Advances in Mathematics **143**, 357–375 (1999)
Article ID aima.1998.1800, available online at http://www.idealibrary.com on IDEAL®

Spherical Categories

John W. Barrett

School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom
E-mail: jwb@maths.nott.ac.uk

and

Bruce W. Westbury

Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom
E-mail: bww@maths.warwick.ac.uk

This paper is a study of monoidal categories with duals where the tensor product need not be commutative. The motivating examples are categories of representations of Hopf algebras. We introduce the new notion of a spherical category. In the first section we prove a coherence theorem for a monoidal category with duals following S. MacLane (1963, *Rice Univ. Stud.* **49**, 28–46). In the second section we give the definition of a spherical category, and construct a natural quotient which is also spherical. In the third section we define spherical Hopf algebras so that the category of representations is spherical. Examples of spherical Hopf algebras are involutory Hopf algebras and ribbon Hopf algebras. Finally we study the natural quotient in these cases and show it is semisimple.

INTRODUCTION

In this paper we introduce the notion of a spherical category. The motivation for this and the main application is to generalise the Turaev–Viro state sum model invariant of a closed piecewise-linear 3-manifold. This application is discussed separately in [13].

A spherical category is defined to be a pivotal category which satisfies an extra condition, which states that two trace maps are equal. The universal strict pivotal category has an interpretation in terms of planar graphs. A number of authors have introduced categories whose morphisms are defined in terms of diagrams and which have applications to low dimensional topology (for example, [3, 4, 10]). The main difference between pivotal and spherical categories and these categories is that pivotal and spherical categories are not required to be braided. However, any braided pivotal category satisfies the extra spherical condition. The geometrical

0001-8708/99 $30.00

interpretation of the extra condition, and the reason for the terminology, is that in a pivotal category planar graphs are equivalent under isotopies of the plane whereas in a spherical category closed planar graphs are equivalent under isotopies of the 2-sphere.

The first part of the paper gives the definition of a monoidal category with duals and the proof of a coherence theorem. The proof follows the proof of the coherence theorem for monoidal categories in [7]. A monoidal category with duals can be thought of as an abstraction of the category of representations of a Hopf algebra, except that we require each object to be reflexive. A coherence theorem for categories with duals is given in [5]. The differences between this result and ours are that we do not assume that the monoidal category is symmetric but we do assume that all objects are reflexive. The difference between their proof and ours is that they emphasise the adjunctions involving Hom and $\otimes$ whereas we emphasise the contravariant duality functor (following [3]). The reason we need this result is that the category of representations of a Hopf algebra is not strict.

The main sources of spherical categories are categories of representations of Hopf algebras with a distinguished element satisfying conditions. These are called spherical Hopf algebras. The two main examples of such Hopf algebras are involutory Hopf algebras and ribbon Hopf algebras (such as quantised enveloping algebras). The reason we consider the more general notion of a spherical category instead of spherical Hopf algebras is that in the applications to 3-manifold invariants it is necessary to pass to a quotient. This quotient is not the category of representations of a Hopf algebra. The definition of this quotient applies to any additive spherical category.

In the last section we show that the quotient of the category of representations of a spherical Hopf algebra is semisimple. This is the main result in the paper. The proof uses properties of the matrix trace and does not apply to the quotient of an arbitrary additive spherical category. Although this quotient is semisimple the set of isomorphism classes of simple objects may be infinite even if the original Hopf algebra is finite dimensional. The construction of 3-manifold invariants depends on finding a spherical subcategory such that, in the semisimple quotient of this subcategory, the set of isomorphism classes of simple objects is finite. For each quantised enveloping algebra at a root of unity such a subcategory can be constructed using the results in [1, 2].

1. MONOIDAL CATEGORIES WITH DUALS

In this section we study monoidal categories with duals. Examples are finitely generated free modules over a commutative ring. In these examples the tensor product is commutative, but we do not require this. The motivating examples of monoidal categories with duals are the categories of representations of Hopf algebras. These categories have a contravariant functor given by the contragradient dual but are not categories with duals, in general, as we require that each object be reflexive.

The only result in this section is a coherence theorem. The proof follows the coherence theorem for monoidal categories in [7]. The definition of coherence involves an infinite set of diagrams. A coherence theorem gives a finite set of conditions for coherence to hold.

In the following definition it is convenient to regard the distinguished object e as a functor from the category with one morphism to $\mathscr{C}$. Then λ and ρ can be defined to be natural transformations.

DEFINITION 1.1. A monoidal category is a category, $\mathscr{C}$, together with a functor $\otimes : \mathscr{C} \times \mathscr{C} \to \mathscr{C}$, an object e, and natural transformations

$$\begin{aligned} \alpha &: (\otimes \times 1) \otimes \to (1 \times \otimes) \otimes \\ \lambda &: e \otimes 1 \to 1 \\ \rho &: 1 \otimes e \to 1 \end{aligned}$$

Each component of each of the natural transformations α, λ and ρ is required to be an isomorphism. These data are also required to be coherent which is equivalent to requiring that α satisfy the pentagon identity and that ρ and λ satisfy a triangular identity. Also, $\rho(e) = \lambda(e) : e \otimes e \to e$. These conditions are given in [8].

If each component of the natural transformations α, λ, and ρ is the identity map then the category is called strict monoidal.

If C is a category, then the opposite category C^- is the category obtained by reversing morphisms. If $F: C \to D$ is a functor, then $F^-: C^- \to D^-$ is the corresponding functor on the opposites. A functor $C^- \to D$ is the same as a contravariant functor $C \to D$.

DEFINITION 1.2. Let $(\mathscr{C}, \otimes, e, \alpha, \lambda, \rho)$ be a monoidal category. Then the data needed to make this a category with duals is a functor, $*: \mathscr{C}^- \to \mathscr{C}$, and natural transformations

(1) $\tau: 1 \to **$,

(2) $\gamma: (* \times *) \circ \otimes \to \otimes^{op} \circ *$

and an isomorphism $v: e \to e*$.

Each component of the natural transformations τ and γ is required to be an isomorphism. A category with data satisfying these conditions is called a category with dual data.

DEFINITION 1.3. A category with dual data is strict if each component of the natural transformations γ and τ is the identity map and the map v is the identity map. In this case the category is called a category with strict dual data.

In the following definition, if $\alpha = (\pm, \pm, ..., \pm)$ is an n-tuple of signs, then $\mathscr{C}^\alpha$ denotes the category $\mathscr{C}^\pm \times \mathscr{C}^\pm \times \cdots \times \mathscr{C}^\pm$, with $\mathscr{C}^+ \equiv \mathscr{C}$. Also, if 0 is the 0-tuple, then $\mathscr{C}^0$ is the category with one morphism, so that $e: \mathscr{C}^0 \to \mathscr{C}$.

It is convenient to regard v as a natural transformation from the functor e to the functor $e^- \circ *$.

DEFINITION 1.4. The set of iterates of the functors $\otimes$, $*$, and e is the smallest set of functors $F: \mathscr{C}^\alpha \to \mathscr{C}$ which satisfies the following conditions:

(1) The functor $e: \mathscr{C}^0 \to \mathscr{C}$ is an iterate.

(2) The identity functor $1: \mathscr{C} \to \mathscr{C}$ is an iterate.

(3) If F and G are iterates then the functor $F \otimes G = F \times G \circ \otimes$ is an iterate.

(4) If F is an iterate then $F^* = F^- \circ *$ is an iterate.

The length of α is called the multiplicity of the iterated functor. There is a category with objects the set of all iterates of $\otimes$, $*$, and e and morphisms all natural transformations.

DEFINITION 1.5. Let F, G, and H be any three iterated functors. Then we have natural transformations

$$\alpha: (F \otimes G) \otimes H \to F \otimes (G \otimes H)$$

$$\lambda: e \otimes F \to F$$

$$\rho: F \otimes e \to F$$

$$\gamma: (F^*) \otimes (G^*) \to (G \otimes F)^*$$

$$\tau: F \to F^{**}$$

$$v: e \to e^*$$

Each of these is an isomorphism. Each of these and their inverses are called instances. More particularly, the first one, for example, is called an instance of α, and its inverse an instance of the inverse of α.

DEFINITION 1.6. The set of expansions is the smallest subset of the set of natural transformations which satisfies the following:

(1) The identity functor, 1, is an expansion.

(2) Any instance is an expansion.

(3) If β is an expansion then $\beta \otimes 1$ and $1 \otimes \beta$ are expansions.

(4) If β is an expansion then β^* is an expansion.

The smallest set which satisfies all but the last condition is the set of reduced expansions.

DEFINITION 1.7. The smallest subcategory which contains instances and which is closed under $\otimes$ and $*$ is equal to the smallest subcategory which contains expansions. This is called the category of iterates and any morphism in this category is called an iterated natural transformation.

The smallest subcategory of the category of iterates which contains instances and which is closed under $\otimes$ is equal to the smallest subcategory of the category of iterates which contains reduced expansions. Any morphism in this category is called a reduced iterated natural transformation.

DEFINITION 1.8. Since every expansion is an isomorphism and the inverse of any expansion is an expansion it follows that every iterated natural transformation is an isomorphism and that every inverse of an iterated natural transformation is an iterated natural transformation. The category $\mathscr{C}$ with dual data is coherent if any two iterated natural transformations between the same two iterated functors are equal. Equivalently, the dual data is coherent if any diagram with vertices iterated functors and arrows expansions commutes.

A category with dual data which is coherent is called a category with duals.

The following is a coherence theorem, in the sense that this is a finite set of conditions on the data which are necessary and sufficient for the data to be coherent.

THEOREM 1.9. *A category $\mathscr{C}$ with dual data is coherent if, and only if, the following conditions are satisfied*:

(1) *$\tau(e)$ is the composite*

$$e \xrightarrow{v} e^* \xrightarrow{v^{*-1}} e^{**}.$$

(2) *For each object c, $\tau(c^*)$ and $\tau(c)^*$ are inverse isomorphisms.*

(3) *For any object a, the following diagram commutes*

$$\begin{CD} a^* \otimes e @>{1 \otimes v}>> a^* \otimes e^* \\ @V{\rho(a^*)}VV @VV{\gamma(a, e)}V \\ a^* @>>{\lambda^*(a)}> (e \otimes a)^* \end{CD}$$

(4) *For any object a, the following diagram commutes*

$$\begin{CD} e \otimes a^* @>{v \otimes 1}>> e^* \otimes a^* \\ @V{\lambda(a^*)}VV @VV{\gamma(e, a)}V \\ a^* @>>{\rho^*(a)}> (a \otimes e)^* \end{CD}$$

(5) *For all objects a, b,* and *c, the following diagram commutes*

$$\begin{CD} (c^* \otimes b^*) \otimes a^* @>{\gamma(c, b) \otimes 1}>> (b \otimes c)^* \otimes a^* @>{\gamma(b \otimes c, a)}>> (a \otimes (b \otimes c))^* \\ @V{\alpha(c^*, b^*, a^*)}VV @. @VV{\alpha^*(a, b, c)}V \\ c^* \otimes (b^* \otimes a^*) @>{1 \otimes \gamma(b, a)}>> c^* \otimes (a \otimes b)^* @>{\gamma(c, a \otimes b)}>> ((a \otimes b) \otimes c)^* \end{CD}$$

(6) *For all objects a and b the following diagram commutes*

$$\begin{CD} a \otimes b @>>{\tau(a \otimes b)}> (a \otimes b)^{**} \\ @V{\tau(a) \otimes \tau(b)}VV @VV{\gamma^*(b, a)}V \\ a^{**} \otimes b^{**} @>{\gamma(a^*, b^*)}>> (b^* \otimes a^*)^* \end{CD}$$

Note that the following diagram shows that $\tau(a)^{**} = \tau(a^{**})$ for any object a follows from the condition that τ is a natural transformation with each component an isomorphism.

$$\begin{CD} a @>{\tau(a)}>> a^{**} \\ @V{\tau(a)}VV @VV{\tau^{**}(a)}V \\ a^{**} @>>{\tau(a^{**})}> a^{****} \end{CD}$$

This is implied by condition (2) but not conversely.

It is trivial that if the natural transformations are coherent then each of these diagrams commutes. It remains to prove the converse.

LEMMA 1.10. *Any iterated natural transformation is equal to a reduced iterated natural transformation.*

Proof. There are six natural transformations α, λ, ρ, γ, τ and ν and six conditions above. Furthermore each condition has exactly one instance which is not reduced and for each natural transformation there is exactly one condition which contains an instance of that natural transformation which is not reduced. Hence, by applications of these diagrams and their images by $*$, any instance is equal to a reduced iterated natural transformation and it follows that any iterated natural transformation is equal to a reduced iterated natural transformation.

This shows that it is sufficient to prove the coherence theorem for reduced iterated natural transformations. Any reduced iterate of (α, λ, ρ) and their inverses is called a central iterate. Mac Lane proves the coherence theorem for central iterates in [7].

LEMMA 1.11. *Any expansion of an instance of γ, τ, or ν (but not their inverses) composed with a central iterate is equal to a central iterate composed with an expansion of an instance of γ, τ, or ν.*

Proof. This follows from the condition that the central iterate is a natural transformation and that each component of the expansion is a morphism. The commutative square which contains both of these provides the proof. An example is provided by the following commutative diagram:

$$\begin{array}{ccc} ((F^*\otimes G^*)\otimes H)\otimes J & \xrightarrow{\alpha(F^*\otimes G,\, H,\, J)} & (F^*\otimes G^*)\otimes(H\otimes J) \\ {\scriptstyle (\gamma(F,G)\otimes 1_H)\otimes 1_J}\big\downarrow & & \big\downarrow{\scriptstyle \gamma(F,G)\otimes 1_{H\otimes J}} \\ ((G\otimes F)^*\otimes H)\otimes J & \xrightarrow[\alpha((G\otimes F)^*,\, H,\, J)]{} & (G\otimes F)^*\otimes(H\otimes J) \end{array}$$

DEFINITION 1.12. The set of standard iterated functors is the smallest set of iterated functors which satisfies:

(1) The iterates e, 1, and $*$ are standard.

(2) If $F \neq e$ is standard then $F\otimes 1$ and $F\otimes *$ are standard.

DEFINITION 1.13. The set of reduced iterated functors is the smallest set of iterated functors which satisfies:

(1) The iterates e, 1, and $*$ are reduced.

(2) If F and G are reduced then $F\otimes G$ is reduced.

Define two iterated functors, F and G, to be equivalent if there is an iterated natural transformation from F to G. This is obviously an equivalence relation.

DEFINITION 1.14. The rank r and luminosity, l, of an iterated functor are defined recursively as follows:

(1) $r(e)=1$ and $r(1)=1$

(2) $r(F\otimes G)=r(F)+r(G)+1$

(3) $r(F^*)=r(F)$.

(1) $l(e)=0$ and $l(1)=0$

(2) $l(F\otimes G)=l(F)+l(G)$

(3) $l(F^*)=l(F)+r(F)$.

The rank is a modified multiplicity, and the luminosity is a measure of the number of stars. The luminosity is unaffected by central iterated natural transformations, but is increased by instances of γ, τ, and v.

LEMMA 1.15. *For any iterated functor* F, *there is a unique reduced iterated functor* R *and a unique natural transformation* $R\to F$ *which is an iterate of instances of* γ, τ, *and* v.

Proof. The proof is by induction on the sum of the luminosity and the rank of F. Any iterated functor is of the form

(1) F^{**},

(2) $(F\otimes G)^*$,

(3) e^*,

(4) $F\otimes G$,

(5) 1, $*$, or e;

and the only instances of γ, τ, and v which have these as range are

(1) $\tau\colon F\to F^{**}$

(2) $\gamma\colon (G^*)\otimes(F^*)\to(F\otimes G)^*$

(3) $v\colon e\to e^*$

(4) $A\otimes 1$ or $1\otimes B$

(5) none.

In the first three cases, the conclusion follows from the induction hypothesis applied to the domain of the natural transformation, which has the same rank and lower luminosity. In (4), the induction hypothesis can be applied to F and G separately, which have lower rank and luminosity which is no greater. The conclusion follows as the natural transformations $A \otimes 1$ and $1 \otimes B$ commute. The functors in (5) are reduced iterated functors, and $F = R$.

The proof of the coherence theorem [1.9] now follows. It is only necessary to consider reduced iterated natural transformations. A reduced iterated natural transformation is said to be directed if it is the product of instances which do not include the inverses of (γ, τ, ν). By the preceding Lemma [1.15], each iterated functor is equivalent to a reduced iterated functor, and hence to a standard iterated functor, by a directed natural transformation.

By a standard argument (see [7]), it is sufficient to prove the coherence result for directed natural transformations between an iterated functor F and a standard iterated functor S. According to Lemma 1.11, any directed natural transformation $S \to F$ can be written as a central iterate $S \to R$ followed by an iterate of instances of (γ, τ, ν): $R \to F$. By Lemma 1.15, R and the second natural transformation are unique, while the first natural transformation is unique by the coherence of central iterates.

Definition 1.16. If $\mathscr{C}$ is a category with duals then define the category $P\mathscr{C}$ to have objects of the form $(F, \{a_1, ..., a_n\})$ where F is a standard iterated functor of multiplicity n and each a_i is an object of $\mathscr{C}$. The set of morphisms from $(F, \{a_1, ..., a_n\})$ to $(G, \{a_1, ..., a_m\})$ is the set of morphisms in $\mathscr{C}$ from $F(a_1, ..., a_n)$ to $G(a_1, ..., a_m)$.

There is a functor $i: \mathscr{C} \to P\mathscr{C}$ defined on objects by $a \mapsto (1, a)$. There is also a functor $\pi: P\mathscr{C} \to \mathscr{C}$ defined on objects by $(F, \{a_1, ..., a_n\}) \mapsto F(a_1, ..., a_n)$. It is trivial that the composite $i \circ \pi$ is the identity functor of $\mathscr{C}$. It is also clear that there is a canonical natural transformation from $\pi \circ i$ to the identity functor of $P\mathscr{C}$. Hence the categories $P\mathscr{C}$ and $\mathscr{C}$ are canonically equivalent.

The coherence theorem implies that the category $P\mathscr{C}$ is a strict monoidal category with strict duals. The functor i has the universal property that if $\mathscr{D}$ is a strict monoidal category with strict duals and $F: \mathscr{C} \to \mathscr{D}$ is a functor of monoidal categories with duals then there is a unique functor $G: P\mathscr{C} \to \mathscr{D}$ of strict monoidal categories with strict duals such that $F = i \circ G$.

2. SPHERICAL CATEGORIES

In this section we introduce spherical categories. A pivotal category is a category with duals and with some additional structure and a spherical category is a pivotal category in which the additional structure satisfies an extra condition.

Strict pivotal categories are discussed in [3], where it is shown that the category of oriented planar graphs up to isotopy with labelled edges and with a distinguished edge at each vertex is a strict pivotal category.

If the sphere S^2 is regarded as the plane with the point at infinity attached, then a closed graph in the plane can be regarded as a closed graph on the sphere. There is an isotopy of the sphere which takes a closed graph of the form of Fig. 1 to the graph obtained by closing M in a loop to the left. This isotopy moves the loop in Fig. 1 past the point at infinity. Taken together with planar isotopies, such an operation on planar graphs generates all the isotopies on the sphere. It follows that the evaluation in a spherical category of a closed graph with labelled edges and with a distinguished edge at each vertex is invariant under isotopies of the sphere S^2.

DEFINITION 2.1. A pivotal category is a category with duals together with a morphism $\varepsilon(c)\colon e \to c \otimes c^*$ for each object $c \in \mathscr{C}$.

The conditions on the components of ε are the following:

(1) For all morphisms, $f\colon a \to b$, the following diagram commutes

$$\begin{array}{ccc} e & \xrightarrow{\varepsilon(a)} & a \otimes a^* \\ {\scriptstyle \varepsilon(b)}\downarrow & & \downarrow{\scriptstyle f \otimes 1} \\ b \otimes b^* & \xrightarrow[1 \otimes f^*]{} & b \otimes a^* \end{array}$$

(2) For all objects a, the following composite is the identity map of a^*:

$$a^* \xrightarrow{\lambda^{-1}(a^*)} e \otimes a^* \xrightarrow{\varepsilon(a^*) \otimes 1} (a^* \otimes a^{**}) \otimes a^* \xrightarrow{\alpha(a^*, a^{**}, a^*)} a^* \otimes (a^{**} \otimes a^*)$$

$$\xrightarrow{1 \otimes \gamma(a^*, a)} a^* \otimes (a \otimes a^*)^* \xrightarrow{1 \otimes \varepsilon(a)^*} a^* \otimes e^* \xrightarrow{1 \otimes \nu^{-1}} a^* \otimes e \xrightarrow{\rho(a^*)} a^*$$

(3) For all objects a and b the following composite is required to be $\varepsilon(a \otimes b)$:

$$e \xrightarrow{\varepsilon(a)} a \otimes a^* \xrightarrow{1 \otimes \lambda^{-1}(a^*)} a \otimes (e \otimes a^*)$$

$$\xrightarrow{1 \otimes (\varepsilon(b) \otimes 1)} a \otimes ((b \otimes b^*) \otimes a^*) \xrightarrow{1 \otimes \alpha(b, b^*, a^*)} a \otimes (b \otimes (b^* \otimes a^*))$$

$$\xrightarrow{\alpha^{-1}(a, b, b^* \otimes a^*)} (a \otimes b) \otimes (b^* \otimes a^*) \xrightarrow{1 \otimes \gamma(b, a)} (a \otimes b) \otimes (a \otimes b)^*$$

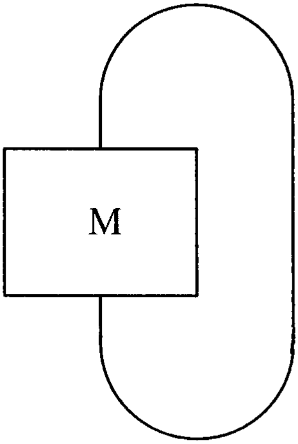

FIGURE 1

These conditions imply that the strict extension is a pivotal category in the sense of [3]. This is clear once it has been shown that ε is well- defined on the strict extension.

The functor $*$ and the maps ε are not independent. The maps ε determine $*$. This observation is in [5] where the maps ε are taken as more fundamental than $*$. In [12] it is noted that the maps ε also determine γ.

LEMMA 2.2. *In any pivotal category, for any morphism $f: a \to b$ the following composite is f^*:*

$$b^* \xrightarrow{\rho^{-1}(b^*)} b^* \otimes e \xrightarrow{1 \otimes \varepsilon(a)} b^* \otimes (a \otimes a^*) \xrightarrow{1 \otimes (f \otimes 1)} b^* \otimes (b \otimes a^*)$$

$$\xrightarrow{\alpha^{-1}(b^*, b, a^*)} (b^* \otimes b) \otimes a^* \xrightarrow{(\tau(b^*) \otimes \tau(b)) \otimes 1} (b^{***} \otimes b^{**}) \otimes a$$

$$\xrightarrow{\gamma(b^{**}, b^*) \otimes 1} (b^* \otimes b^{**})^* \otimes a^* \xrightarrow{\varepsilon^*(b^*) \otimes 1} e^* \otimes a^* \xrightarrow{\nu^{-1} \otimes 1} e \otimes a^* \xrightarrow{\lambda(a^*)} a^*$$

Proof. This follows directly from conditions (1) and (2) of the preceding definition.

Remark 2.3. Note also that a pivotal category is closed, in the sense that there are natural isomorphisms, for all a, b, and c;

$$\mathrm{Hom}(a \otimes b, c) \cong \mathrm{Hom}(b, a^* \otimes c), \qquad \mathrm{Hom}(a, b \otimes c) \cong \mathrm{Hom}(a \otimes c^*, b)$$

$$\mathrm{Hom}(a \otimes b, c) \cong \mathrm{Hom}(a, c \otimes b^*), \qquad \mathrm{Hom}(a, b \otimes c) \cong \mathrm{Hom}(b^* \otimes a, c).$$

DEFINITION 2.4. Let a be any object in a pivotal category. Then the monoid $\mathrm{End}(a)$ has two trace maps, tr_L, tr_R: $\mathrm{End}(a) \to \mathrm{End}(e)$. In a pivotal category $\mathrm{tr}_L(f)$ is defined to be the composite

$$e \xrightarrow{\varepsilon(a^*)} a^* \otimes a^{**} \xrightarrow{1 \otimes \tau^{-1}(a)} a^* \otimes a \xrightarrow{1 \otimes f} a^* \otimes a \xrightarrow{\tau(a^*) \otimes \tau(a)}$$

$$a^{***} \otimes a^{**} \xrightarrow{\gamma(a^{**}, a^*)} (a^* \otimes a^{**})^* \xrightarrow{\varepsilon^*(a^*)} e^* \xrightarrow{\nu} e$$

and $\mathrm{tr}_R(f)$ is defined to be the composite

$$e \xrightarrow{\varepsilon(a)} a \otimes a^* \xrightarrow{f \otimes 1} a \otimes a^* \xrightarrow{\tau(a) \otimes 1}$$

$$a^{**} \otimes a^* \xrightarrow{\gamma(a^*, a)} (a \otimes a^*)^* \xrightarrow{\varepsilon^*(a)} e^* \xrightarrow{\nu} e.$$

In a strict pivotal category these definitions simplify to

$$\mathrm{tr}_L(f) = \varepsilon(a^*)(1 \otimes f)\, \varepsilon(a^*)^*$$

$$\mathrm{tr}_R(f) = \varepsilon(a)(f \otimes 1)\, \varepsilon(a)^*$$

These are called trace maps because they satisfy $\mathrm{tr}_L(fg) = \mathrm{tr}_L(gf)$ and $\mathrm{tr}_R(fg) = \mathrm{tr}_R(gf)$.

DEFINITION 2.5. A pivotal category is spherical if, for all objects a and all morphisms $f: a \to a$,

$$\mathrm{tr}_L(f) = \mathrm{tr}_R(f).$$

An equivalent condition is that $\mathrm{tr}_L(f) = \mathrm{tr}_L(f^*)$, for all $f: a \to a$. Also, in a spherical category, $\mathrm{tr}_L(f \otimes g) = \mathrm{tr}_L(f) \cdot \mathrm{tr}_L(g)$ (where the product is in $\mathrm{End}(e)$) for all $f: a \to a$ and all $g: b \to b$.

In the rest of this paper we study additive spherical categories. This means that each Hom set is a finitely generated abelian group; and that the data defining the spherical structure is compatible with the additive structure. This means that $\otimes$ is $\mathbb{Z}$-bilinear, that $*$ is $\mathbb{Z}$-linear, and all components of all the natural transformations in the pivotal data are multilinear over $\mathbb{Z}$.

In any additive monoidal category $\mathrm{End}(e)$ is a commutative ring (see [5]) and we denote this ring by $\mathbb{F}$. In particular the trace map, tr_L, takes values in this ring. It follows that an additive monoidal category is $\mathbb{F}$-linear, by using the structure maps of λ and ρ, and that if the category is pivotal then the data defining the pivotal structure is $\mathbb{F}$-*linear*.

The main examples of additive spherical categories arise as the category of representations of a Hopf algebra with some additional structure. This is discussed in detail in the next section.

EXAMPLE 2.6. An example of an additive spherical category which cannot be regarded as a category whose objects are finitely generated free modules is given by taking the free $\mathbb{Z}[q, q^{-1}, z]$-linear category on the category of oriented tangles and then taking the quotient by the well-known skein relation for the HOMFLY polynomial. This is an additive spherical category and for each pair of objects a and b, $\mathrm{Hom}(a, b)$ is a finitely generated free $\mathbb{Z}[q, q^{-1}, z]$-module. However, the objects cannot be taken to be finitely generated modules unless z is a quantum integer.

DEFINITION 2.7. For any two objects a and b there is a bilinear pairing

$$\Theta\colon \mathrm{Hom}(a, b) \times \mathrm{Hom}(b, a) \to \mathbb{F}$$

defined by $\Theta(f, g) = \mathrm{tr}_L(fg) = \mathrm{tr}_L(gf)$.

DEFINITION 2.8. A spherical category is non-degenerate if, for all objects a and b, the pairing Θ is non-degenerate.

The next theorem shows that every additive spherical category has a natural quotient which is a non-degenerate spherical category. This construction is mentioned in [12].

THEOREM 2.9. *Let $\mathscr{C}$ be an additive spherical category. Define the subcategory $\mathscr{J}$ to have the same set of objects and morphisms defined by*

$$\mathrm{Hom}_{\mathscr{J}}(c_1, c_2) = \{ f \in \mathrm{Hom}_{\mathscr{C}}(c_1, c_2) : \mathrm{tr}_L(fg) = 0 \textit{ for all } g \in \mathrm{Hom}_{\mathscr{C}}(c_2, c_1)\}.$$

Then $\mathscr{C}/\mathscr{J}$ is a non-degenerate additive spherical category.

Proof. It is clear that $\mathscr{J}$ is closed under composition on either side by arbitrary morphisms in $\mathscr{C}$. Hence the quotient is an additive category. It is also clear that $f \in \mathscr{J}$ if and only if $f^* \in \mathscr{J}$ and so the functor $*$ is well defined on the quotient. The functor $\otimes$ is well-defined on the quotient since $f \in \mathscr{J}$ implies $f \otimes g_1 \in \mathscr{J}$ and $g_2 \otimes f \in \mathscr{J}$ for arbitrary morphisms in $\mathscr{C}$. This follows from the observation that $\mathrm{tr}_L((f \otimes g)h)$ can always be written as $\mathrm{tr}_L(fh')$ for a morphism h', which can be constructed from g and h.[1]

The structure maps of the natural transformations α, λ, ρ, γ, v, τ and the morphisms $\varepsilon(a)$ are taken to be the images in the quotient of the given morphisms in $\mathscr{C}$. The conditions on this structure which imply that this quotient is spherical follow from the same conditions in $\mathscr{C}$.

Each pairing Θ is non-degenerate by construction.

[1] We thank Marco Mackaay for pointing out that this argument was given incorrectly in [13].

3. SPHERICAL HOPF ALGEBRAS

In this section we introduce spherical Hopf algebras. These are defined to be Hopf algebras with some additional structure which implies that the category of finitely generated modules is a spherical category.

DEFINITION 3.1. A spherical Hopf algebra over $\mathbb{F}$ consists of $\mathbb{F}$-module A together with the following data

(1) a multiplication μ

(2) a unit η: $\mathbb{F} \to A$

(3) a comultiplication Δ: $A \to A \otimes A$

(4) a counit ε: $A \to \mathbb{F}$

(5) an antipodal map γ: $A \to A$

(6) an element $w \in A$.

The data $(A, \mu, \eta, \Delta, \varepsilon, \gamma)$ are required to define a Hopf algebra. The conditions on the element w are the following:

(1) $\gamma^2(a) = waw^{-1}$ for all $a \in A$.

(2) $\Delta(w) = w \otimes w$.

(3) $\gamma(w) = w^{-1}$.

(4) $\varepsilon(w) = 1$.

(5) $\mathrm{tr}(\theta w) = \mathrm{tr}(\theta w^{-1})$ for any left A-module, V, which is finitely generated and projective as an $\mathbb{F}$-module and for all $\theta \in \mathrm{End}_A(V)$.

It follows from the condition $\Delta(w) = w \otimes w$ that $\gamma(w) = w^{-1}$ and that $\varepsilon(w) = 1$. Such elements are called group-like.

EXAMPLE 3.2. Examples of Hopf algebras which are spherical are:

(1) Any involutory Hopf algebra is spherical. The element w can be taken to be 1.

(2) Any ribbon Hopf algebra, as defined in [10], is spherical. The element w can be taken to be uv^{-1} where the element u is determined by the quasi-triangular structure and the element v is the ribbon element.

Remark. 3.3. A Hopf algebra with an element w that satisfies the first two conditions of Definition 3.1 is spherical if either $w^2 = 1$, or all modules are self-dual.

Remark. 3.4. If A is a Hopf algebra there may exist more than one element w such that (A, w) is a spherical Hopf algebra. However, if w_1 is one such element then $w_2 = gw_1$ is another such element if and only if g satisfies the conditions:

(1) g is central
(2) g is group-like
(3) g is an involution.

In particular, if A is finite dimensional over a field, the set of g satisfying these conditions forms a group of the form $\mathbb{Z}_2^k$ for some integer k.

EXAMPLE 3.5. This is an example of a finite dimensional Hopf algebra over $\mathbb{C}$ which satisfies all the conditions for a spherical Hopf algebra except that the left and right traces are distinct. This example is the quantized enveloping algebra of the Borel subalgebra of $SL_2(\mathbb{C})$.

Let s be a primitive $2r$th root of unity with $r > 1$. Let B be the unital algebra generated by elements X and K subject to the defining relations

$$KX = sXK$$
$$K^{4r} = 1$$
$$X^r = 0.$$

Then B is a finite dimensional algebra and also has a Hopf algebra structure defined by:

(1) The coproduct, Δ, is defined by $\Delta(K) = K \otimes K$ and $\Delta(X) = X \otimes K + K^{-1} \otimes X$
(2) The augmentation, ε, is defined by $\varepsilon(K) = 1$ and $\varepsilon(X) = 0$
(3) The antipode, γ, is defined by $\gamma(K) = K^{-1}$ and $\gamma(X) = -sX$.

The element $w = K^2$ satisfies the conditions

$$\Delta(w) = w \otimes w$$
$$\varepsilon(w) = 1$$
$$\gamma(w) = w^{-1}$$
$$\gamma^2(b) = wbw^{-1} \quad \text{for all } b \in B.$$

The trace condition is not satisfied since B has $4r$ one-dimensional representations with $X = 0$ and K a $4r$th root of unity and it is clear that the trace condition is not satisfied in these representations.

In the rest of this paper a left A-module will mean a left A-module which is finitely generated and free as a $\mathbb{F}$-module.

PROPOSITION 3.6. *If A is a spherical Hopf algebra over $\mathbb{F}$ then the category of left A-modules is a spherical category.*

Proof. The category of modules of a bialgebra is monoidal. The category of modules of a Hopf algebra has a functor $*$ which sends a module to its contragradient dual and a morphism to its adjoint.

An element w in a Hopf algebra which satisfies conditions (1) and (2) determines a pivotal structure for the category of modules. For each module a, the map $\tau(a)$ given by the natural transformation $\tau: 1 \to **$ is defined to be the natural identification of $\mathbb{F}$-modules $a \to a^{**}$ followed by multiplication by w^{-1}. This is a morphism of modules. The morphism ε_a sends 1 to the canonical element of $a \otimes a^*$.

The two traces are given by $\mathrm{tr}_L(\theta) = \mathrm{tr}(\theta w)$ and $\mathrm{tr}_R(\theta) = \mathrm{tr}(\theta w^{-1})$. This shows that the category of modules of a spherical Hopf algebra is an additive spherical category.

The rest of the paper is a discussion of the properties of a spherical category which is constructed by taking the non-degenerate quotient of a spherical subcategory of the category of modules of a spherical Hopf algebra. In the rest of this paper we assume that $\mathbb{F}$ is an algebraically closed field. These results show that this quotient is semisimple. If the Hopf algebra is itself semisimple this is clear.

PROPOSITION 3.7. *Let A be a spherical Hopf algebra. Let x be a left A-module. Then* $\mathrm{End}(x)$, *the endomorphism algebra in the non-degenerate quotient, is semisimple.*

Proof. Let $\mathrm{End}_A(x)$ be the endomorphism algebra of x in the category of left A-modules. Then we show that the kernel of the map $\mathrm{End}_A(x) \to \mathrm{End}(x)$ contains the nilpotent radical of $\mathrm{End}_A(x)$. Since the nilpotent radical is a two-sided ideal it is sufficient to show that if $\theta \in \mathrm{End}_A(x)$ is nilpotent then $\mathrm{tr}_L(\theta) = 0$. However, θ and ω commute and so $\theta\omega$ is also nilpotent. This shows that $\mathrm{tr}_L(\theta)$ is the matrix trace of a nilpotent endomorphism of a finite dimensional vector space over $\mathbb{F}$ and so $\mathrm{tr}_L(\theta) = 0$.

PROPOSITION 3.8 *Let A be a spherical Hopf algebra. Take the non-degenerate quotient of the spherical category of left A-modules. Take J to be a set of objects such that if* $\mathrm{End}(x) \cong \mathbb{F}$ *then there is a unique $a \in J$ such that $x \cong a$. Then, for each object x in the non-degenerate quotient, there is a natural isomorphism*

$$x \cong \bigoplus_{a \in J} \mathrm{Hom}(x, a) \otimes a.$$

Proof. First we define, for each idempotent $\pi \in \mathrm{End}(x)$, an object πx in the non-degenerate quotient. Let $\pi' \in \mathrm{End}_A(x)$ be a lifting of π. Then $\pi' x$ is a left A-module and so also gives an object, πx, in the non-degenerate quotient. This object is independent of the choice of lifting since πx is characterised by the property that, for any w and y,

$$\mathrm{Hom}(w, \pi x) = \mathrm{Hom}(w, x) \cdot \pi \qquad \text{and} \qquad \mathrm{Hom}(\pi x, y) = \pi \cdot \mathrm{Hom}(x, y).$$

Next we show that if π is primitive then $\mathrm{End}(\pi x) \cong \mathbb{F}$. More generally, if π_1 and π_2 are primitive idempotents then

$$\mathrm{Hom}\ (\pi_1 x, \pi_2 x) \cong \begin{cases} \mathbb{F} & \text{if } \pi_1 \text{ and } \pi_2 \text{ are isomorphic} \\ 0 & \text{otherwise.} \end{cases}$$

If π is primitive then so is π' and so $\mathrm{End}_A(\pi' x)$ is a local algebra. Hence $\mathrm{End}(\pi x)$ is a non-trivial semisimple quotient of a local algebra and so is $\mathbb{F}$. Alternatively, $\mathrm{End}(\pi x)$ is π. $\mathrm{End}(x) \cdot \pi$ which is $\mathbb{F}$ since π is primitive.

Now write $1 \in \mathrm{End}(x)$ as a sum of orthogonal primitive idempotents. This decomposition is not canonical. However, for each $a \in J$, define E_a to be the sum of the idempotents, $\{\pi\}$, in the decomposition that satisfy $\pi x \cong a$. Then each non-zero E_a is a central idempotent in $\mathrm{End}(x)$ and the decomposition

$$\mathrm{x} \cong \bigoplus_{a \in J} E_a x$$

is canonical. Furthermore it follows from this discussion that, for each $a \in J$, $E_a x$ can be naturally identified with $V(x, a) \otimes a$ for some $V(x, a)$. Hence $x \cong \bigoplus_{a \in J} V(x, a) \otimes a$.

Each $V(x, a)$ can be identified with $\mathrm{Hom}(x, a)$ since $\mathrm{Hom}(a, b) = 0$ if a and b are distinct elements of J. This follows from the non-degeneracy condition which implies that if $\mathrm{Hom}(a, b) = 0$ then $\mathrm{Hom}(b, a) = 0$.

This shows that the non-degenerate quotient is semisimple and that the set J is a complete set of inequivalent simple objects. If the Hopf algebra is finite dimensional and semisimple then it is clear that this set is finite. If the Hopf algebra is finite dimensional but not semisimple then the set J is typically infinite.

There are two 3-manifold invariants constructed from the quantized enveloping algebra of $SL(2)$. One is the construction of an invariant of a framed 3-manifold given a surgery presentation of the 3-manifold given in [9]. The other is a state sum model invariant constructed in [11] using a triangulation. The formalism for extending the definitions from $SL(2)$ to other finite dimensional simple algebras is given in [12, 13]. Each of these approaches encounters the same difficulty which is that the set J is required to be finite. A quantized enveloping algebra at a root of unity is not semisimple and the set J for the category of all modules may be infinite. In order to define 3-manifold invariants it is sufficient to find a spherical subcategory with J finite. This problem was solved for $SL(2)$ in [9].

Let A be the quantized enveloping algebra of a finite dimensional semisimple Lie algebra at a root of unity. Assume that the order, k, of the quantum parameter q is at least the Coxeter number of the Lie algebra. Then it is shown in [2] that the category of tilting modules is the smallest category which contains the irreducibles with highest weight in the interior of the fundamental alcove and which is closed under taking tensor products and direct summands. Furthermore this category is spherical and the isomorphism classes of objects in the semisimple quotient are indexed by the dominant weights in the interior of the fundamental alcove, and in particular, is a finite set.

Remark 3.9. Let A be a finite dimensional spherical Hopf algebra. Then the category of projective left A-modules is a spherical category. This constructs a spherical subcategory with J finite but we know of no applications. If A is semisimple this result is superfluous and in the example of the quantised enveloping algebra of $SL(2)$ given in [9] every indecomposable projective has quantum dimension 0 so that J is empty.

The category of projective left A-modules is a subcategory of the spherical category of left A-modules. Hence to show it is spherical it is sufficient to show it is closed under the operations of tensor product and taking duals. The tensor product of a projective module with any module is projective and so the category of projective modules is closed under tensor products. It is shown in [6] that a finite dimensional Hopf algebra is Frobenius and therefore the dual of a projective module is projective.

REFERENCES

1. H. H. Andersen, Tensor products of quantized tilting modules, *Comm. Math. Phys.* **149** (1992), 149–159.
2. H. H. Andersen and J. Paradowski, Fusion categories arising from semisimple Lie algebras, *Comm. Math. Phys.* **169** (1995), 563–588.
3. P. J. Freyd and D. N. Yetter, Braided compact closed categories with applications to low dimensional topology, *Adv. Math.* **77** (1989), 156–182.
4. A. Joyal and R. Street, The geometry of tensor calculus, I, *Adv. Math.* **88** (1991), 55–112.
5. G. M. Kelly and M. I. Laplaza, Coherence for compact closed categories, *J. Pure Appl. Algebra* **19** (1980), 193–213.
6. R. G. Larson and M. E. Sweedler, An associative orthogonal bilinear form for Hopf algebras, *Amer. J. Math.* **91** (1969), 75–94.
7. S. Mac Lane, Natural associativity and commutativity, *Rice Univ. Stud.* **49** (1963), 28–46.
8. S. Mac Lane, "Categories for the Working Mathematician," Graduate Texts in Mathematics, Vol. 5, Springer-Verlag, New York/Heidelberg/Berlin, 1971.
9. N. Y. Reshetikhin and V. G. Turaev, Invariants of 3-manifolds via link polynomials and quantum groups, *Invent. Math.* **103** (1991), 547–597.

10. N. Y. Reshetikhin and V. G. Turaev, Ribbon graphs and their invariants derived from quantum groups, *Comm. Math. Phys.* **127** (1990), 1–26.
11. V. G. Turaev and O. Y. Viro, State sum invariants of 3-manifolds and quantum 6j-symbols, *Topology* **31** (1992), 865–902.
12. V. G. Turaev, Modular categories and 3-manifold invariants, *Internat. J. Modern Phys.* **6** (1992), 1807–1824.
13. J. W. Barrett and B. W. Westbury, Invariants of piecewise-linear 3-manifolds, *Trans. Amer. Math. Soc.* **348** (1996), 3997–4022.