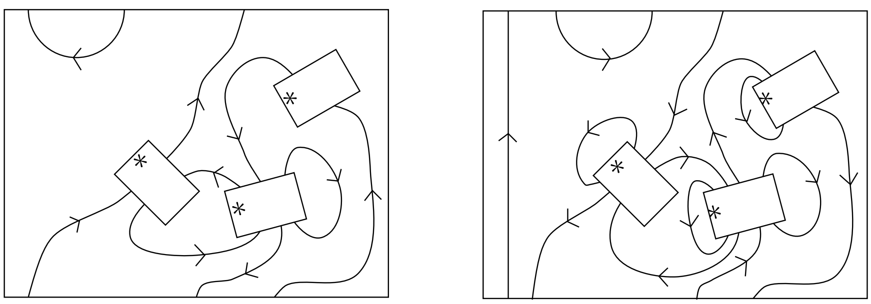


where e is a matrix unit with one subscript fixed. These diagrams span $p_v P_k p_w$ and the sesquilinear form given by $(x, y) = \mathrm{tr}(y^* x)$ is diagonal with non-zero entries. Hence $\dim(p_v P_k p_w) = n_p$. □

3.2. Duality

If $P = \cup_k P_k$ is a planar algebra, the filtered algebras $\lambda_n(P)$, where $\lambda_n(P)_k = P_{n,n+k}$, for fixed n, have natural planar algebra structures. For n even, this is rather obvious — just add n straight vertical lines to the left of a tangle. But if n is odd one must be more careful because of orientations. In fact $\lambda_1(P)$ and P are not isomorphic in general, even as filtered algebras, as one can see from example 2.9. We begin by describing the planar algebra structure on $\lambda_1(P)$.

If T is an unlabelled k-tangle we define the unlabelled $(k+1)$-tangle $\tilde{T}$ to be the tangle consisting of a vertical straight line from (1,0) to (1,1), and the tangle T, with all its orientations reversed, shifted by 1 in the positive x direction. Also in $\tilde{T}$ each internal p-box is replaced by a $(p+1)$-box with the first and last distinguished boundary points connected by a short curve. The procedure is illustrated in Figure 3.2.1.



The tangle T. The tangle $\tilde{T}$.

Figure 3.2.1

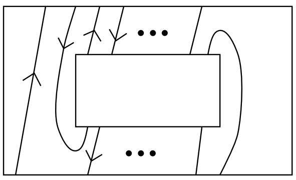
To each internal box B of T there corresponds in the obvious way a box $\tilde{B}$ of $\tilde{T}$. If T is labelled by $L = \coprod_{k>0} L_k$, $\tilde{T}$ will be given the obvious labelling by $\tilde{L}$, $\tilde{L}_k = L_{k-1} (\tilde{L}_1 = \emptyset)$.

Proposition 3.2.2 *Let $P = \cup_k P_k$ be a planar algebra with parameters $\delta_1 = Z(\circlearrowright)$ and $\delta_2 = Z(\circlearrowleft)$. Assume P is presented on itself by Φ. Then $\lambda_1(P)$ is a planar algebra with parameters δ_2, δ_1, presented on itself by $\lambda_1(\Phi)$ where $\lambda_1(\Phi)$ is defined by $\lambda_1(\Phi)(T) = \delta^{-p}\Phi(\tilde{T})$, p being the number of internal boxes in T.*

Proof. First note how the labels in a tangle of $\mathcal{P}(\lambda_1(P))$ give valid labels for $\mathcal{P}(P)$ because of the inclusion $\lambda_1(P)_k \subset P_{k+1}$. That $\lambda_1(\Phi)$ is a filtered algebra homomorphism is obvious. The annular invariance of ker $\lambda_1(\Phi)$ follows immediately from that of Φ, by representing elements of $P_{1,1+k}$ as linear combinations of tangles with vertical first string and applying $\sim$ to linear combinations. Thus $\lambda_1(P)$ is a general planar algebra.

Now $\lambda_1(P)_0 = P_{1,1}$ and $\lambda_1(P)_{1,1} = P_{2,2}$ which has dimension 1 since P is a planar algebra. So $\lambda_1(P)$ is planar. The multiplicativity property for $\lambda_1(P)$ follows immediately from that of P, where orientations on networks without boxes are reversed. $\square$

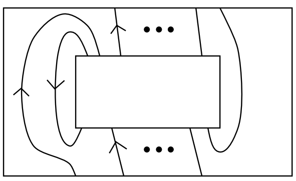
In the next two lemmas, A will be the $\mathcal{A}(\emptyset)$ element



where the actual number of boundary points is as required by context.

Lemma 3.2.3 *If P is a planar algebra, π_A defines a linear isomorphism between P_k and $\lambda_1(P)_k$, for each $k > 0$.*

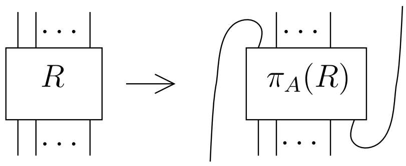
Proof. By a little isotopy and the definition of $P_{1,k+1}$, π_A is onto. But



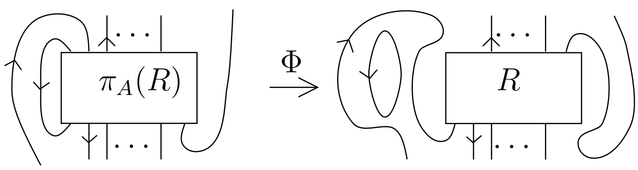
provides an inverse for π_A, up to a non-zero scalar. So π_A is an isomorphism. $\square$

Lemma 3.2.4 *The subset S of the planar algebra generates P as a planar algebra iff $\pi_A(S)$ generates $\lambda_1(P)$ as a planar algebra.*

Proof. (⇒) Given a tangle T in $\mathcal{P}_{1,k+1}(S)$, it suffices to exhibit a tangle T_A in $\mathcal{P}_k(\pi_A(S))$ with $\Phi(\tilde{T}_A)$ being a multiple of $\Phi(T)$. We create T_A from T by eliminating the first string, reversing all orientations and otherwise changing only in small neighborhoods of the internal boxes of G, sending a box labelled $R \in S$ in T to the box labelled $\pi_A(R)$ in T_A as below:



Then by definition $\tilde{T}_A$ will be exactly like T except near its boxes where it will look as below:



Thus $\Phi(\tilde{T}_A)$ is a multiple of $\Phi(T)$.

(⇐) Given $x \in P$, $\pi_A(\boxed{x})$ is in $\lambda_1(P)$ so by hypothesis it is the image under $\lambda_1(\Phi)$ of a linear combination of tangles labelled by elements of $\pi_A(S)$ which are in turn images under Φ of tangles labelled by elements of S (up to nonzero scalars). Using the tangle of Lemma 3.2.3 to invert π_A, we are done. □

By iterating the procedure $P \mapsto \lambda_1(P)$, we see that all the $\lambda_n(P)$ have natural planar algebra structures, but observe that all the $\lambda_{2n}(P)$ are isomorphic to P as planar algebras via the endomorphism (often called "le shift de deux") defined by adding two straight vertical strings to the left of a tangle. We leave the details to the reader.

The planar algebra $\lambda_1(P)$ is said to be the *dual* of the planar algebra P, and we have $\lambda_1(\lambda_1(P)) \simeq P$ as planar algebras.

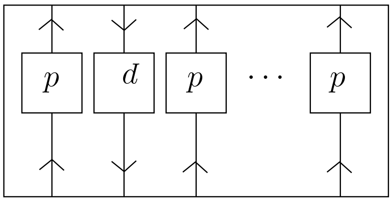
In the case of Example 2.9, P_2^G is the group algebra $\mathbb{C}G$ and $\lambda_1(P^G)_2$ is $\ell^\infty(G)$. The tangle π_A gives a *linear* isomorphism between the two. Thus planar algebra duality extends the duality between a finite group and its dual object.

3.3. Reduction and cabling

We give two ways to produce new planar algebras from a given one. The first is a reduction process which makes "irreducible" planar algebras — those with dim

$P_1 = 1$ — the focus of study. A planar algebra is not reconstructible in any simple way from its irreducible reductions, though, as can be seen from example 2 where the irreducible reductions would be trivial.

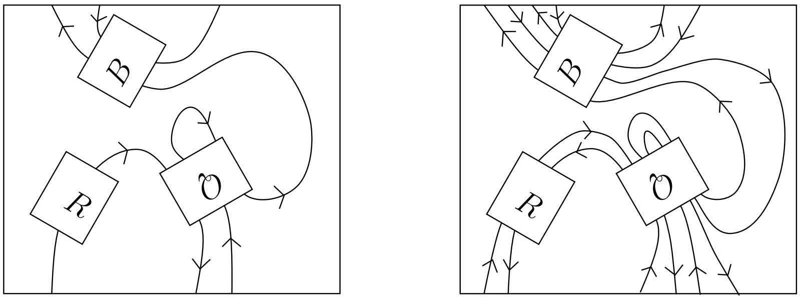
Given a general planar algebra P presented on itself by Φ, and an idempotent $p \in P_1$, we define the *reduced* general planar algebra pPp (by p) as follows: for each k we let p_k be the element



of P_k (illustrated when k is odd) and we set $(pPp)_k = p_k(P_k)p_k$ with identity p_k, and unital inclusion $P_k \hookrightarrow P_{k+1}$ given by $p_k x p_k \hookrightarrow p_{k+1} x p_{k+1}$ (note $p_{k+1}p_k = p_k p_{k+1} = p_{k+1}$). We make pPp into a planar algebra on P as follows. Given a tangle $T \in \mathcal{P}(\mathcal{P})$, define the tangle ${}_pT_p \in \mathcal{P}(P)$ by inserting $\boxed{p}$ in every string of TG. Then $p\Phi p : \mathcal{P}(L) \to pPp$ is $p\Phi p(T) = \Phi(pTp)$. Since p is idempotent, $p\Phi p$ is a filtered algebra homomorphism with annular invariance, obviously surjective, so pPp is a general planar algebra. Planarity is inherited from P and pPp has parameters $Z(\boxed{p})$ and $Z(\boxed{p})$, provided these are nonzero. If P is a C^*-planar algebra and p is a projection ($p = p^2 = p^*$), pPp is clearly also a C^*-planar algebra, spherical if P is.

Note. We have used the canonical labelling set for P to define pPp. If we were given another specific labelling set L, it is not clear that the homomorphism obtained in the same way from $\mathcal{P}(\mathcal{L})$ to pPp is surjective. We do not have any example of this phenomenon.

To use the reduction process we require dim $P_1 > 1$. But even "irreducible" planar algebras can yield this situation by cabling, i.e. grouping several strings together. The term is borrowed from knot theory. Given a general planar algebra P we define the n^{th} cabled (general) planar algebra $\mathcal{C}_n(P)$ by $\mathcal{C}_n(P)_k = P_{nk}$ which we endow with a planar algebra structure as follows. If Φ presents P on itself, we define $\mathcal{C}_n(\Phi) : \mathcal{P}(\mathcal{C}_n(P)) \to \mathcal{C}_n(P)$ by taking a labelled k-tangle T in $\mathcal{P}(\mathcal{C}_n(P))$ and constructing an nk-tangle $\tilde{T}$ in $\mathcal{P}(P)$ with the same labels on boxes, but where every boundary point in TG (both on internal and external boxes) is replaced by n boundary points, and of course orientations alternate. Every string in T is then replaced by n parallel strings. The procedure is illustrated in Figure 3.4.1, where $k = 3$ and $n = 2$.



The 3-tangle T The 3-tangle $\tilde{T}$

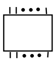
Figure 3.4.1

Then we define $\mathcal{C}_n(\Phi)(T)$ to be $\Phi(\tilde{T})$. It is clear that $\mathcal{C}_n(\Phi)$ is a general planar algebra, connected and multiplicative if P is, with parameters $(\delta_1\delta_2)^{\left[\frac{n}{2}\right]}\delta_1^{n-\left[\frac{n}{2}\right]}$, $(\delta_1\delta_2)^{\left[\frac{n}{2}\right]}\delta_2^{n-\left[\frac{n}{2}\right]}$, (where δ_1 and δ_2 are the parameters of the planar algebra P and $\left[\frac{n}{2}\right]$ is the integer part of $\left[\frac{n}{2}\right]$). Also $\mathcal{C}_n(P)$ is a C^*-planar algebra if P is, spherical if P is.

3.4. Tensor product

Let $P^1 = \cup_k P_k^1$ and $P^2 = \cup_k P_k^2$ be general planar algebras. We will endow the filtered algebra $P^1 \otimes P^2 = \cup_k P_k^1 \otimes P_k^2$ with a general planar algebra structure on the labelling set $L = \coprod_{i \geq 1} P_i^1 \times P_i^2$. Consider P^1 and P^2 presented on themselves by Φ_1 and Φ_2 respectively. First define a linear map $\mathcal{L} : \mathcal{P}(L) \to \mathcal{P}(P_1) \otimes \mathcal{P}(P_2)$ by $\mathcal{L}(T) = T_1 \otimes T_2$ where T is a tangle labelled by $f : \text{Boxes}(T) \to P_1 \times P_2$, T_i have the same unlabelled tangle as T and they are labelled by f composed with the projection $P_1 \times P_2 \to P_i$, $i = 1, 2$. This $\mathcal{L}$ is well defined since the isotopy classes of labelled tangles are a basis of $\mathcal{P}(L)$. Now define the presenting map $\Phi^{P_1 \otimes P_2} : \mathcal{P}(L) \to P_1 \otimes P_2$ by

$$\Phi^{P_1 \otimes P_2} = (\Phi_1 \otimes \Phi_2) \circ \mathcal{L} \ .$$

This obviously gives a homomorphism of filtered algebras. It is surjective because we may consider the tangle  labelled by $x \times y$ for an arbitrary pair (x, y) in $P_k^1 \times P_k^2$. This will be sent on to $x \otimes y \in P_k^1 \times P_k^2$. Thus we need only show the annular invariance of $\ker \Phi^{P_1 \otimes P_2}$. But if A is an annular tangle in $\mathcal{A}(L)$ then it is easy to see that $\mathcal{L} \circ \pi_A = \pi_{A_1} \otimes \pi_{A_2} \circ \mathcal{L}$ where A_1 and A_2 are the annular tangles having the same unlabelled tangle as A but labelled by the first and second components of the labels of A respectively. So if

$$x \in \ker \in \Phi^{P_1 \otimes P_2} \ ,$$

$$\Phi^{P_1 \otimes P_2} \pi_A(x) = \Phi_1 \otimes \Phi_2(\pi_{A_1} \otimes \pi_{A_2}(\mathcal{L}(x))) = \pi_{A_1} \otimes \pi_{A_2}(\Phi_1 \otimes \Phi_2(\mathcal{L}(x)))$$
$$= \pi_{A_1} \otimes \pi_{A_2}(\Phi^{P_1 \otimes P_2}(x)) = 0$$

So $P_1 \otimes P_2$ is a general planar algebra.

It is clear that $P_1 \otimes P_2$ is connected iff both P_1 and P_2 are and that $Z_{P_1 \otimes P_2} = Z_{P_1} Z_{P_2}$ in the sense that a network labelled with $P_1 \times P_2$ is the same as two networks labelled with P_1 and P_2 respectively. Thus $P_1 \otimes P_2$ is a planar algebra if P_1 and P_2 are. Moreover, nondegeneracy, $*$ structure, positivity and sphericity are inherited by $P_1 \otimes P_2$ from P_1 and P_2. So the tensor product of two C^*-planar algebras is a C^*-planar algebra.

Notes. (i) By representing elements of P_1 and P_2 by tangles, one may think of the tensor product planar structure as being a copy of P_1 and one of P_2 sitting in boxes on parallel planes, with no topological interaction between them. This corresponds to presenting $P_1 \otimes P_2$ on the labelling set $\mathcal{P}(P_1) \times \mathcal{P}(P_2)$ in the obvious way.

(ii) It is clear that $P_1 \otimes P_2 \approx P_2 \otimes P_1$ as (general) planar algebras.

3.5. Free product

The notion of free product of planar algebras was developed in collaboration with D. Bisch and will be presented in a future paper. The free product $P^1 \times P^2$ of two planar algebras P^1 and P^2 is by definition the subalgebra of the tensor product linearly spanned by (images of) T_1 consisting of a pair $T_1 \in P^1_k$, $T_2 \in P^2_k$ which can be drawn in a single $2k-1$-box, with boundary points in pairs, alternately corresponding to P^1 and P^2, so that the two tangles T_1 and T_2 are disjoint. An example is given by Figure 3.5.1 where we have used the “colours” to indicate boundary points belonging to P_1 and P_2.

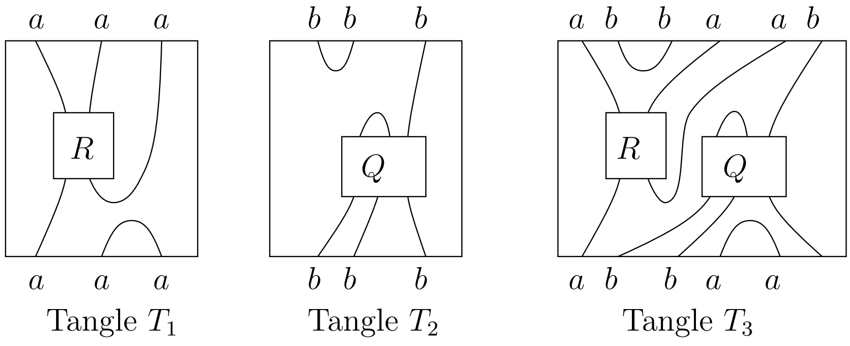


Figure 3.5.1

S. Gnerre defined in [Gn] a notion of free product using detailed connection calculations in the paragroup formalism.

The most interesting result so far of the work with Bisch is a formula, at least for finite dimensional C^*-planar algebras, for the Poincaré series of $P^1 \times P^2$ in terms of those of P^1 and P^2, using Voiculescu's free multiplicative convolution ([V]).

3.6. Fusion algebra

The reduced subalgebras of the cables on a planar algebra form a "fusion algebra" along the lines of [Bi]. This is to be thought of as part of the *graded algebra* structure given by a planar algebra and will be treated in detail in a future paper with D. Bisch.

4. Planar Algebras and Subfactors

In this section we show that the centralizer tower for an extremal finite index type II_1 subfactor admits the structure of a spherical C^*-planar algebra, and vice versa. We need several results from subfactors, some of which are well known.

4.1 Some facts about subfactors

Let $N \subset M$ be II_1 factors with $\tau^{-1} = [M : N] < \infty$. We adopt standard notation so that $M_i,\ i = 0, 1, 2, \ldots$ is the tower of [J1] with $M_0 = M,\ \ M_1 = \langle M, e_n \rangle$, $M_{i+1} = \langle M_i, e_{i+1} \rangle$ where $e_i : L^2(M_{i-1}) \to L^2(M_{i-2})$ is orthogonal projection, $e_N = e_1$. Let $B = \{b\}$ be a finite subset of M (called a basis) with

$$(4.1.1) \qquad \sum_{b \in B} b e_1 b^* = 1 \ .$$

Then by [PP] and [Bi],

$$(4.1.2) \qquad x(\sum_b be_1b^*) = (\sum_b be_1b^*)x = x \quad \text{for } x \in M.$$

$$(4.1.3) \qquad \sum_b bb^* = \tau^{-1}$$

$$(4.1.4) \qquad \sum_b bE_N(b^*x) = x = \sum_b E_N(xb)b^* \quad \text{for } x \in M.$$

$$(4.1.5) \quad \text{For } x \in N' \text{ (on } L^2(M)), \quad \tau \sum_b bxb^* = E_{M'}(x)\ .$$

4.1.6 Recall that the subfactor is called *extremal* if the normalized traces on N' and M coincide on $N' \cap M$ in which case the traces on $N' \cap M_k$, realized on $L^2(M)$, coming from M_k and N' coincide for all k, and $E_{M'}(e_1) = \tau$.

4.1.7 By standard convex averaging procedures on $L^2(M_k)$, given a finite subset $X = \{x\}$ of M_k and $\varepsilon > 0$, there is a finite set $U = \{u\}$ of unitaries in M and $\lambda_u \in \mathbb{R}^+$, $\sum_{u \in U} \lambda_u = 1$ with

$$\| \sum_u \lambda_u uxu^* - E_{M'}(x) \|_2 < \varepsilon \quad \text{for } x \in X.$$

Further averaging does not make the estimate worse, so by averaging again with $\sum_u \lambda_u \mathrm{Ad}\ u^*$ and gathering together repeated terms if necessary, we may assume $U^* = U$ and $\lambda_{u^*} = \lambda_u$.

It will be convenient to renormalize the e_i's so we set $E_i = \delta e_i$ with $\delta^2 \tau = 1$ ($\delta > 0$). (Note that there is a very slight notational clash with §2, but we will show it to be consistent.) We then have the formulae

$$(4.1.8) \qquad E_i^2 = \delta E_i\ , \quad E_iE_j = E_jE_i \ \text{ if } |i-j| \geq 2, \quad E_iE_{i\pm 1}E_i = E_i\ ,$$

$E_1xE_1 = \varepsilon E_N(a)E_1$ and $\sum_b bE_1e_2E_1b^* = 1$ so that $\{bE_1 \mid b \in B\}$ is a basis for M_1 over M.

Definition 4.1.9. For $k = 1, 2, 3, \ldots$ let $v_k = E_kE_{k-1} \ldots E_1 \in N' \cap M_k$. (Note $v_kxv_k^* = \delta E_N(x)E_k$ for $x \in M$, $v_k^*v_k = \delta E_1$.)

Theorem 4.1.10 *If x_i, $i = 1, \ldots, k+1$ are elements of M then $x_1v_1x_2v_2 \ldots v_kx_{k+1} = x_1v_k^*x_2v_{k-1}^* \ldots v_1^*x_{k+1}$, and the map $x_1 \otimes x_2 \otimes \cdots \otimes x_{k+1} \mapsto x_1v_1x_2v_2 \ldots v_kx_{k+1}$ defines an $M-M$ bimodule isomorphism, written θ, from $M \otimes_N M \otimes_N \cdots \otimes_N M$ (with $k+1$ M's), written $\bigotimes_N^{k+1} M$, onto M_k.*

Proof. See [J5], Corollary 11.

Recall that if R is a ring and B is an $R-R$ bimodule, an element b of B is called *central* if $rb = br \ \ \forall\, r \in R$.

Corollary 4.1.11 *The centralizer $N' \cap M_k$ is isomorphic under θ to the vector space V_{k+1} of central vectors in the $N-N$ bimodule $\bigotimes_N^{k+1} M$.*

We now define the most interesting "new" algebraic ingredient of subfactors seen from the planar point of view. It is the "rotation", known to Ocneanu and rediscovered by the author in specific models. See also [BJ1].

Definition 4.1.12. For $x \in M_k$ we define

$$\rho(x) = \delta^2 E_{M_k}(v_{k+1}\ E_{M'}(xv_{k+1}))$$

Proposition 4.1.13 *$\rho(M_k) \subseteq N' \cap M_k$ and if B is a basis, ρ coincides on $N' \cap M_k$ with $r : M_k \to M_k,\ \ r(x) = E_{M_k}(v_{k+1} \sum_{b\in B} bxv_{k+1}b^*)$.*

Proof. This is immediate from 4.1.5. □

Lemma 4.1.14 *With θ as above,*

$$\theta^{-1}r\theta(x_1 \otimes x_2 \otimes \cdots \otimes x_{k+1}) = \sum_{b\in B} E_N(bx_1)x_2 \otimes x_3 \otimes \cdots \otimes x_{k+1} \otimes b^*$$

Proof.

$$\begin{aligned} r(x_1v_1x_2v_2\ldots v_kx_{k+1}) &= \sum_b E_{M_k}(v_{k+1}bx_1v_{k+1}^*x_2v_k^*\ldots v_1^*b^*) \\ &= \delta\sum_b E_{M_k}(E_{k+1}E_N(bx_1)x_2v_k^*\ldots v_1^*b^*) \quad \text{(by(4.1.9))} \\ &= \sum_b E_N(bx_1)x_2v_k^*\ldots v_1^*b^* \\ &= \sum_b \theta(E_N(bx_1)x_2 \otimes x_3 \otimes \cdots \otimes x_{k+1} \otimes b^*) \end{aligned}$$

□

Note that the rotation on $\bigotimes_{\mathbb{C}}^{k+1} M$ does *not* pass to the quotient $\bigotimes_N^{k+1} M$, however we have the following.

Lemma 4.1.15 *Suppose $N \subset M$ is extremal, then*

$$\rho(\theta(x_1 \otimes x_2 \otimes \cdots \otimes x_{k+1})) = E_{N'}(\theta(x_2 \otimes x_3 \otimes \cdots \otimes x_{k+1} \otimes x_1))$$

Proof. If $x = \theta(x_1 \otimes x_2 \otimes \cdots \otimes x_{k+1})$ and $y = \theta(x_2 \otimes x_3 \otimes \cdots \otimes x_1)$, it suffices to show that $\mathrm{tr}(\rho(x)a) = \mathrm{tr}(ya)$ for all a in $N' \cap M_k$. Let $\varepsilon > 0$ be given and choose by 4.1.7 a finite set U of unitaries in M, with $U=U^*$, and $\lambda_u\in\mathbb{R}^+$, $\sum_{u\in U}\lambda_u=1$, $\lambda_{u^*}=\lambda_u$ so that

$$\|\sum_u \lambda_u v_{k+1}uxv_{k+1}u^* - v_{k+1}E_{M'}(xv_{k+1})\|_2 < \varepsilon \ ,$$

and, by extremality, $\|\sum_u \lambda_u uE_1u^* - \delta^{-1}\|_2 < \varepsilon$. So, if $y\in M$,

$$|\sum_u \lambda_u uE_N(u^*y) - \tau y\|_2 < \varepsilon\|y\| \qquad \text{(by(4.1.8)and } \|ab\|_2 \le \|a\|\ \|b\|_2. \tag{4.1.16}$$

So

$$\begin{aligned}
|\mathrm{tr}((\rho(x)-y)a| &= |\mathrm{tr}(\delta^2 E_{M_k}(v_{k+1}E_{M'}(xv_{k+1}))-y)a| \\
&< |\mathrm{tr}(\delta^2 E_{M_k}(v_{k+1}\sum_u \lambda_u uxv_{k+1}u^*)-y)a| + \delta^2\varepsilon\|a\|_2 \\
&= |\mathrm{tr}(\delta^2 \sum_u \lambda_u E_N(ux_1)\theta(x_2\otimes x_3\otimes\cdots\otimes x_{k+1}\otimes u^*)-y)a| + \delta^2\varepsilon\|a\|_2 \\
&\qquad \text{(as in the proof of 4.11)} \\
&= |\delta^2\big((\mathrm{tr}\sum_u \lambda_u\theta(x_2\otimes x_3\otimes\cdots\otimes x_{k+1}\otimes u^*E_N(ux_1)) - \theta(y)\big)a\big) + \delta^2\varepsilon\|a\|_2 \\
&\qquad \text{(since } a\in N'\text{)} \\
&= |\delta^2\mathrm{tr}(\theta(x_2\otimes x_3\otimes\cdots\otimes x_{k+1}\otimes(\sum_u \lambda_u uE_N(u^*x_1)-x_1))a)\|_2 + \delta^2\varepsilon\|a\|_2 \\
&\qquad \text{(since } U=U^*,\quad \lambda_u=\lambda_{u^*})\ .
\end{aligned}$$

For fixed $x_1, x_2, \ldots, x_{k+1}$ and a, this can clearly be made as small as desired by 4.1.16, by choosing ε small. □

Corollary 4.1.17 *If $x\in N$ and $\xi\in L^2(M_k)$, $\rho(x\xi-\xi x)=0$.*

Proof. By its definition, ρ extends to a bounded linear map from $L^2(M_k)$ to itself, so it suffices to show the formula for ξ of the form $\theta(x_1\otimes x_2\otimes\cdots\otimes x_{k+1})$. But if $n\in N$, $E_{N'}(\theta(x_2\otimes\cdots\otimes x_{k+1}\otimes nx_1)) = E_{N'}(\theta(x_2\otimes\cdots\otimes x_{k+1}n\otimes x_1))$, so by 4.1.15 we are done. □

Theorem 4.1.18 *If $N\subset M$ is extremal, $\rho^{k+1}=\mathrm{id}$ on $N'\cap M_k$.*

Proof. Recall that if $\mathcal{H}$ is an $N-N$ bimodule (correspondence as in [Co]) then $\langle\eta, x\xi-\xi x\rangle = \langle x^*\eta-\eta x^*,\xi\rangle$ for $x\in N$ and $\xi,\eta\in\mathcal{H}$. So η is central iff it is orthogonal to commutators.

Hence Lemma 4.1.15 reads

$$\rho(\theta(x_1\otimes x_2\otimes\cdots\otimes x_{k+1})) = \theta(x_2\otimes x_3\otimes\cdots\otimes x_1)+\xi$$

where $\xi \in \kappa$, which we define to be the closure of the linear span of commutators in the $N-N$ correspondence $L^2(M_k)$. Applying ρ to both sides of this equation, and 4.1.17, we obtain

$$\rho(\theta(x_1 \otimes \cdots \otimes x_{k+1})) - \theta(x_3 \otimes x_4 \otimes \cdots \otimes x_1 \otimes x_2) \in \kappa$$

and continuing,

$$\rho^{k+1}(\theta(x_1 \otimes \cdots \otimes x_{k+1})) - \theta(x_1 \otimes x_2 \otimes \cdots \otimes x_{k+1}) \in \kappa \ .$$

So, by linearity, if $x \in N' \cap M_k$, $\rho^{k+1}(x) - x \in \kappa$. But both $\rho^{k+1}(x)$ and x are central, so orthogonal to κ and $\rho^{k+1}(x) = x$. □

We now define five types of maps between centralizers using the isomorphism θ. Choose a basis B.

Definition 4.1.19. If $x = x_1 \otimes x_2 \otimes \cdots \otimes x_k \in \bigotimes_N^k M$,

(1) For $j = 2, 3, \ldots, k$,
$a_j(x) = \delta(x_1 \otimes x_2 \otimes \cdots \otimes x_{j-1}E(x_j) \otimes \cdots \otimes x_k) \in \bigotimes_N^{k-1} M$,
and $a_1(x) = \delta E(x_1)x_2 \otimes \cdots \otimes x_k$.

(2) For $j = 2, 3, \ldots, k$,
$\mu_j(x) = x_1 \otimes x_2 \otimes \cdots \otimes x_{j-1}x_j \otimes \cdots \otimes x_k) \in \bigotimes_N^{k-1} M$

(3) For $j = 2, 3, \ldots, k$,
$\eta_j(x) = x_1 \otimes x_2 \otimes \cdots \otimes x_{j-1} \otimes 1 \otimes x_j \otimes \cdots \otimes x_k) \in \bigotimes_N^{k-1} M$,
and $\eta_1(x) = 1 \otimes x_1 \otimes \cdots \otimes x_k$.

(4) For $j = 1, 2, \ldots, k$,
$\kappa_j(x) = \delta^{-1} \sum_{b \in B} x_1 \otimes \cdots \otimes x_j b \otimes b^* \otimes x_{j+1} \otimes \cdots \otimes x_k) \in \bigotimes_N^{k-1} M$

(5) If $\theta(c) \in N' \cap M_{n-1}$ define $\alpha_{j,c} : \otimes_N^k M \to \otimes_N^{k+n} M$ by
$\alpha_{j,c}(x) = x_1 \otimes \cdots \otimes x_{j-1} \otimes c \otimes x_j \otimes \cdots \otimes x_k$
($\alpha_{1,c}(x) = c \otimes x$, $\alpha_{k+1,c}(x) = x \otimes c$. Note also $\alpha_{j,1} = \eta_j$.)

Note all these maps are N middle linear (for (5) this requires b to be central; for (4) we use 4.1.2), so they are defined on the tensor product over N. They are all $N-N$ bimodule for maps so they preserve central vectors and are thus defined on the space V_k of N-central elements of $\otimes_N^k M$ $(= \theta^{-1}(N' \cap M_{k-1}))$. We will use the same notation for the restrictions of these maps to the V_k. Note that κ_j does not depend on the basis, indeed $\kappa_j(x) = \mu_{j+1}\alpha_{j+1,\mathrm{id}}(x)$ where $\theta(\mathrm{id})$ is the identity of $N' \cap M$, is a basis-independent formula for κ_j.

Lemma 4.1.20 *If $c \in V_k$ and $d \in V_{k'}$, then* *(i)* $\alpha_{d,j+k}\alpha_{c,i} = \alpha_{c,i}\alpha_{d,j}$ *for* $i \leq j$.

(ii) For $i < j$, $a_{j-1}a_i = a_i a_j$, $\mu_{j-1}\mu_i = \mu_i\mu_j$.

For $i \leq j$, $\eta_{j+1}\eta_i = \eta_i\eta_j$, $\kappa_{j+1}\kappa_i = \kappa_i\kappa_j$.

(iii) For $i \leq j$,

$$\begin{aligned} a_{j+k}\alpha_{c,i} &= \alpha_{c,i}a_j & a_{j-1}\alpha_{c,j} &= \alpha_{c,j-1}a_{i-1} \\ \mu_{j+k+1}\alpha_{c,i} &= \alpha_{c,i}\mu_{j+1} & \mu_{i-1}\alpha_{c,j} &= \alpha_{c,j-1}\mu_{i-1} \\ \eta_{j+k}\alpha_{c,i} &= \alpha_{c,i}\eta_j & \eta_{i-1}\alpha_{c,j-1} &= \alpha_{c,j}\eta_{i-1} \\ \kappa_{j+k}\alpha_{c,i} &= \alpha_{c,i}\kappa_j & \kappa_{i-1}\alpha_{c,j} &= \alpha_{c,j+1}\kappa_{i-1} \end{aligned}$$

(iv) For $i < j$,

$$\begin{aligned} \mu_j a_i &= a_i\mu_{j+1} & \mu_i a_j &= a_{j-1}\mu_i \\ \eta_{j-1}a_i &= a_i\eta_j & \eta_{i+1}a_j &= a_{j+1}\eta_{i+1} \\ \kappa_{j-1}a_i &= a_i\kappa_j & \kappa_i a_j &= a_{j+1}\kappa_i \\ \eta_{j-1}\mu_i &= \mu_i\eta_j & \eta_i\mu_j &= \mu_{j+1}\eta_i \\ \kappa_{j-2}\mu_i &= \mu_i\kappa_{j-1} & \kappa_i\mu_j &= \mu_{j+1}\kappa_i \\ \kappa_j\eta_i &= \eta_i\kappa_{j-1} & \kappa_i\eta_j &= \eta_{j+1}\kappa_i \end{aligned}$$

(v)

$$\begin{aligned} a_i\kappa_i &= \mathrm{id} \\ \mu_i\eta_i &= \mathrm{id} \\ a_{i+1}\kappa_i &= \mathrm{id} \\ \mu_{i+1}\eta_i &= \mathrm{id} \end{aligned}$$

(vi)

$$\begin{aligned} a_j\eta_j &= \delta\mathrm{id} \\ \mu_{j+1}\kappa_j &= \delta\mathrm{id}. \end{aligned}$$

(These identities hold when i, j and k are such that all the maps involved are defined by 4.1.19.)

Proof. Almost all cases of identities (i)–(iv) are trivial as they can be written so as to involve distant tensor product indices: thus they just amount to a renumbering. The ones that involve some interaction between the tensor product components are

$$\mu_i\mu_i = \mu_i\mu_{i+1}\ , \quad \kappa_{i+1}\kappa_i = \kappa_i\kappa_i\ , \quad \kappa_{i-1}\mu_i = \mu_i\kappa_i\ , \quad \kappa_i\mu_{i+1} = \mu_{i+2}\kappa_i\ .$$

These all follow easily from associativity of multiplication and
$\displaystyle\sum_{b\in B} b\otimes b^* x = \sum_{b\in B} xb\otimes b^*$ for $x\in M$, which is 4.1.2.

$$\begin{aligned}
\text{For (v): } a_i\kappa_i &= \text{id follows from} \sum_{b\in B} E_N(xb)b^* = x \quad (4.1.4)\\
\mu_i\eta_i &= \text{id follows from } x1 = x\\
a_{i+1}\kappa_i &= \text{id follows from} \sum_{b\in B} bE(b^*x) = x \quad (4.1.4)\\
\mu_{i+1}\eta_i &= \text{id follows from } x1 = x
\end{aligned}$$

$$\begin{aligned}
\text{For (vi): } a_j\eta_j &= \delta\text{id follows from } E_N(1) = 1\\
\mu_{j+1}\kappa_j &= \delta\text{id follows from } \sum_{b\in B} bb^* = \delta^2\text{id} \quad (4.1.3)
\end{aligned}$$

□

Lemma 4.1.21 *If* $x\in M$, $2\le r\le k$, *then* $v_kxv_r = v_{r-2}xv_k$ *(where* $v_0 = 1$*), and* $v_kxv_1 = \delta E_N(x)v_k$.

Proof. Simple manipulation of 4.1.8 and 4.1.9. □

Lemma 4.1.22 *If* $x\in\otimes_N^k M$, *then*

(i) $\theta^{-1}E_{M_{k-2}}\theta(x) = \delta^{-1}a_{m+1}(x)$ *if* k *is odd,* $k = 2m+1$.

(ii) $\theta^{-1}E_{M_{k-2}}\theta(x) = \delta^{-1}\mu_{m+1}(x)$ *if* k *is even,* $k = 2m$.

Proof. Let x be of the form $x_1\otimes x_2\otimes\cdots\otimes x_k$. Then

$$E_{M_{k-2}}(\theta(x)) = E_{M_{k-2}}(x_1v_1x_2v_2\dots v_{k-1}x_k) = \delta^{-1}x_1v_1x_2v_2\dots x_{k-2}v_{k-2}x_{k-1}v_{k-2}x_k$$

Case (i). If $k = 2m+1$ we may apply 4.1.21 $m-1$ times to obtain

$$\begin{aligned}
E_{M_{k-2}}(\theta(x)) &= \delta^{-1}x_1v_1\dots x_mv_mx_{m+1}v_1x_{m+2}v_{m+1}\dots x_{k-1}v_{k-2}x_k\\
&= x_1v_1\dots x_mv_mE(x_{m+1})x_{m+2}v_{m+1}\dots x_{k-1}v_{k-2}x_k\\
&= \delta^{-1}\theta(a_{m+1}(x))
\end{aligned}$$

Case (ii). If $k = 2m$ we apply 4.1.21 $m-1$ times to obtain

$$\begin{aligned}
E_{M_{k-2}}(\theta(x)) &= \delta^{-1}x_1v_1\dots x_mv_mx_{m+1}v_0x_{m+2}v_{m+1}\dots x_{k-1}v_{k-2}x_k\\
&= \delta^{-1}\theta(\mu_{m+2}(x)) \quad \text{(since } v_0 = 1).
\end{aligned}$$

□

Lemma 4.1.23 *If* $y \in \otimes_N^k M$ *and* $x, z \in M$, *then*

$$xv_k^*\theta(y)v_{k+1}z = \theta(x \otimes y \otimes z)$$

Proof. Simple commutation of E_i with M and E_j's $j \leq i-2$. □

In the next lemma, let $\chi_j = a_j\mu_{j+1}$, $j = 1, 2, \dots k-1$.

Lemma 4.1.24 *If* $x, y \in \otimes_N^k M$, *then* $\theta(x)\theta(y) =$

(i) $\theta(\chi_{m+1}\chi_{m+2}\dots\chi_k(x \otimes_N y))$ *if* k *is even,* $k = 2m$.

(ii) $\theta(\mu_{m+2}\chi_{m+2}\dots\chi_k(x \otimes_N y))$ *if* k *is odd,* $k = 2m+1$.

Proof. By induction on k. Let $x = x_1 \otimes x_2 \cdots \otimes x_{k+1}$, $y = y_1 \otimes y_2 \cdots \otimes y_{k+1}$, then

$$\begin{aligned}
\theta(x)\theta(y) &= x_1 v_k^* k x_2 v_{k-1}^* \dots v_1^* x_{k+1} y_1 v_1 y_2 v_2 \dots y_k v_k y_{k+1} \\
&= \delta x_1 v_{k-1}^* E_{M_{k-2}}(\theta(x_2 \otimes x_3 \otimes \dots \otimes x_{k+1})\theta(y_1 \otimes \dots \otimes y_k))v_k y_{k+1}
\end{aligned}$$

(i) If k is even, $k = 2m$, by the inductive hypothesis we have

$$\begin{aligned}
\theta(x)\theta(y) &= \delta x_1 v_{k-1}^* E_{M_{k-2}}(\theta(\chi_{m+1}\chi_{m+2}\dots\chi_k(x_2 \otimes \dots \otimes x_{k+1} \otimes y_1 \otimes \dots \otimes y_k)))v_k y_{k+1} \\
&= x_1 v_{k-1}^* \theta(\mu_{m+1}\chi_{m+1}\dots\chi_k(x_2 \otimes \dots \otimes y_k))v_k y_{k+1} \quad \text{(by Lemma 4.1.22)} \\
&= \theta(x_1 \otimes (\mu_{m+1}\chi_{m+1}\chi_{m+2}\dots\chi_k(x_2 \otimes \dots \otimes y_k) \otimes y_{k+1})) \quad \text{(by Lemma 4.1.23)} \\
&= \theta(\mu_{m+2}\chi_{m+2}\chi_{m+3}\dots\chi_{k+1}(x_1 \otimes x_2 \otimes x_{k+1} \otimes y_1 \otimes \dots \otimes y_{k+1}))
\end{aligned}$$

(ii) If k is odd, $k = 2m+1$,

$$\begin{aligned}
\theta(x)\theta(y) &= x_1 v_{k-1}^* E_{M_{k-2}}(\theta(\mu_{m+2}\chi_{m+2}\chi_{m+3}\dots\chi_k(x_2 \otimes \dots x_{k+1} \otimes y_1 \otimes \dots \otimes y_k)))v_k y_{k+1} \\
&\quad \text{(by the induction hypothesis)} \\
&= x_1 v_{k-1}^* \theta(a_{m+1}\mu_{m+2}\chi_{m+2}\chi_{m+3}\dots\chi_k(x_2 \otimes \dots \otimes y_k))v_k y_{k+1} \\
&\quad \text{(by Lemma 4.1.22)} \\
&= \theta(x_1 \otimes \chi_{m+1}\chi_{m+2}\dots\chi_k(x_2 \otimes \dots \otimes y_k) \otimes y_{k+1})) \quad \text{(by Lemma 4.1.23)} \\
&= \theta(\chi_{m+2}\chi_{m+3}\dots\chi_{k+1}(x_1 \otimes x_2 \dots \otimes y_{k+1})) \ .
\end{aligned}$$

It only remains to check the formula for $k = 1$, $(m = 0)$. Then $\theta(x) = x$, $\theta(y) = y$ and the formula reads

$$\theta(x)\theta(y) = xy = \theta(\mu_2(x \otimes y)) = \theta(xy) \ .$$

□

Lemma 4.1.25 *For $m = 1, 2, \ldots,$ and $x_{m+1}, x_{m+2}, \ldots x_{2m} \in M$,*

$$v_m v_{m+1} x_{m+1} v_{m+2} x_{m+2} v_{m+3} \ldots v_{2m} x_{2m} = v_m x_{m+1} v_{m+1} x_{m+2} \ldots v_{2m-1} x_{2m} \ .$$

Proof. Induction on m.

For $m = 1$ the formula reads $E_1 E_2 E_1 x_2 = E_1 x_2$ which is correct. Now suppose the formula holds for m, then

$$v_{m+1} v_{m+2} x_{m+2} = v_{m+1} x_{m+2} E_{m+2} E_{m+1} \ldots E_3 \qquad \text{(by 4.1.8)}$$

so

$$v_{m+1} v_{m+2} x_{m+2} v_{m+3} x_{m+3} \ldots v_{2m+2} x_{2m+2} = (v_{m+1} x_{m+2}) V_m V_{m+1} y_{m+1} V_{m+2} y_{m+2} \ldots V_{2m} y_{2m}$$

where $V_n = E_{n+2} E_{n+1} \ldots E_3$ and $y_n = E_2 E_1 x_{n+2}$.

We may now apply the inductive hypothesis to the subfactor $M_1 \subset M_2$ (for which the E_i's are just those for $N \subset M$, shifted by 2), to obtain

$$v_{m+1} x_{m+2} V_m y_{m+1} \ldots V_{2m} y_{2m} = v_{m+1} x_{m+2} v_{m+2} x_{m+3} \ldots v_{2m+1} x_{2m+2}$$

□

Corollary 4.1.26 *With notation as above, for $m = 2, 3, 4, \ldots,$*

$$\begin{aligned} &E_m E_{m-1} \ldots E_2 (v_{m+1} x_{m+1})(v_{m+2} x_{m+2}) \ldots (v_{2m-1} x_{2m-1}) \\ &\quad = (v_m \ldots x_{m+1})(v_{m+1} x_{m+2}) \ldots (v_{2m-2} x_{2m-1}) \end{aligned}$$

Proof. Write $V_n = E_{n+1} E_{n-1} \ldots E_2$, $y_n = E_1 x_{n+1}$ and apply 4.1.25 to the subfactor $M \subset M_1$. □

Lemma 4.1.27 *For $p = 1, 2, \ldots,$*

$$\sum_{b_1, b_2, \ldots b_p \in B} (b_1 v_1)(b_2 v_2) \ldots (b_p v_p) v_{p+1} v_{p+2} b_p^* v_{p+3} b_{p-1}^* v_{p+4} \ldots v_{2p+1} b_1^* = \delta^p E_{2p+1}$$

Proof. By induction on p. For $p = 1$ the formula reads

$$\sum_{b \in B} b E_1 E_2 E_1 E_3 E_2 E_1 b^* = E_3 \sum_{b \in B} b E_1 b^* = \delta E_3 \ ,$$

which is correct. Now observe that

$$\sum_b b v_p v_{p+1} v_{p+2} b^* v_{p+3} = \delta E_p E_{p-1} \ldots E_2 V_{p-1} V_p V_{p+1} E_2 E_1$$

where $V_n = E_{n+2}E_{n+1}\dots E_3$, and that

$$b_1v_1b_2v_2\dots b_{p-1}v_{p-1}E_pE_{p-1}\dots E_2 = y_1V_1y_2V_2\dots y_{p-2}V_{p-2}y_{p-1}$$

where $y_i = b_iE_1E_2$, so that

$$\sum_{b_1,\dots,b_p\in B} b_1v_1b_2v_2\dots b_pv_pv_{p+1}v_{p+2}b_p^*v_{p+3}b_{p-1}^*v_{p+4}\dots v_{2p+1}b_1^*$$
$$= \delta \sum_{y_1\dots y_{p-1}\in BE_1E_2} y_1V_1y_2V_2\dots y_{p-2}V_{p-2}y_{p-1}V_{p-1}V_pV_{p+1}y_{p-1}^*V_{p+2}\dots V_{2p-1}y_1^*$$

so since $\{bE_1E_2 \mid b\in B\}$ is a basis for M_2 over M_1, we are through by induction. □

Corollary 4.1.28 *For* $p = 1, 2, \dots,$

$$\sum_{b_1,b_2,\dots,b_{p+1}\in B} (b_1v_1)(b_2v_2)\dots(b_pv_p)b_p^*b_{p+1}v_{p+1}b_{p+1}^*v_{p+2}b_{p-1}^*\dots v_{2p}b_1^* = \delta^{p+1}E_{2p}$$

Proof. Observe that

$$\sum_{b_p,b_{p+1}\in B} b_pv_pb_p^*b_{p+1}v_{p+1}b_{p+1}^*v_{p+2} = \delta^2V_{p-1}V_pV_{p+1}E_1\ ,$$

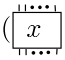
where $V_n = E_{n+1}E_n\dots E_2$, so the left-hand side of the equation becomes

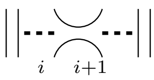
$$\delta^2 \sum_{y_1,\dots,y_{p-1}} y_1V_1y_2V_2\dots y_{p-1}V_{p-1}V_pV_{p+1}y_{p-1}^*V_{p+2}\dots V_{2p-1}y_1^*$$

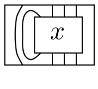
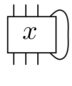
where $y_n = b_nE_1$, and this is $\delta^{p+1}E_{2p}$ by 4.1.27 applied to the subfactor $M\subset M_1$ with basis $\{bE_1 \mid b\in B\}$. □

4.2 Subfactors give planar algebras

We keep the notation of §4.1. The next theorem legitimizes the use of pictures to prove subfactor results.

Theorem 4.2.1 *Let $N\subset M$ be an extremal type II_1 subfactor with $[M:N]^{\frac{1}{2}} = \delta < \infty$. For each k let $P_k^{N\subset M} = N'\cap M_{k-1}$ (isomorphic via θ to V_k, i.e. N-central vectors in $\bigotimes_N^k M$). Then $P^{N\subset M} = \bigcup_k P_k^{N\subset M}$ has a spherical C^*-planar algebra structure (with labelling set $P^{N\subset M}$) for which* $\Phi($$) = x$ *and, suppressing the presenting map Φ,*

(i) For $i = 1, 2, \dots k-1$,  $= E_i$

(ii)  $= \delta E_{M'}(x)$,  $= \delta E_{M_{k-2}}(x)$

(iii) $\boxed{x}\ \Big|$ $=$ $\boxed{x}$ *(where on the right, x is considered as an element of M_{k+1})*

(iv) $Z(\boxed{x}) = \delta^k \mathrm{tr}(x) \quad (x \in P_k^{N\subset M})$

Moreover, any other spherical planar algebra structure Φ' with $\Phi'(\boxed{x}) = x$ and (i),(ii),(iv) *for Φ' is equal to Φ.*

PROOF. The idea of the proof is fairly simple but it will involve a lot of details, so we begin with an informal description of the idea. We must show how to associate an element $\Phi(T)$ in $N' \cap M_{k-1}$ to a tangle T whose boxes are labelled by elements of the appropriate $N' \cap M_j$. An example with $k = 5$ is given in Figure 4.2.2.

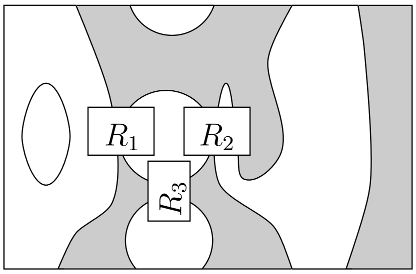


Figure 4.2.2

Shade the regions black and white and observe that a smooth oriented curve starting and ending on the left-hand boundary, and missing the internal boxes, will generically pass through a certain number of black regions. Number the connected components of the intersection of the curve with the black regions $1, 2, \dots n$ in the order they are crossed. The regions on the curve will be used to index the tensor product components in $\bigotimes_N^n M$.

We will start with our curve close to the boundary so that it crosses no black regions, and allow it to bubble outwards until it is very close to the outside boundary at which point it will cross k black regions. As the curve bubbles out, it will pass through non-generic situations with respect to the strings of the tangle, and it will envelop internal boxes. At generic times we will associate an N-central element of $\bigotimes_N^n M$ with the curve. As the curve passes through exceptional situations we will change the element of $\bigotimes_N^n M$ according to certain rules, the main one of which being that, when the curve envelops a box labelled by a tensor, we will insert that label into the tensor on the curve at the appropriate spot, as illustrated in Figure 4.2.3.

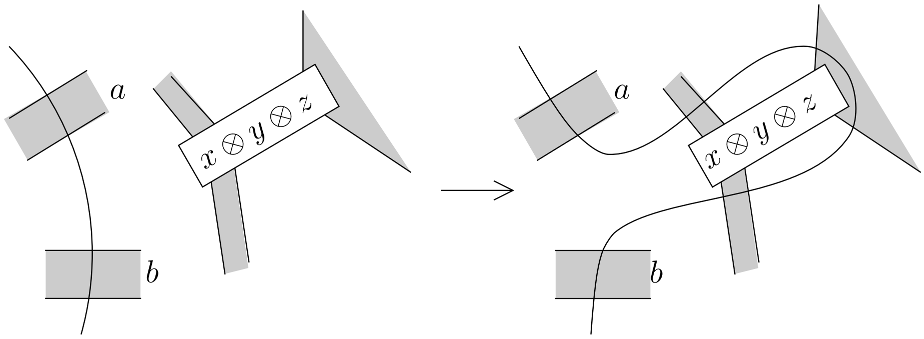


Figure 4.2.3: Tensor on bubbling curve:
$\cdots \otimes a \otimes b \otimes \ldots \ \rightarrow\ \cdots \otimes a \otimes x \otimes y \otimes z \otimes b \otimes \ldots$

When the curve arrives very close to the boundary it will have associated to it a central element of $\bigotimes_N^k M$ which gives an element of $N' \cap M_{k-1}$ via θ. This element will be $\Phi(T)$.

This strategy meets several obstacles.

(1) We must show that $\Phi(T)$ is well defined – note that the insertions of 4.2.3 are **not** well defined for the tensor product over N.

(2) $\Phi(T)$ must be central. This will require either enveloping the boxes only starting from the white region touching the first boundary point, or projecting onto central vectors at each step. We will adopt the former policy.

(3) $\Phi(T)$ must be independent of isotopy of T and the choice of the path. Since we must show isotopy invariance eventually, we might as well suppose that the tangles are in a convenient standard form, since any two ways of arriving at that standard form from a given T will only differ by an isotopy.

We begin the formal proof by describing the standard form.

A *k-picture* (or just "picture" if the value of k is clear) will be the intersection of the unit square $[0,1] \times [0,1]$ in the $x-y$ plane with a system of smooth curves, called strings, meeting only in finitely many isolated singularities, called "cusps", where $2m$ strings meet in a single point. The strings must meet the boundary of $[0,1] \times [0,1]$ transversally in just $2k$ points on the boundary line $[0,1] \times \{1\}$. A cusp (x,y) will be said to be in *standard form* if, in some neighborhood of (x,y) the

y-coordinates of points on the strings are all greater than y. A picture may always be shaded black and white with the y-axis being part of the boundary of a white region.

A picture Θ will be said to be *standard* if

(i) All its cusps are in standard form, and the region immediately below the cusp is white.

(ii) The y-coordinate, restricted to strings, has only generic singularities, i.e., isolated maxima and minima.

(iii) The y-coordinates of all cusps and all maxima and minima are distinct. This set will be written $\mathfrak{S}(\Theta) = \{y_1, y_2, y_3, \ldots, y_c\}$ with $y_i < y_{i+1}$ for $1 \leq i < c$.

An example of a standard k-picture with $k = 6$ is in Figure 4.2.4.

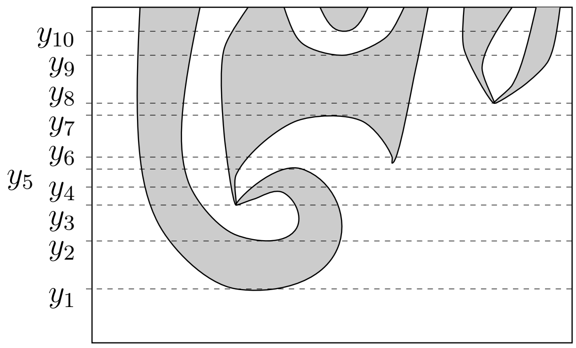
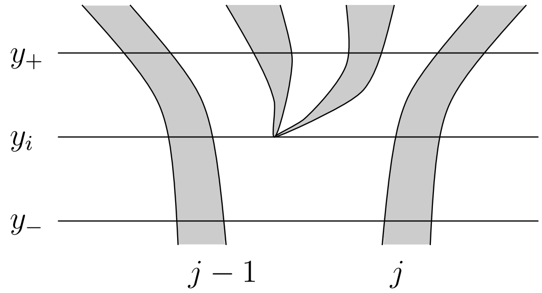


Figure 4.2.4: A standard 6-picture with $\mathfrak{S}(\Theta) = \{y_1, y_2, y_3, \ldots, y_{10}\}$.

A standard k picture Θ will be *labelled* if there is a function from the cusps of Θ to $\coprod_m V_m$ so that a cusp where $2m$ strings meet is assigned an element of V_m (or $N' \cap M_{m-1}$, via θ). We now describe how to associate an element Z_Θ of $N' \cap M_{k-1}$ to a labelled standard k-picture Θ, using the operators of definition 4.1.19.

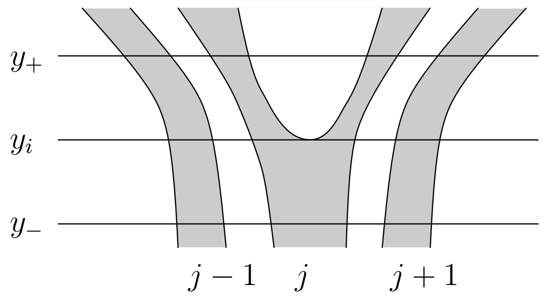
Let Θ be a labelled standard k-picture with $\mathfrak{S}(\Theta) = \{y_1, y_2, y_3, \ldots, y_c\}$. We define a locally constant function $Z : [0,1] \backslash \mathfrak{S}(\Theta) \to V_{k(y)}$, where $k(y)$ is the number of distinct intervals of $[0,1] \times \{y\}$ which are the connected components of its intersection with all the black regions of Θ. For instance, if Θ is as in Figure 4.2.4 and $y_8 < y < y_9$ then $k(y) = 4$. Obviously $k(y)$ is locally constant and $k(y) = k$ for $y_c < y < 1$. For $y < y_1$, $V_0 = \mathbb{C}$ and we set $Z(y) = 1$. There are five possible ways for Z to change as y goes from a value y_-, just less than y_i, to y_+, just bigger than y_i. We define $Z(y_+)$ from $Z(y_-)$ in each case:

Case (i). y_i is the y-coordinate of a cusp, between the $(j-1)^{\text{th}}$ and j^{th} connected components of the intersection of $[0,1] \times y_-$ with the black regions, as below



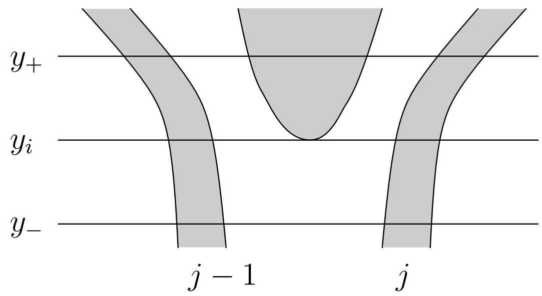
Set $Z(y_+) = \alpha_{j,c}(Z(y_-))$ where c is the label (in $V_{k(y_-)}$) associated with the cusp.

Case (ii). y_i is the y-coordinate of a minimum, with numbering and shading as below:



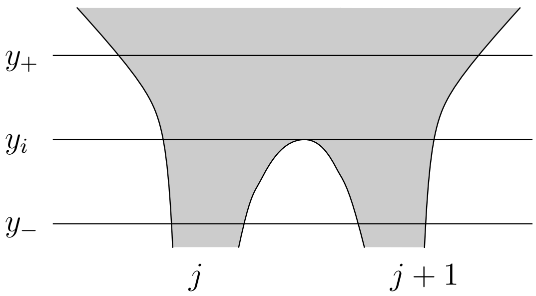
Set $Z(y_+) = \kappa_j(Z(y_-))$.

Case (iii). y_i is the y-coordinate of a minimum, with numbering and shading as below:



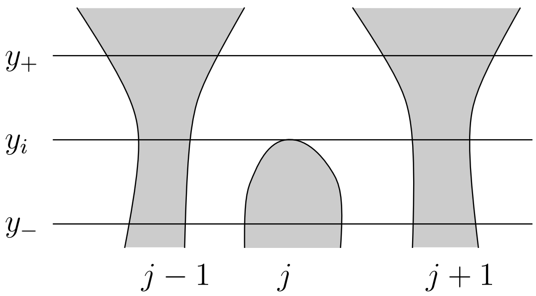
Set $Z(y_+) = \eta_j(Z(y_-))$.

Case (iv). y_i is the y-coordinate of a maximum, with numbering and shading as below:



Set $Z(y_+) = \mu_{j+1}(Z(y_-))$.

Case (v). y_i is the y-coordinate of a maximum with shading as below:



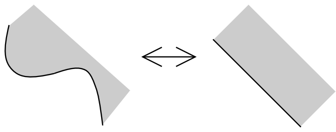
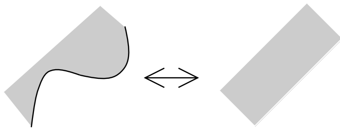
Set $Z(y_+) = a_j(Z(y_-))$.

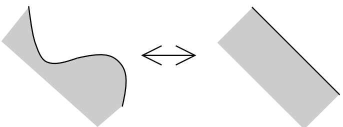
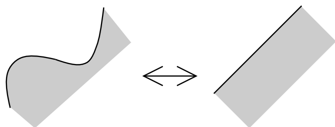
Finally, we define $Z_\Theta = \theta(Z(y))$ for $y > y_c$, also written $Z_1(\Theta)$. Our main job is now to prove that Z_Θ is unchanged if Θ is changed by isotopy to another standard picture Θ', with labels transported by the isotopy. If the isotopy passes only through standard pictures, critical points can never change order or be annihilated or created. The pattern of connected components of the intersection of the black regions with horizontal lines cannot be changed either as such a change would have to involve two maxima, minima or cusps having the same y coordinate. So isotopies through standard pictures do not change Z_Θ.

We next argue that if ϕ_t, $0 \leq t \leq 1$ is an isotopy which preserves the standard form of each cusp, then $Z_\Theta = Z_{\phi_1(\Theta)}$. For now $Z_{\phi_t(\Theta)}$ can only change if the singularities of the y-coordinate function change. By putting the isotopy in general position we see that this can be supposed to happen in only two ways (see e.g. [Tu], or note that this argument can be made quite combinatorial by using piecewise linear strings).

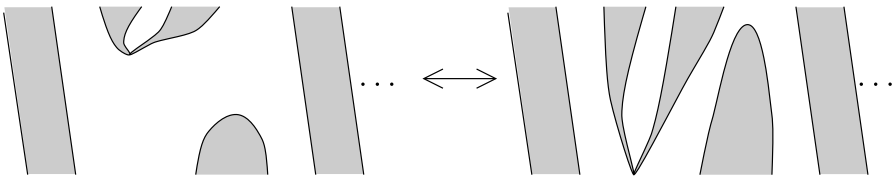
(1) The y-coordinates of two of the singularities coincide then change order while the x-coordinates remain distinct.

(2) The y-coordinate along some string has a point of inflection and the picture, before and after, looks locally like one of the following

(A)  (C) 

(B)  (D) 

In case 2), invariance of Z_Θ is guaranteed by (v) of 4.1.20 (in the order C, D, B, A) and in case 1), (i)→(iv) of 4.1.20 is a systematic enumeration of all 25 possibilities. For instance, the case



is covered by the first equation of (iii) of 4.1.20 with $j = i$. Thus Z_Θ is invariant under isotopies preserving standardness of the cusps.

Now we argue that a general isotopy ϕ_t may be replaced by a $\tilde{\phi}_t$ for which $\tilde{\phi}_t$ preserves the standard form of cusps for all t, without changing Z_Θ. To see this, construct a small disc around each cusp of Θ, sufficiently small so that the y-coordinates of all points in a given disc are distinct from those in any other disc and distinct from any maxima or minima of y on the strings, and such that the same is true for the images of these discs under ϕ_1. (Remember that $\phi_1(\Theta)$ is also a standard picture.) Now in each disc D, construct a smaller disc D_0 inside D, centered at the cusp, sufficiently small so that one can construct a new isotopy $\tilde{\phi}_t$ having the properties

(i) $\tilde{\phi}_t$ restricted to each D_0 is just translation in the plane (and $\tilde{\phi}_t$ (a cusp) $= \phi_t$ (that cusp)).

(ii) $\tilde{\phi}_t = \phi_t$ on the complement of the discs D.

Thus inside $\tilde{\phi}_t(D_0)$, the cusp remains standard and $\tilde{\phi}_t$ is extended somehow to the annular region between $\tilde{\phi}_t(D_0)$ and $\phi_t(D)$. But the mapping class group of diffeomorphisms of the annulus that are the identity on the boundary is generated by a Dehn twist of 360°. So in a neighborhood of each cusp point, ϕ_1 and $\tilde{\phi}_1$ differ only by some integer power of a single full twist. Figure 4.2.5 illustrates how $\phi_1(\Theta)$ and $\tilde{\phi}_1(\Theta)$ would differ if the twist incurred were a single clockwise twist.

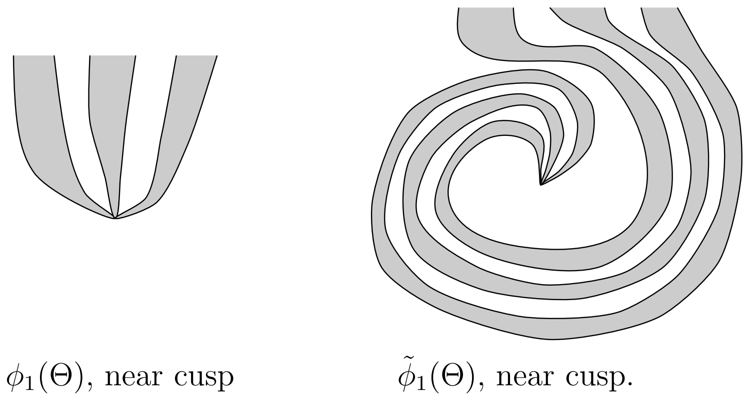


Figure 4.2.5

We want to show, first in this case and then in the case of an arbitrary integral power of a full twist, that Z_Θ is unchanged. For this, consider Figure 4.2.6 which is supposed to be part of a standard labelled picture in which the maxima and minima of y, in the figure, and its cusps occur as an uninterrupted sequence in $\mathfrak{S}(\Theta)$:

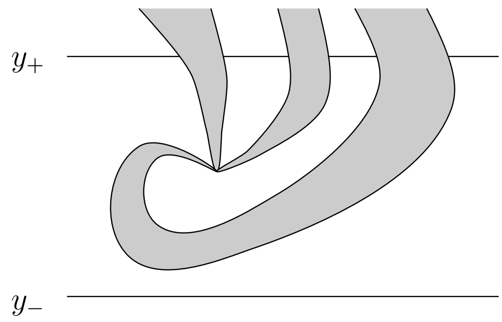


Figure 4.2.6

From the definition of $Z(y)$ we see that

$$\begin{aligned} Z(y_+) &= a_j \mu_{j+1} \alpha_{j+1,c} \kappa_j \eta_j (Z(y_-)) \\ &= \alpha_{j,\rho(c)}(Z(y_-)) \quad \text{(by 4.1.13 and 4.1.14)} \end{aligned}$$

(Here the cusp is labelled by $c \in N' \cap M_{n-1}$ where $n = 3$ in Figure 4.2.6.) Similarly we see that if the cusp is surrounded by a full 360° twist,

$$\begin{aligned} Z(y_+) &= \alpha_{j,\rho^n(b)}(Z(y_-)) \\ &= \alpha_j, b(Z(y_-)) \qquad \text{(by Theorem 4.1.18)} \end{aligned}$$

Thus if a cusp (or any part of a picture that is just a scaled down standard labelled picture) is surrounded by a single clockwise full twist, the effect on Z_Θ is as if the twist were not there. If there were anticlockwise full twists around a cusp, surround it further by the same number of clockwise twists. This does not change Z_Θ, but the cancelling of the positive and negative twists involves only isotopies that are the identity near the cusps. Thus by our previous argument clockwise full twists around cusps do not change Z_Θ either. We conclude

$$Z_{\phi_1(\Theta)} = Z_{\tilde{\phi}_1(\Theta)} = Z_\Theta \ .$$

We have established that Z_Θ may be assigned to a standard labelled picture by a product of elementary maps $\alpha, \kappa, \eta, a, \mu$ in such a way that Z_Θ is unchanged by isotopies of Θ. We can now formally see how this makes $\coprod_n V_n$ a planar algebra according to §1. The labelling set will be $\coprod_n V_n$ itself. The first step will be to associate a labelled picture $\beta(T)$ with a labelled tangle T. To do this, shrink all the internal boxes of T to points and isotope the standard k-box to $[0,1] \times [0,1]$ with all the marked boundary points going to points in $[0,1] \times 1$. Then distort the shrunk boxes of T to standard cusps, by isotopy, so that the string attached to the first boundary point of the box becomes the first string (from the left) attached to the cusp. The procedure near a 4-box of T is illustrated in Figure 4.2.7:

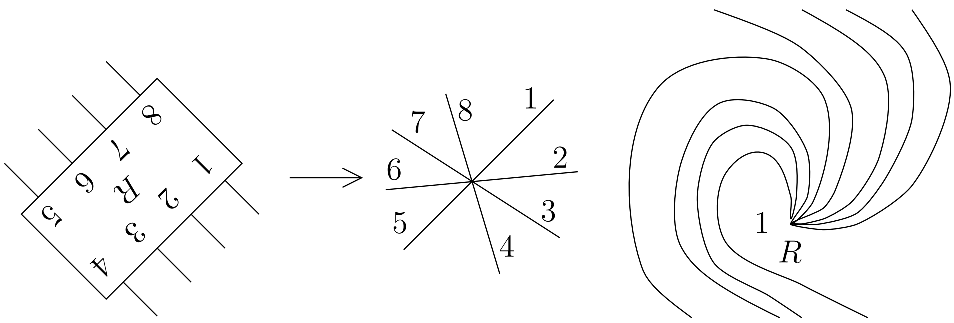
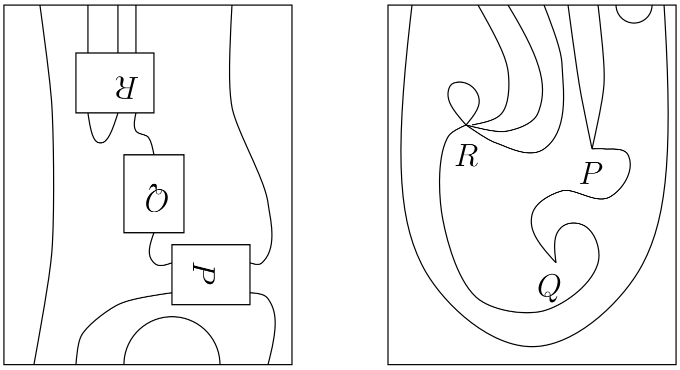


Figure 4.2.7

The label associated to the cusp is just θ^{-1} of the label associated to the box but we may reasonably suppress θ^{-1}.

Figure 4.2.8 illustrates a labelled tangle T and a labelled standard picture $\beta(T)$:

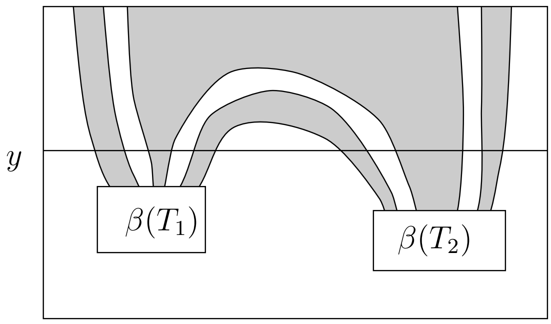


The tangle T A standard picture $\beta(T)$

Figure 4.2.8

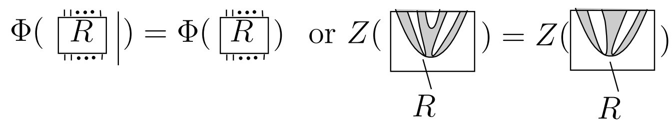
Note that $\beta(T)$ is not well defined, but two different choices of $\beta(T)$ for a given T will differ by an isotopy so the map Φ, $\Phi(T) = Z_{\beta(T)}$, gives a well defined linear map from the universal planar algebra on $\coprod_k N' \cap M_{k-1}$.

We now check that Φ makes $(P^{N\subset M})$ into a connected spherical C^*-planar algebra. The first thing to check is that Φ is a homomorphism of filtered algebras. But if T_1 and T_2 are labelled k-tangles, a choice of $\beta(T_1T_2)$ is shown below ($k = 3$).



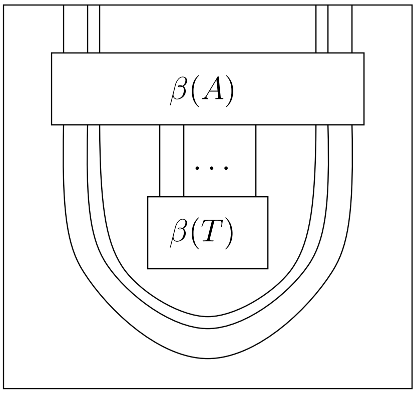
If y is as marked, from the definition of Z, $Z(y) = Z(\beta(T_1)) \otimes Z(\beta(T_2))$. Moreover each pair of maximal  contributes a factor χ_i $(= a_i\mu_{i+1})$ to $Z(y)$ as y increases, so if k is even, part (1) of 4.1.24 gives $\Phi(T_1T_2) = \Phi(T_1)\Phi(T_2)$ and if k is odd (as in the figure) the last maximum has the black region above so contribute a factor μ and part (ii) of 4.1.24 applies.

That Φ is compatible with the filtrations amounts to showing that



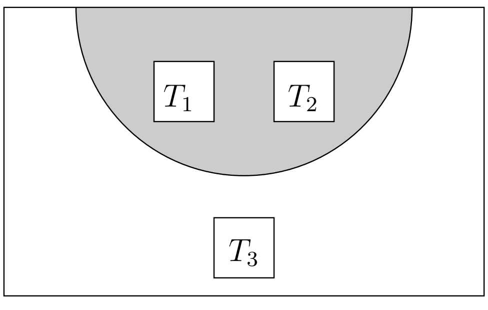
If k is even this follows from 4.1.25 and if k is odd it follows from 4.1.26 (together with $\sum_{b\in B} bv_k b^* = \delta E_k E_{k-1}\ldots E_2$ to take care of the factor κ introduced by the minimum in the picture). So Φ is a homomorphism of filtered algebras.

Annular invariance (and indeed the whole operadic picture) is easy. If T is an element of $\mathcal{P}(\coprod_k N' \cap M_{k-1})$ (linear combination of tangles) with $\Phi(T) = 0$, then if T is surrounded by an annular labelled tangle A, then we may choose $\beta(\pi_A(T))$ to look like

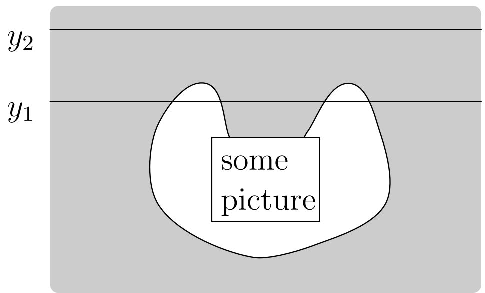


(Strictly speaking, one needs to consider such a picture for each tangle in the linear combination forming T, and add.) Clearly if $Z_{\beta(T)} = 0$, so is $Z_{\beta(\pi_A(T))}$, since the map $\alpha_{j,\theta^{-1}(Z_{\beta(T)})}$ is applied in forming $Z_{\beta(\pi_A(T))}$.

We now turn to planarity. By definition $V_0 = \mathbb{C}$ so we only need to show dim $V_{1,1} = 1$. A basis element of $\mathcal{P}_{1,1}(\coprod_{k=1}^{\infty} N' \cap M_{k-1})$ is a 1-tangle T with a vertical straight line and planar networks to the left and right. We may choose $\beta(T)$ to be as depicted below

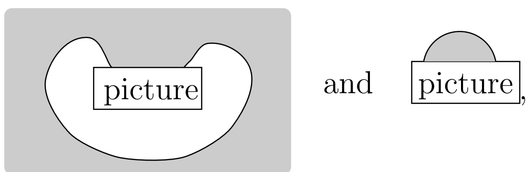


where there are 0-pictures inside the regions $T_1, T_2, \ldots$. It is a simple consequence of our formalism that a closed picture surrounded by a *white* region simply contributes a scalar in a multiplicative way. This is because one may first isotope the big picture so that all the maxima and cusps in the 0-picture have y coordinates in an uninterrupted sequence in $\mathfrak{S}$, and the last singularity must be a maximum, shaded below, the first being a minimum, shaded above. The final map will be an a_j and will send the contribution of the 0-picture to an element of $N' \cap N = \mathbb{C}$. Thus we only need to see that 0-pictures inside a black region contribute a scalar in a multiplicative way. But we may isotope the big picture so that the singular y-values of the 0-picture occur in uninterrupted succession, and near the 0-picture the situation is as below:



With y_1 as marked, $Z(y_1)$ will be $\sum_{b \in B} b \otimes x \otimes b^*$ for some element x in $N' \cap M$. But then $Z(y_2)$ will be $\sum_{b \in B} bxb^* \in M' \cap M = \mathbb{C}$, by 4.1.5. So $P^{N \subset M}$ is a planar algebra.

The spherical property is easy: comparing



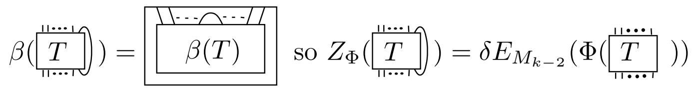
we see that the partition functions are the same since one gives $E_{M'}$ applied to an element of $N' \cap M$, and the other gives E_N applied to the same element with the correct powers of δ contributed from the minimum in the first picture and the maximum in the second. Either way we get the trace by extremality.

For the $*$-structure, observe first that

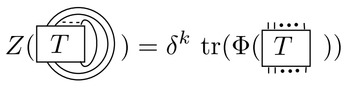
$$\theta(x_1 \otimes x_2 \cdots \otimes x_k)^* = \theta(x_k^* \otimes x_{k-1}^* \otimes \cdots \otimes x_1^*)$$

so $\theta(\alpha_{j,c}(x))^* = \theta(\alpha_{k-j+2,c^*}(x^*))$, $\theta(a_j(x))^* = \theta(a_{k-j+1}(x^*))$, $\theta(\mu_j(x))^* = \theta(\mu_{n-j+2}(x^*))$, $\theta(\eta_j(x))^* = \theta(\eta_{k-j+2}(x^*))$ and $\theta(\kappa_j(x))^* = \theta(\kappa_{k-j+1}(x^*))$. Moreover if T is a labelled tangle, $\beta(T^*)$ is $\beta(T)$ reflected in the line $x = \frac{1}{2}$ and with labels replaced by their adjoints (via θ). We conclude that $(Z_{\beta(T)})^* = Z_{\beta(T^*)}$ by applying the relations above at each of the y-values in $\mathfrak{S}(\beta(T)) = \mathfrak{S}(\beta(T^*))$.

The C^*-property is just the positive definiteness of the partition function. But if T is a labelled k-tangle,



by 4.1.22. Applying this k times we get



so the positive definiteness of Z follows from that of tr.

We now verify (i)→(iv) in the statement of the theorem.

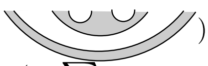
(i) Since Φ is a homomorphism of filtered algebras, it suffices to prove the formula when $k = i$.
If $k = 2p+1$,

$Z_{\beta(|\,|\ldots\ )} = \theta(Z_1($  $))$

$$= \delta^{-P}\theta\Big(\sum_{b_1,b_2,\ldots,b_p} b_1 \otimes b_2 \otimes \cdots \otimes b_p \otimes 1 \otimes 1 \otimes b_p^* \otimes \cdots \otimes b_2^* \otimes b_1^*\Big)$$

$$= E_{2p+1} \qquad \text{by Lemma 4.1.27.}$$

If $k = 2p$,

$Z_{\beta(|\,|\ldots\ )} = \theta(Z_1($  $))$

$$= \delta^{-p-1}\theta\Big(\sum_{b_1,\ldots,b_{p+1}} b_1 \otimes b_2 \otimes \cdots \otimes b_p \otimes b_p^* b_{p+1} \otimes b_{p+1}^* \otimes b_{p-1}^* \otimes \cdots \otimes b_1^*\Big)$$

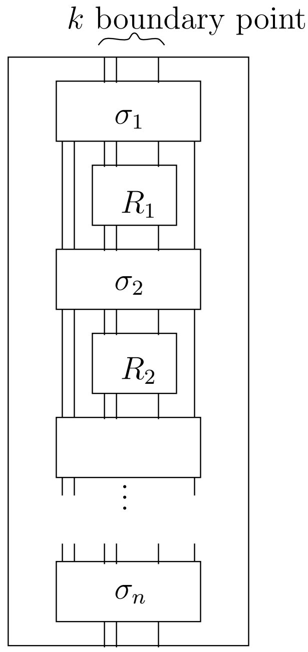
$$= E_{2p} \qquad \text{by Corollary 4.1.28.}$$

(ii) The first formula follows from 4.1.5 and we showed the second when we proved the positive definiteness of the partition function.

(iii) This is just the filtered algebra property.

(iv) We also showed this in the positive definiteness proof.

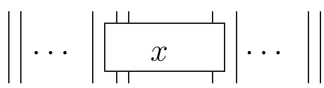
All that remains is to prove the uniqueness of the planar algebra structure. First observe that, as in Proposition 1.14, a labelled tangle may be arranged by isotopy so that all of its boxes occur in a vertical stack. After further isotopy and the introduction of kinks or redundant loops, one may obtain the picture below for the tangle (in $\mathcal{P}_k$)

$T =$ 

where the regions marked $\sigma_1, \ldots, \sigma_n$ contain only strings and $\sigma_2, \ldots, \sigma_{n-1}$ have a fixed number p of boundary strings top and bottom, $\geq k$. Clearly $p - k$ is even so we conclude that, if Φ_1 is some other planar algebra structure satisfying (i) and (ii) then

$$\Phi_1(T) = \delta^{-\frac{p-k}{2}}\ \Phi_1\Big(\boxed{\ T\ }\Big)\ ,$$

where we have introduced $\frac{p-k}{2}$ maxima and minima. (To see this just apply the second formula of (ii) $p - k$ times.) Thus we find that it suffices to prove that $\Phi' = \Phi$ on a product of Temperley-Lieb tangles and tangles of the form

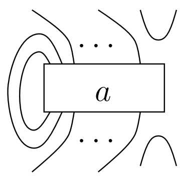


But the Temperley-Lieb algebra is known to be generated by $\{\|\ldots \underset{i\quad i+1}{\overset{\cup}{\cap}} \|\}$ whose images are the E_i's by (i). By condition (ii), we see that it suffices to show that

$\Phi'(||\boxed{x}) = \Phi(||\boxed{x})$. To this end we begin by showing

$$\Phi'(\,|\,|\boxed{a}) = \Phi(\,|\,|\boxed{a}) \quad \text{for } a \in N' \cap M_k\ .$$

This follows from the picture below

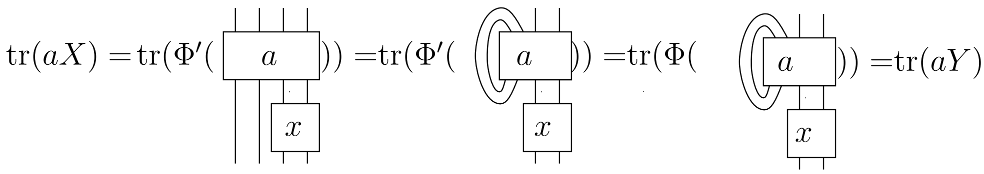


for we know that

$$\Phi'(|\boxed{a}) = E_{M'}(a) = \Phi(|\boxed{a}).$$

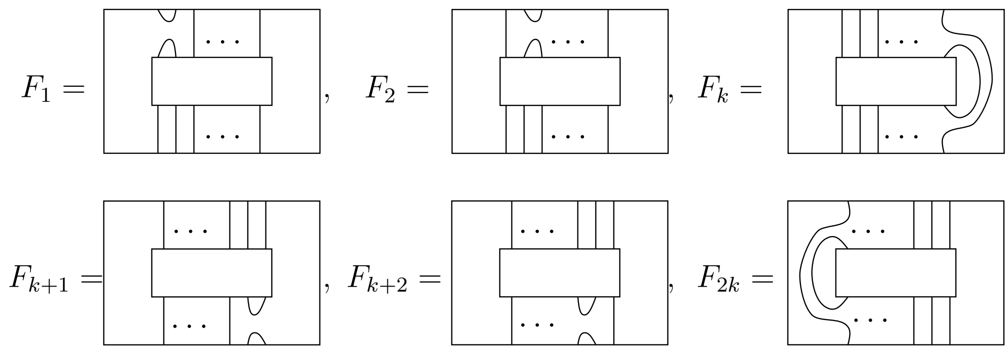
Now to show that $X = \Phi'(||\boxed{x}) = \Phi(||\boxed{x}) = Y$ it suffices to show that

$\mathrm{tr}(aX) = \mathrm{tr}(aY)$ for all $a \in N' \cap M_k$. But up to powers of δ,



and we are done. □

Definition 4.2.8. The *annular Temperley Lieb* algebra $AT(n, \delta)$, for n even, will be the $*$-algebra with presentation:



Remark. Since $(F_1F_2 \ldots F_k)(F_1F_2 \ldots F_k)^* = \delta F_1$ and $(F_1F_2 \ldots F_k)^*(F_1F_2 \ldots F_k) =$

δF_k, if $\mathcal{H}$ is a Hilbert space carrying a $*$-representation of $AT(n, \delta)$, $\dim(F_i\mathcal{H})$ is independent of i.

Corollary 4.2.9 *If $N \subset M$ is an extremal subfactor of index $\delta^{-2} > 4$, each $N' \cap M_{k-1}$ is a Hilbert space carrying a $*$-representation of $AT(2k, \delta)$. And* $\dim(F_i(N' \cap M_{k-1})) = \dim(N' \cap M_{k-2}) = \dim(M' \cap M_{k-1})$.

Proof. Let F_i be the elements of $\mathcal{A}(\phi)$ defined as follows
The relations are easily checked, the Hilbert space structure on $N' \cap M_{k-1}$ being given by the trace $-\langle a, b\rangle = \mathrm{tr}(b^* a)$. Since $F_k(N' \cap M_{k-1}) = N' \cap M_{k-2}$ and $F_{2k}(N' \cap M_{k-1}) = M' \cap M_{k-1}$ (by (ii) of 4.2.1). We are through. $\square$

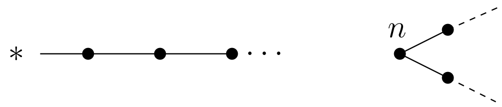
Lemma 4.2.10 *If $\mathcal{H}$ carries an irreducible $*$-representation of $AT(n, \delta)$ for $\delta > 2$, $n > 4$, and* $\dim(F_i\mathcal{H}) = \infty$, *then* $\dim \mathcal{H} = n$ *(remember n is even).*

Proof. Let v_i be a unit vector in $F_i\mathcal{H}$ for each i. Then $F_j v_i$ is a multiple of v_j so the linear span of the v_i's is invariant, thus equal to $\mathcal{H}$ by irreducibility. Here $\dim \mathcal{H} \leq n$ (this does not require $\delta > 2$). Moreover the commutation relations imply $|\langle v_i, v_{i+1}\rangle| = \delta^{-1}$ and $\langle v_i, v_j\rangle = 0$ or 1 if $i \neq j \pm 1$. The case $\langle v_i, v_j\rangle = 1$ forces $i = j$ so it only happens if $n = 4$. So by changing the v_i's by phases we may assume that the matrix $\delta\langle v_i, v_j\rangle$ is

$$\Delta_n(\omega) = \begin{pmatrix} \delta & 1 & 0 & 0\ldots\omega \\ 1 & \delta & 1 & 0\ldots 0 \\ 0 & 1 & \delta & 1\ldots 0 \\ \vdots & & & \vdots \\ \bar{\omega} & 0 & \ldots & 1 \;\; \delta \end{pmatrix}.$$

It is easy to check that $\det(\Delta_n(\omega)) = P_{2n}(\delta) - P_{2n-2}(\delta) - 2\mathrm{Re}(\omega)$ where $P_n(\delta)$ are Tchebychev polynomials. Thus $\det(\Delta_n(\omega))$ is smallest, for fixed δ, when $\omega = 1$. But then
$\|\Delta_n(1) - \delta\ \mathrm{id}\| = 2$ by Perron-Frobenius so $\det \Delta_n(1) > 0$ for $\delta > 2$. $\square$

Corollary 4.2.11 *Suppose the principal graph of the subfactor $N \subset M$, $[M : N] > 4$, has an initial segment equal to the Coxeter-Dynkin diagram D_{n+2} with $*$ as shown:*



Then there are at least two edges of the principal graph connecting the two points at distance $n+1$ from $*$ to points of distance $n+2$.

Proof. Since $\delta > 2$, the Temperley-Lieb algebra generated by $\{1, e_1, e_2, \ldots, e_k\}$ in $N' \cap M_k$ has dimension $\frac{1}{k+2}\binom{2k+4}{k+1}$. The information on the principal graph then gives the following Bratteli diagram for the inclusions $N' \cap M_{n-1} \subset N' \cap M_n \subset N' \cap M_{n+1}$:

$$\begin{array}{ccc} N' & \cap & M_{n+1} \\ & \cup & \\ N' & \cap & M_n \\ & \cup & \\ N' & \cap & M_{n-1} \end{array}$$

Here we have shown only that part of the Bratteli diagram relevant to the proof. Only the two 1's in the middle row can be connected to anything in the top row other than vertices corresponding to the ideal generated by e_{n+1}. We have to show that it is impossible for just one of these 1's to be connected, with multiplicity one, to a new principal graph vertex. By contradiction, suppose this were the case. Then we would have

$$\dim(N' \cap M_{n+1}) = \frac{1}{n+3}\binom{2(n+2)}{n+2} - (n+1)^2 + (n+2)^2$$

since the only difference between the $N' \cap M_{n+1}$ level of the Bratteli diagram and the Temperley-Lieb Bratteli diagram (see [GHJ]) is that the "$n+1$" in Temperley-Lieb has become "$n+2$". Thus

$$\dim(N' \cap M_{n+1}) = \frac{1}{n+3}\binom{2(n+2)}{n+2} + 2n+3\ .$$

But consider $N' \cap M_{n+1}$ as a module over $AT(2n+4, \delta)$. The Temperley-Lieb subalgebra is invariant and so therefore is its orthogonal complement $TL^{\perp}$ of dimension $2n+3$. But consider the image of $F_{n+2} = \delta E_{N' \cap M_n}$. It is $\dfrac{1}{n+2}\dbinom{2(n+1)}{n+1} + 1$ because of the single extra vertex on the principal graph at distance $n+1$ from $*$. But, by pictures, the image of F_{n+2} restricted to the Temperley-Lieb subalgebra of $N' \cap M_{n+1}$ is $\dfrac{1}{n+2}\dbinom{2(n+1)}{n+1}$. Hence on $TL^{\perp}$, $\dim(F_1(TL^{\perp}) = 1$. So by Lemma 4.2.10, $\dim(TL^{\perp}) \geq 2n+4$, a contradiction. □

We have obtained far more powerful results than the following by a study of the representation theory of $AT(n, \delta)$. These results will be presented in a future paper of this series. We gave the result here because it was announced some time ago. It is a version of the "triple point obstruction" of Haagerup and Ocneanu (see [Ha]) but proved by a rather different method!

If we apply the argument we have just given when $\delta \leq 2$ we obtain nontrivial but known results. The argument is very simple so we present it.

Definition 4.2.12 The *critical depth* of a planar algebra P will be the smallest k for which there is an element in P_k which is not in the Temperley-Lieb subalgebra TL_k.

In the C^*- case, if $\delta < 2$,the norm of the principal graph is less than 2 so as in [GHJ] it follows that the principal graph is an A,D or E Coxeter graph with $*$ as far as possible from a vertex of valence 3. In particular if k is the critical depth, the dimension of the quotient $\frac{P_k}{V_k}$ is at most 1 so the rotation acts on it by multiplication by a k-th root of unity. We will use the term "chirality" for this root of unity in an appropriate planar algebra.

The restrictions on the principal graph in the following theorem were first obtained by Ocneanu.

Theorem 4.2.13 *If P is a C^*-planar algebra with $\delta < 2$ then the principal graph can be neither D_n with n odd nor E_7. If the principal graph is D_{2n} the chirality is -1, if it is E_6 the chirality is $e^{\pm 2\pi i/3}$ and if it is E_8 the chirality is $e^{\pm 2\pi i/5}$.*

PROOF. If k is the critical depth, by drawing diagrams one sees that the ω in the $(2k+2) \times (2k+2)$ matrix $\Delta_n(\omega)$ is the chirality. Also if $\delta = z + z^{-1}$, we have $\det(\Delta_n(\omega)) = z^{2k+2} + z^{-(2k+2)} - \omega - \omega^{-1}$. If κ is the Coxeter number of the principal graph we have $z = e^{\pm \pi i/\kappa}$.

On the other hand, by the argument of Corollary 4.2.11, the dimension of P_{k+1} would be too great if the determinant were non-zero. Thus we have, whatever the Coxeter graph may be,

$$e^{(2k+2)\pi i/\kappa} + e^{-(2k+2)\pi i/\kappa} = \omega + \omega^{-1}$$

for a k-th root of unity ω.

The critical depth for D_m would be $m-2$ so the left hand side of the equation is -2 so that ω has to be -1. But if m is odd, -1 is not an $(m-2)^{\text{th}}$ root of unity so D_m cannot be a principal graph. If m is even we conclude that the chirality is -1.

The critical depth for E_7 is 4 and the Coxeter number is 18. The above equation clearly has no solution ω which is a fourth root of unity.

For E_6 the critical depth is 3 and the Coxeter number is 12 so $\omega = e^{\pm 2\pi i/3}$. For E_8 the critical depth is 5 and the Coxeter number is 30 so $\omega = e^{\pm 2\pi i/5}$.

$\square$

The above analysis may also be carried out for $\delta = 2$ where the Coxeter graphs are replaced by the extended Coxeter graphs. In the D case the presence of the Fuss Catalan algebra of example 2.3 makes it appropriate to replace the notion of critical depth by the first integer such that P_k is bigger than the Fuss Catalan algebra. One obtains then that the chirality, together with the principal graph, is a complete invariant for C^*-planar algebras with $\delta = 2$ (see [EK] p. 586).

We end this section by giving more details of the planar structure on $N \subset M$. In particular we give the subfactor interpretations of duality, reduction, cabling and tensor product. The free product for subfactors is less straightforward.

Corollary 4.2.14 *If $N \subset M$ is an extremal II_1 subfactor then, with the notation of §3.2, $\lambda_1(P^{N\subset M}) = P^{M\subset M_1}$, as planar algebras. (Note that $P_k^{M\subset M_1} = M' \cap M_k$ which is a subset of $P_{k+1}^{N\subset M}$. We are saying that the identity map is an isomorphism of planar algebras.)*

Proof. Equality of $P_k^{N\subset M}$ and $\lambda_1(P^{N\subset M})$ as sets follows immediately from (ii) of 4.2.1. To show equality of the planar algebra structure we use the uniqueness part of 4.2.1.

By definition of $\lambda_1(\Phi)$, for $x \in M' \cap M_k \subset N' \cap M_k$,

$$x = \Phi^{M\subset M_1}(\boxed{x}) = \Phi^{N\cap M}(\boxed{x}) = \lambda_1(\Phi)(\boxed{x}).$$

Properties (i) and (iv) are straightforward as is the second equation of (ii). So we only need to check the first equation of (ii), i.e.,

$$\lambda_1(\Phi^{N\subset M})(\,\big|\,\boxed{x}\,) = \delta E_{M_1}(x) \text{ for } x \in M' \cap M_k.$$

But by definition

$$\lambda_1(\Phi^{N\subset M})(\,\big|\,\boxed{x}\,) = \frac{1}{\delta}\Phi(\,\big|\;\big|\,\boxed{x}\,).$$

If we define $g : \otimes_N^k M \to \otimes_N^k M$ and $f : \otimes_N^k M \to \otimes_N^{k-2} M$ by $g(y) = \sum_{b\in B} byb^*$ and $f(x_1 \otimes \cdots \otimes x_k) = E(x_1)x_2 \otimes \cdots \otimes x_{k-1}E(x_k)$, and if $\theta(y) = x$ for $x \in M' \cap M_k$, by the definition of $\Phi^{N\subset M}$ in Theorem 4.2.1,

$$\Phi^{N\subset M}(\,\big|\;\big|\,\boxed{x}\,) = \theta(\sum_{b\in B} b \otimes f(g(y)) \otimes b^*).$$

But since θ is an $M - M$ bimodule map and x commutes with M, $g(y) = \delta^2 y$. But by definition of θ and 4.1.8,

$$\theta(\sum_{b\in B} b \otimes f(y) \otimes b^*) = \delta^{-2} \sum_{b\in B} bE_1xE_1b^* \ .$$

Since $\{bE_1\}$ is a basis for M_1 over M, we are done by 4.1.5. $\square$

Iterating, we see that $\lambda_n(P^{N\subset M})$ is the planar algebra for the subfactor $M_{n-1} \subset M_n$.

For cabling we have the following

Corollary 4.2.15 *If $N \subset M$ is an extremal II_1 subfactor, the cabled planar algebra $\mathcal{C}_n(P^{N\subset M})$ of* §3.3 *is isomorphic to $P^{N\subset M_{n-1}}$.*

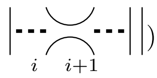
PROOF. We will again use the uniqueness part of 4.2.1. It follows from [PP2] that if we define E_i^n to be

$$(E_{ni}E_{ni-1}E_{ni-2}\dots E_{n(i-1)+1})(E_{ni+1}E_{ni}\dots E_{n(i-1)+2})\dots(E_{n(i+1)-1}E_{n(i+1)}\dots E_{ni})$$

(a product of n products of E's with indices decreasing by one), and v_i^n with respect to E_i^n as in 4.1.9, then the map $x_1 \otimes_N x_2 \cdots \otimes_N x_k \to x_1 v_1^n x_2 v_2^n \dots v_k^n x_k$ establishes, via the appropriate θ's, a $*$-algebra isomorphism between the $(k-1)$-th algebra in the tower for $N \subset M_{n-1}$ and M_{kn-1}, hence between $P_k^{N\subset M_{n-1}}$ and $P_{nk}^{N\subset M} = \mathcal{C}_n(\mathcal{P}^{\mathcal{N}\subset\mathcal{M}})$.

If γ is the inverse of this map, $\gamma \circ \mathcal{C}_n(\Phi)$ thus defines a spherical planar algebra structure on $P^{N\subset M_{n-1}}$. The labelling set is identified with $P^{N\subset M_{n-1}}$ via γ, so

$$\gamma \circ \mathbb{C}_n(\Phi^{N\subset M})(\boxed{x}) = \Phi^{N\subset M})(\boxed{x}) = x.$$

Condition (i) of 4.1 for $\gamma \circ \mathbb{C}_n(\Phi)$ follows by observing that $\mathbb{C}_n(\Phi)($$) =$ E_i^n. Condition (ii) follows from 4.2.14, and condition (iv) is clear. $\square$

Reduction is a little more difficult to prove.

Corollary 4.2.16 *Let $N \subset M$ be an extremal II_1 subfactor and p a projection in $N' \cap M$. The planar algebra $p(P^{N\subset M})p$ of §3.3 is naturally isomorphic to the planar algebra of the reduced subfactor $pN \subset pMp$.*

PROOF. We first claim that the tautological map

$$\alpha : \bigotimes_{pN}^{k} pMp \to \bigotimes_{N}^{k} M \ , \qquad \alpha(\bigotimes_{i=1}^{k} x_i) = \bigotimes_{i=1}^{k} x_i$$

is an injective homomorphism of $*$-algebras when both domain and range of α are equipped with their algebra structures via the respective maps θ as in 4.1.10. To see this observe that the conditional expectation $E_{pN} : pMp \to pN$ is just $\frac{1}{\mathrm{tr(p)}}E_N$. It follows that $\alpha a_j = a_j \alpha$ and clearly $\alpha\mu_j = \mu_j\alpha$, so by 4.1.24, α is a $*$-algebra homomorphism.

The next thing to show is that α takes pN-central vectors to $\theta^{-1}(p_k(N' \cap M_{k-1})p_k)$, with p_k as in 3.3, using the planar algebra structure on $N' \cap M_k$. Let $\pi : \otimes_N^k M \to \otimes_N^k M$ be the map $\pi(x_1 \otimes x_2 \cdots \otimes x_k) = px_1p \otimes px_2p \otimes \cdots \otimes px_kp$, which is well defined since p commutes with N. Then a diagram shows that, for $x \in N' \cap M_{k-1}$, $p_kxp_k = \theta(\pi(\theta(x)))$, and conversely, if $\theta^{-1}(x)$ is in the image of α, $\pi(\theta^{-1}(x)) = \theta^{-1}(x)$ so $p_kxp_k = x$. Hence α induces a $*$-algebra isomorphism between $P_k^{pN\subset pMp}$ and $p_k(N' \cap M_{k-1})p_k$. To check that this map induces the right planar algebra structure, we first observe that α commutes suitably with the maps η and κ of 4.1.19. For η we have $\alpha \circ \eta = \pi \circ \eta \circ \alpha$ by definition. For κ, note that if we perform the basic construction of [J1] on $L^2(M)$, $\frac{1}{\mathrm{tr(p)}}pE_Np$ is the basic construction projection for $pN \subset pMp$ on $L^2(pMp)$. Thus if $\{b\}$ is a basis for pMp over pN we have $\sum_b bE_Nb^* = \mathrm{tr}(p)p$. So if

$$x = x_1 \otimes \cdots \otimes x_k \in \bigotimes_{pN}^{k} pMp \ ,$$

$$\alpha(\kappa_j(x)) = \frac{1}{\mathrm{tr(p)}\delta} \sum_b x_1 \otimes \cdots \otimes x_j b \otimes b^* \otimes x_{j+1} \otimes \cdots \otimes x_k$$

and

$$\kappa_j(\alpha(x)) = \frac{1}{\delta} \sum_c x_1 \otimes \cdots \otimes x_j c \otimes c^* \otimes x_{j+1} \otimes \cdots \otimes x_k \ .$$

Applying θ we see that $\alpha\kappa_j = \kappa_j\alpha$.

With these two commutation results it is an easy matter to check that α defines planar algebra isomorphism between $P^{Np\subset pMp}$ and $p(P^{N\subset M})p$, using the uniqueness part of 4.2.1 or otherwise.

Corollary 4.2.17 *If $N_1 \subset M_1$ and $N_2 \subset M_2$ are extremal finite index subfactors, then $P^{N_1\otimes N_2 \subset M_1 \otimes M_2}$ is naturally isomorphic to the tensor product $P^{N_1\subset M_1} \otimes P^{N_2 \subset M_2}$ of §3.4.*

Proof. We leave the details to the reader. □

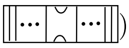
4.3 Planar algebras give subfactors

The following theorem relies heavily on a result of Popa [Po2].

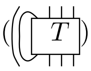
Theorem 4.3.1 *Let (P, Φ) be a spherical C^*-planar algebra with invariant Z and trace* tr. *Then there is a subfactor $N \subseteq M$ and isomorphisms $\Omega : N' \cap M_i \to P_i$ with* *(i)* Ω *is compatible with inclusions*

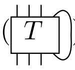
(ii) $\mathrm{tr}(\Omega(x)) = \mathrm{tr}(x)$

(iii) $\Omega(M' \cap M_i) = P_{1,i}$ *(linear span of tangles with vertical first string)*

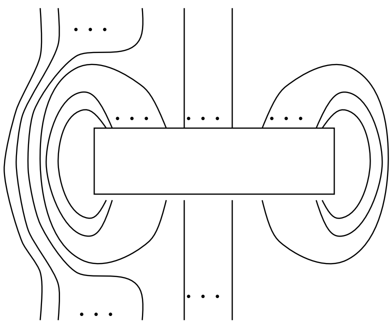
(iv) $\Omega(e_i) = \frac{1}{\delta}(\Phi($  $))$

(v) $[M : N] = \delta^2$. *If* $x \in N' \cap M_i$, *choose* $T \in \mathcal{P}$ *with* $\Phi(T) = \Omega(x)$.

(vi) $\Omega(E_{M'}(x)) = \frac{1}{\delta}\Phi($  $)$

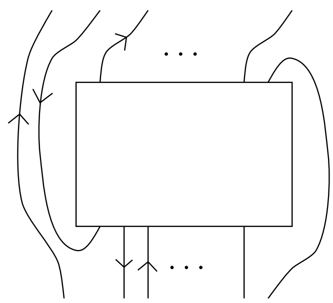
(vii) $\Omega(E_{M_{i-1}}(x)) = \frac{1}{\delta}\Phi($  $)$

Proof. By theorem 3.1 of [Po2], the pair $N \subset M$ exists given a system (A_{ij}), $0 \leq i < j < \infty$ of finite dimensional C^*-algebras with $A_{i,j} \subset A_{k,\ell}$ if $k \leq i$, $j \leq \ell$ and a faithful trace on $\bigcup_{n=0}^{\infty} A_{0n}$ satisfying 1.1.1, 1.1.2, 1.3.3$'$ and 2.1.1 of [Po2]. We set $A_{i,j} = P_{i,j}$ (Definition 1.20). Then $A_{ii} = \mathbb{C}$ since Z is multiplicative and non-degenerate. The conditions of [Po2] involve e_i's and conditional expectations $E_{A_{ij}}$. We define the e_i's in P to be what we have called $\Omega(e_i)$ (note our e_i is Popa's "e_{i+1}"). The map $E_{A_{ij}}$ is defined by the relation $\mathrm{tr}(xE_{A_{ij}}(y)) = \mathrm{tr}(xy)$ for x in A_{ij} and y arbitrary. Since Z is an S^2 invariant one easily checks that $E_{A_{ij}}$ is given by the element of $\mathcal{A}_{kj}(\emptyset)$ given in the figure below (for $x \in A_{0,k}$)

$E_{A_{ij}} : \frac{1}{\delta^{i+k-j}}$ 

$1 \; 2 \qquad i \;\; i+1 \;\; j$

Popa's (1.1.1) and (1.1.2) and b)$'$ of (1.3.3) follow immediately from pictures (note that the power of $\frac{1}{\delta}$ is checked by applying $E_{A_{ij}}$ to 1). Condition a)$'$ of 1.3.3$'$ is dim A_{ij} = dim $A_{i,j+1}e_j$ = dim $A_{i-1,j+1}$. But it is easy from pictures that $E_{A_{ij}}$ defines a linear map from $A_{i,j+1}e_j$ onto $A_{i,j}$, whose inverse is to embed in $A_{i,j+1}$ and multiply on the right by e_j. Moreover the element of $\mathcal{A}(\emptyset)$ (illustrated for $i=0$) in the figure defines a linear isomorphism from $A_{i,j}$ to $A_{i+1,j+1}$ — the inverse is a similar picture.



Finally the commutation relations 2.1.1, $[A_{ij}, A_{k\ell}] = 0$ for $i \leq j \leq k \leq \ell$ are trivial since they are true in $\mathcal{P}$, involving non-overlapping strings.

It is standard theory for external subfactors that $E_{M'}$ restricted to $N' \cap M_i$ is $E_{A_{1,i}}$ and $E_{M_{i-1}} = E_{A_{1,i-1}}$. So (vi) and (vii) are clear. $\square$

Corollary 4.3.2 *If* (P, Φ) *is as before, the Poincaré series* $\sum_{n=0}^{\infty} \dim(P_n) z^n$ *has radius of convergence* $\geq \frac{1}{\delta^2}$.

This result could be proved without the full strength of Theorem 4.3.1 (as pointed out by D.Bisch), using the principal graph and the trace. The Poincaré series of planar subfactors enjoy many special properties as we shall explore in future papers.

References

[A] V. Arnold, *Remarks on the enumeration of plane curves*, Amer. Math. Soc. Transl. **173**(1996), 17–32.

[Ban] T.Banica, *Hopf algebras and subfactors associated to vertex models*, math.QA/9804016

[BW] J.Barrett and B.Westbury, *Spherical Categories*, hep-th preprint #9310164.

[Ba] R. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, New York, 1982.

[BiW] J.Birman and H.Wenzl, *Braids, link polynomials and a new algebra*, Trans. Amer. Math. Soc. **163** (1989), 249–273.

[Bi] D.Bisch, *Bimodules, higher relative commutants and the fusion algebra associated to a subfactor*, The Fields Institute for Research in Mathematical Sciences Communications Series **13** (1997), 13–63.

[BH] D.Bisch and U.Haagerup, *Composition of subfactors: new examples of infinite depth subfactors*, Ann. scient. Éc. Norm. Sup. **29** (1996), 329–383.

[BJ1] D.Bisch and V.Jones, *Singly generated planar algebras of small dimension*, preprint 1998.

[BJ2] D.Bisch and V.Jones, *Algebras associated to intermediate subfactors*, Invent. Math. **128** (1997), 89–157.

[BJ3] D.Bisch and V.Jones, *A note on free composition of subfactors*, Geometry and physics (Aarhus, 1995), 339–361 in *Lecture Notes in Pure and Appl. Math.* **184**, Dekker, New York, 1997

[Co1] A.Connnes, *Noncommutative Geometry*, Academic Press (1994).

[dlH] P. de la Harpe, *Spin models for link polynomials, strongly regular graphs and Jaeger's Higman-Sims model*, Pacific J. Math **162** (1994), 57–96.

[Dr] V.Drinfeld, *Quantum groups*, Proceedings ICM 1986, vol. **1**, pp.798–820.

[EK] D.Evans and Y.Kawahigashi, *Quantum Symmetries on Operator Algebras*, Oxford University Press (1998).

[F+] P.Freyd, D.Yetter, J.Hoste, W.Lickorish, K.Millett, and A.Ocneanu, *A new polynomial invariant of knots and links*, Bull. AMS **12** (1985), 183–190.

[FY] P.Freyd and D.Yetter, *Braided compact closed categories with applications to low dimensional topology*, Adv. in Math. **77** (1989), 156–182.

[FRS] K. Fredenhagen, K-H. Rehren and B. Schroer, *Superselection sectors with braid group statistics and exchange algebras*, Comm. Math. Phys. **125** (1989), 201–226.

[Gn] S. Gnerre, *Free composition of paragroups*, Thesis, UC Berkeley, (1997).

[GHJ] F.Goodman, P. de la Harpe and V.F.R.Jones, *Coxeter Graphs and Towers of Algebras*, Springer Verlag, MSRI publications (1989).

[GL] J. Graham and G. Lehrer, *The representation theory of affine Temperley-Lieb algebras*, preprint, University of Sydney (1997).

[Ha] U.Haagerup, *Principal graphs of subfactors in the index range* $4 < [M : N] < 3 + \sqrt{2}$, *Subfactors*, World Scientific, Singapore-New Jersey-London-Hong Kong (1994), 1–39.

[Ja] F.Jaeger, *Strongly regular graphs and spin models for the Kauffman polynomial*, Geom. Dedicata **44** (1992), 23–52.

[JMN] F.Jaeger, M.Matsumoto and K.Nomura, *Bose-Mesner algebras related with type II matrices and spin models*, RIMS preprint (1995).

[J1] V.Jones, *Index for subfactors*, Invent. Math. **72** (1983), 1–25.

[J2] V.Jones, *A polynomial invariant for knots via von Neumann algebras,* Bulletin of the Amer. Math. Soc. **12** (1985), 103–112.

[J3] V.Jones, *Hecke algebra representations of braid groups and link polynomials*, Ann. Math. **126** (1987), 335–388.

[J4] V.Jones, *On a certain value of the Kauffman polynomial*, Comm. Math. Phys. **125** (1989), 459-467.

[J5] V.Jones, *Index for subrings of rings*, Contemporary Math. **43** (1985), 181–190.

[JS] V.F.R.Jones and V.Sunder, *Introduction to Subfactors*, LMS lecture note series no. **234** (1997) 162 pages.

[Ka1] L.Kauffman, *State models and the Jones polynomial*, Topology **26** (1987), 395–407.

[Ka2] L.Kauffman, *An Invariant of regular isotopy*, Trans. Amer. Math. Soc. **318** (1990), 417–471.

[KS] Krishnan amd V.Sunder, *On biunitary permutation matrices and some subfactors of index* 9, Trans. Amer. Math. Soc. **348** (1996), 4691–4736.

[Ku] G.Kuperberg, *Spiders for rank 2 Lie algebras*, Commun. Math. Phys. **180** (1996), 109–151, q-alg/9712003.

[La] Z. Landau, Thesis, UC Berkeley, 1998

[LM] W.Lickorish and K.Millett, *Some evaluations of link polynomials*, Comment. Math. Helv. **61** (1986), 349–359.

[Lo] R.Longo, *Index of subfactors and statistics of quantum fields, I, Comm. Math. Phys.* **126** (1989), 217–247.

[LR] R.Longo and J. Roberts, *A theory of dimension*, Journal of K-theory **11** (1997), 103–159, funct-an/9604008.

[Ma] J.May, *Definitions: operads, algebras and modules*, Contemporary Mathematics **202** (1997) 1–7.

[MvN] F.Murray and J.von Neumann, *On rings of operators, IV*, Annals Math. **44** (1943), 716–808.

[Mu] J.Murakami, *The Kauffman polynomial of links and representation theory*, Osaka Journal of Mathematics **24** (1987), 745–758.

[O1] A.Ocneanu, *Quantum symmetry, differential geometry of finite graphs and classification of subfactors*, University of Tokyo Seminary Notes (notes recorded by Y. Kawahigashi).

[Pe] R.Penrose, *Application of negative dimensional tensors*, in *Combinatorial Mathematics and Its Applications*, Academic Press (1971) 221–244.

[PP] M.Pimsner and S.Popa, *Entropy and index for subfactors*, Ann. Sci. Ecole Norm. Sup. **19** (1986), 57–106.

[Po1] S. Popa, *Classification of Subfactors and their Endomorphisms*, CBMS Lecture Notes, **86** (1995).

[Po2] S. Popa, *An axiomatization of the lattice of higher relative commutants of a subfactor*, Invent. Math. **120** (1995), 427–445.

[TL] H.Temperley and E.Lieb, *Relations between the "percolation" and "colouring" problem and other graph-theoretical problems associated with regular planar lattices: some exact results for the "percolation" problem*, Proc. Royal Soc. London **A322** (1971), 251–280.

[Tu] V.Turaev, *Quantum Invariants for Knots and 3-manifolds*, de Gruyter Berlin (1994).

[TV] V.Turaev and O.Viro, *State sum invariants of 3-manifolds and quantum 6-j symbols*, Topology **31** (1992) 86–902.

[Tut] W.Tutte, *On dichromatic polynomials*, J. Combinatorial Theory **2** (1967), 301–320.

[V] D.Voiculescu, *Multiplication of certain non-commuting random variables*, J. Operator Theory **18** (1987), 223–235.

[Wa] A.Wassermann, *Operator algebras and conformal field theory, III: Fusion of positive energy representations of LSU(N) using bounded operators*, Inventiones Math. **133** (1998) 467–538, math.OA/9806031.

[We1] H. Wenzl, *Hecke algebras of type A_n and subfactors*, Invent. Math. **92** (1988), 249–383.

[We2] H. Wenzl, *Quantum groups and subfactors of type B, C and D*, Comm. Math. Phys. **133** (1990), 383–432.

[We3] H. Wenzl, *C^* tensor categories from quantum groups*, J. Amer. Math. Soc. **11** (1998), 261–282.

[X] F.Xu, *Standard λ-lattices from quantum groups*, Invent. Math. **134** (1998).

Department of Mathematics
University of California
Berkeley, California 94720–3840