7.4 Inclusions Inducing the Unit and Counit for Right-Dual Objects 92
7.5 Inclusion Inducing Naturality Isotopies for Braiding . . 94
7.6 Schematic of Balancing for Categories of Tangles 95

9.1 The Functor from RFC to $F(C)$ Induced by Inclusion and Freeness Factors through **FTC** 129

12.1 The Braiding Associated to a Multiplication 147

16.1 The Associahedron 164
16.2 The "Chinese Lantern" 173

17.1 Diagram Relating Unit Conditions for Strong Monoidal Functors 179

19.1 Coörienting the Finite Codimensional Strata 187
19.2 Vassiliev-type Extension Formulae 187
19.3 Relations on Chord and Bead Diagrams 193
19.4 A Dual Basis for Degree 3 Framed Weight Systems for Knots 194
19.5 Relations on Bead-and-Chord Tangle Diagrams 196

Chapter 1

Introduction

One of the most remarkable developments in the recent history of mathematics has been the discovery of an intimate connection between the central objects of study in low-dimensional geometric topology—classical knots and links and low-dimensional manifolds themselves—and what had heretofore been somewhat exotic algebraic objects—Hopf algebras, monoidal categories, and even more abstract-seeming structures. A correspondence between geometric and algebraic structures has, of course, been central to the development of mathematics, at least since Descartes provided the world with the coördinate plane.

The connection disclosed by the rise of "quantum topology" is, however, of a different character from that classically known. The classical connection is mediated by algebras of (possibly quite generalized) functions, so that the correspondence between geometric and algebraic objects is contravariant, as, for example, the correspondence between manifolds and algebras of smooth functions, or between affine schemes and commutative rings. In the connections between topology and algebra which have come to light since the discovery of the Jones polynomial, the topological objects (usually parts or relative versions of the primary objects of interest) are themselves the elements of an algebraic object. Topological information is then wrung from this algebraic object by representing it in other algebraic objects of the same type.

Another feature of these recent developments is the difference between the role categories and functors have usually played since their discovery and the role they now play in quantum topology. Rather than serving a foundational role, as a clean way of encoding "natural" constructions of one kind of mathematical object from another, categories in quantum topology stand as algebraic objects in their own right. This difference has not always been generally understood, even by quite brilliant mathematicians working in related areas, as the following personal anecdote involving the late Moshé Flato illustrates. One evening at a Joint Summer Research Conference in the early 1990's Nicholai Reshetikhin and I button-holed Flato, and explained at length Shum's coherence theorem and the role of categories in "quantum knot invariants". Flato was persistently dismissive of categories as a "mere language". I retired for the evening, leaving Reshetikhin and Flato to the discussion. At the next morning's session, Flato tapped me on the shoulder, and, giving a thumbs-up sign, whispered, "Hey! Viva les catégories! These new ones, the braided monoidal ones."

It is the purpose of this book to lay out clearly and in one place much of the scattered lore concerning the categories most intimately related with classical knot theory, and to relate these categories both to knot polynomials, which were the original motivation for their study, and to the theory of Vassiliev invariants. No claim is made that this treatment is exhaustive of the current state of knowledge, but it is the author's hope that it will prove useful to students and established researchers alike. One area specifically not touched in this work (though some of the requisite definitions are mentioned as examples) is the connection between the theory of monoidal categories and the known algebraic constructions of topological quantum field theories. We have also steered clear of any areas in which the universal constructions charateristic of category theory in its foundational role are needed, as for example limits or colimits of diagrams. By doing this, we emphasize the algebraic nature of the subject at hand.

Part I lays out the fundamentals of "functorial knot theory", recalling the necessary facts and theorems from both category theory and knot theory, and even providing proofs of some "folk theorems" which are universally assumed. Part II shows that Vassiliev theory, at least in its combinatorial guise, falls within the scope of functorial knot theory, and thus understood can be viewed as a species of algebraic deformation theory. Part I is intended to be fairly self-contained, with only standard topics in first year graduate courses as prerequisites. Part II assumes some familiarity with algebraic deformation theory (in particular, Gerstenhaber [23, 24] and Gerstenhaber and Schack [25]) and homological algebra (see, for example, Weibel [57]).

Chapter 2

Basic Concepts

In this chapter we introduce basic concepts from low-dimensional topology and category theory which will be required in this study. We will begin with concepts from classical knot theory, and then turn to categorical structures. Whenever possible, we will illustrate categorical notions with both of classically known "categories-as-foundations" examples, and with more recent "categories-as-algebra" examples, these latter being chosen to emphasize the close connection between the categorical concept and low-dimensional topology.

Throughout this study, unless otherwise specified, terms like "manifold", "map", "embedding" and "homotopy" will refer to the piecewise linear (PL) version of the concept. Due to various classical smoothing and triangulation theorems, it would generally be a matter of indifference if the smooth versions were being used. Although there are some concepts, such as framed links, which are more natural in the smooth setting, we prefer the PL setting to avoid some niceties involving restrictions on germs near boundaries which are needed to develop the theory of smooth tangles. We will address these in Chapter 8. In the earlier chapters we will attempt to point out the adjustments which would be needed in the smooth setting, either in asides or in footnotes.

Throughout this work the unit interval $[0,1] \subset \mathbb{R}$ is denoted $\mathbb{I}$.

2.1 Knots, Links and Tangles

Knots and links, that is to say, compact 1-submanifolds of $\mathbb{R}^3$ or $\mathbf{S}^3$, play a remarkably important role in the theory of smooth or piecewise linear 3- and 4-manifolds, and in a variety of other parts of mathematics and the sciences.

When equipped with a framing (or in the presence of orientations, a smooth field of normal vectors), they provide the data for the attaching of 2-handles to $\mathbf{B}^4$. Theorems of Kirby [35] show that every compact oriented 3-manifold arises as the boundary of a 4-dimensional handle-body with only 0- and 2-handles, and provides a calculus of "moves" to relate any two presentations of the same (diffeomorphic) 3-manifold(s). Similarly, the 2-handle structure turns out to be central to the properties of smooth 4-manifolds.

Many properties of singularities of complex plane curves are intimately related to the "link" of the singularity, that is, the intersection of the curve with the bounding $\mathbf{S}^3$ of a sufficiently small ball about the singularity. Finite families of closed trajectories of 3-dimensional dynamical systems can form links of arbitrary complexity.

Bacterial DNA forms a closed loop, and is thus reasonably modeled by a knot. Certain enzyme actions lead to very complex knots. More remarkable still, knots and links arise naturally from considerations in the quantization of general relativity.

For all of these reasons, the study of knots and links is of great interest, and it behooves us to consider precise definitions:

Definition 2.1 *A* (classical) knot *is an embedding of* $\mathbf{S}^1$ *into* $\mathbf{S}^3$ *(or* $\mathbb{R}^3$*).*

A (classical) link *is an embedding of* $\coprod_{i=1}^{n} \mathbf{S}^1$ *into* $\mathbf{S}^3$ *(or* $\mathbb{R}^3$*), for some* $n \in \mathbb{N}$*. (Note: we include 0, so that there is an "empty link").*

In all of the applications noted above, and whenever knots and links are studied topologically, the important thing is not the embedding itself, but its class under a suitable notion of equivalence defined in terms of geometric deformations. The naive notions of geometric deformation, homotopy, or even isotopy (that is, homotopy through embeddings) turn out to be unsuitable. Therefore we make

Definition 2.2 *Two knots or links* K_1, K_2 *are* ambient isotopic *or simply* equivalent *if there is an isotopy* $H : \mathbf{S}^3 \times \mathbb{I} \to \mathbf{S}^3$ *(or similarly for* $\mathbb{R}^3$ *instead of* $\mathbf{S}^3$*) which carries one to the other.*

More precisely, H is a PL map, satisfying $H(-,0) = Id_{\mathbf{S}^3}$*;* $H(-,t)$ *is a PL-homeomorphism for each t; and*

$$H(K_1(x), 1) = K_2(x)$$

(using K_i *to denote the mapping, with implied domain.)*

In this study, it is important to consider also a "relative" or local version of knots and links confined to a rectangular solid:

Definition 2.3 *A* tangle *is an embedding* $T : X \to \mathbb{I}^3$ *of a 1-manifold with boundary into the rectangular solid* $\mathbb{I}^3$ *satisfying*

$$T(\partial X) = T(X) \cap \partial \mathbb{I}^3 = T(X) \cap (\mathbb{I}^2 \times \{0,1\}).$$

The relevant notion of equivalence for tangles is then given by

Definition 2.4 *Two tangles* $T_1 : X_1 \to \mathbb{I}^3$ *and* $T_2 : X_2 \to \mathbb{I}^3$ *are* equivalent *or* isotopic rel boundary *if there exist a PL homeomorphism* $\Phi : X_1 \to X_2$ *and a map* $H : \mathbb{I}^3 \times \mathbb{I} \to \mathbb{I}^3$ *satisfying*

1. $H|_{\partial \mathbb{I}^3 \times \mathbb{I}} = p_{\partial \mathbb{I}^3}$
2. $H(-,t)$ *is a PL homeomorphism for all t*
3. $H(-,0) = Id_{\mathbb{I}^3}$
4. $H(T_1, 1) = T_2(\Phi) : X_1 \to \mathbb{I}^3$

The following lemma about ambient isotopies in $\mathbb{I}^3$ will be useful in what follows:

Lemma 2.5 *Given an isotopy H of a closed set $F = [\epsilon, 1-\epsilon]^3 \subset \mathbb{I}^3$, there is an isotopy $\tilde{H}$ of $\mathbb{I}^3$ to itself whose restriction to F is H, and whose restriction to $\partial\mathbb{I}^3$ is the trivial isotopy $p_{\partial\mathbb{I}^3} : \partial\mathbb{I}^3 \times \mathbb{I} \to \partial\mathbb{I}^3$.*

proof: Consider triangulations of $F \times \mathbb{I}$ and F on which the H is given by linear maps of the simplexes. Now, choose triangulations of $\partial\mathbb{I}^3 \times \mathbb{I}$ and $\partial\mathbb{I}^3$ subordinate to which the projection is given by linear maps of simplexes. Subdivide these triangulations so that the triangulation of ∂F and the triangulation of $\partial\mathbb{I}^3$ are isomorphic by the map given by radial projection from the center of $\mathbb{I}^3$.

Now, $\overline{\mathbb{I}^3 \setminus F}$ is PL homeomorphic to $[\partial F] \times \mathbb{I}$. Choose a PL homeomorphism $\phi_1 \times \phi_2 = \phi : \overline{\mathbb{I}^3 \setminus F} \to [\partial F] \times \mathbb{I}$ with the property that $\phi_2(\partial F) = 1$ and $\phi_2(\partial\mathbb{I}^3) = 0$. Then there is a piecewise smooth isotopy $S : \overline{\mathbb{I}^3 \setminus F} \times \mathbb{I} \to \overline{\mathbb{I}^3 \setminus F}$ given by $S(x,t) = \phi^{-1}(H(\phi_1(x), \phi_2(x)\cdot t), \phi_2(x))$ whose restrictions to ∂F and $\partial\mathbb{I}^3$ are linear. Now, let Σ be a PL approximation to S agreeing with S on ∂F and $\partial\mathbb{I}^3$. The desired isotopy is then given by

$$\tilde{H}(x,t) = \begin{cases} H(x,t) & \text{if } x \in F \\ \Sigma(x,t) & \text{if } x \in \mathbb{I}^3 \setminus F \end{cases}.$$

□

There are two particularly important auxiliary structures with which knots, links and tangles may be equipped: orientations and framings. The first may be defined either homologically or combinatorially in the PL setting.[1] We prefer the combinatorial approach:

Definition 2.6 *A knot, link or tangle is* oriented *if every edge is equipped with a choice of one of its vertices as "first", in such a way that no vertex is chosen as "first" for both edges with which it is incident. We encode this choice diagrammatically by equipping each edge with an arrow pointing from the first vertex to the other (last) vertex.*

[1] Of course in the smooth setting, we could also define orientations in terms of orientation on the tangent bundle.

Observe that it suffices to equip one arrow in each connected component of a knot, link, or tangle with an arrow to specify completely an orientation on it.

The second notion, that of framing, exists most naturally in the smooth setting as a choice of a framing for the normal bundle of the (smooth) knot, link, or tangle. We may, however, easily translate it into the PL setting as follows: in the presence of the standard orientation on the ambient $\mathbb{R}^3$, $\mathbb{S}^3$ or $\mathbb{I}^3$, and an orientation on the knot, link or tangle, the specification of a framing on the normal bundle can be reduced to the specification of a field of normal vectors, since a second normal vector may be obtained as the cross-product of the unit tangent vector with the given normal vector. Using the exponential map of the standard metric, we can replace this normal vector field with a thin ribbon, one edge of which is the knot, link, or tangle. We can then take this "ribbon" version of framed links and translate them into the PL setting:

Definition 2.7 *A* framing *of a (PL) knot, link, or tangle is an extension of the embedding $T : M^1 \to X^3$ (for $X^3 = S^3$, $\mathbb{R}^3$ or $\mathbb{I}^3$) defining the knot, link, or tangle to an embedding $T_f : M^1 \times \mathbb{I} \to X^3$ such that $T_f(x,0) = T(x)$, and (in the case of tangles) if $x \in \partial X^1$, then $T_f(x,t) \in \mathbb{I}^2 \times \{0,1\}$ for all $t \in \mathbb{I}$.*

In Chapter 8 we will consider the smooth approach in more detail.

We can also encode a framing by attaching an integer to each component of the knot, link, or tangle. In the case of knots and links, this integer is simply the linking number of the two boundaries of the ribbon (with the orientation on the opposite boundary reversed).

In the case of tangles, an encoding of framings by integers can be given, but either it will be non-canonical and involve a choice of which framing is the 0-framing for each interval component, or it will involve further restrictions on the intersections of the tangle with $\partial\mathbb{I}^3$.

In cases where we consider the tangles to be oriented or framed, we require that the ambient isotopy in the definition of equivalence respect the orientation or framing in the obvious sense.

In all cases, of knots, links, or tangles, with or without orientations or framings, the abuse of language which ignores the distinction between a thing and its equivalence class is commonplace. For example, "the unknot" refers to the equivalence class of a planar circle.

Although the fundamental topological notion of equivalence is that of ambient isotopy, or ambient isotopy rel boundary, it is convenient in practice to replace this notion with a more combinatorial notion. The relevant notion was given in the classic treatise on knot theory, *Knottentheorie*, by K. Reidemeister [44]:

Definition 2.8 *Two PL knots, links, or tangles are* isotopic by moves *if they can be related by a sequence of moves of the following form:*

> *Let Δ be a closed triangle (in some triangulation in the PL structure on $\mathbb{R}^3$, $\mathbf{S}^3$, or $\mathbb{I}^3$ as relevant) such that the intersection of the knot, link or tangle, T, is exactly one or two of the closed edges of Δ. Replace $\Delta \cup T$ with the closure of the edges of Δ not contained in T.*

We then have

Proposition 2.9 *Two knots (resp. links, tangles) T_1 and T_2 are equivalent if and only if they are isotopic by moves.*

In the case of knots and links, the proof is given in Reidemeister [44]. For tangles, Reidemeister's proof together with Lemma 2.5 give the desired result.

One important fact about knots, links and tangles is that they can be completely characterized up to equivalence by certain planar drawings, called "diagrams". A sequence of propositions and definitions make this precise:

Proposition 2.10 *Almost every (orthogonal) projection of a knot or link K onto a plane is "at-most-two-to-one", in the sense that the inverse image of any point of the plane contains zero, one or two points of K, with only (isolated) transverse double points. We call such a projection a* regular projection.

proof: The PL proof may be found in detail in [44]. We sketch it here. Observe that the (orthogonal) projections in $\mathbb{R}^3$ are parameterized by $\mathbf{S}^2$. "Almost every" then indicates all except a set of measure zero in $\mathbf{S}^2$, in particular, all projections except a family parameterized by a curve (perhaps with isolated points) in $\mathbf{S}^2$.

One must avoid the directions of the edges (a finite set of points) so that many-to-one image points do not arise by the projection of an edge to a point. For each pair of edges, the directions of secant lines joining a point of one edge to a point of the other form two (topological) disks or arcs on $\mathbf{S}^2$. In the case where they form arcs, we must avoid these arcs to ensure transversality of double points, and we must likewise avoid directions of secant lines from any vertex to any point for the same reason (a finite set of arcs and points). For each pair of edges, one must avoid directions of secant lines from a point on one edge to a point on the other which also hit other points, to avoid image points with multiplicity greater than two. The secant lines themselves fill a closed region of $\mathbb{R}^3$ in such a way that every point of the region, except those on the two edges, lies on exactly one secant line. We must thus avoid a curve of directions described by the intersection of the other edges of K with the region. □

In the case of tangles, an analogous result holds, though here we wish to consider only projections onto the "back wall" of the cube $\mathbb{I}^3$. Therefore we consider non-orthogonal projections onto the plane of the "back wall" followed by linear scaling into a standard square.

Of course, information is lost in the process of projection: one no longer knows the height of the points above or below the plane of projection. Since we are concerned with knots, links and tangles only up to equivalence, most of the lost information is irrelevant: there are ambient isotopies (or isotopies by moves) which preserve the projection, but change the height of the points. What cannot be changed by an ambient isotopy that preserves the projection is which of the preimage points of a double point lies above the other.

In fact, it is the case that this information about the preimages of each double point is enough to recover the knot or link up to equivalence.

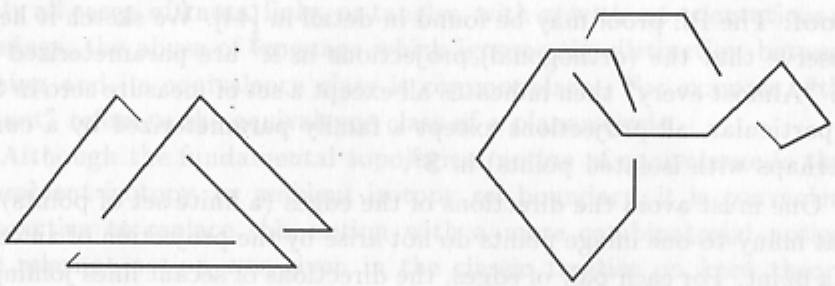

Figure 2.1: Examples of Knot Diagrams

By convention, the information is given by a *knot (or link) diagram*: a drawing of the projection in which the arc containing the lower of the two preimages is broken on either side of the double point, as, for example, in Figure 2.1. As is standard practice, we refer to these as knot diagrams, or simply diagrams, even in the case of links, and refer to the double points with the lower preimage indicated by the broken arc as *crossings*.

We then have

Theorem 2.11 *A knot or link is determined up to equivalence by any of its diagrams.*

The double points of a link diagram are called *crossings*. In the case where the link is oriented, we can distinguish two different types of crossings:

Definition 2.12 *Crossings in an oriented link diagram are* positive *or* negative *if the over- and under-crossing arcs are oriented as in Figure 2.2.*

Mnemonically, a crossing is positive if the right-hand rule curling from the out-bound over-crossing arc to the out-bound under-crossing arc gives a vector pointing up out of the plane of projection.

This then raises the question of when two diagrams determine the same equivalence class of knot of links. The answer is given by the classical theorem of Reidemeister [44]:

Figure 2.2: Crossing Signs

Theorem 2.13 *Two knot diagrams determine equivalent links if and only if they are related by a sequence of moves of the forms given in Figure 2.3.*

Before giving the proof of Theorem 2.13 we should comment on the fact that our set of moves is the original, larger set of combinatorial moves given in [44] rather than the smaller set, $\Omega.1$, $\Omega.2$ and $\Omega.3$, which is usually given under the name "Reidemeister moves" (cf. for example Burde and Zieschang [11]). The moves $\Delta.\pi.1$ and $\Delta.\pi.2$ are usually collected together in the phrase "isotopies of the plane of projection". Their inclusion, however, is both a convenience in the proof and, once the categorical structure of tangles is considered, a necessity for this study.

proof of Theorem 2.13:

The key to the proof is Reidemeister's other result: that ambient isotopy is equivalent to isotopy by moves. Consider a move across a triangle: if the projection of the triangle is an arc, the projection is unchanged by the move; otherwise, the projection of the triangle is itself a triangle.

To see that equivalence of diagrams under the diagrammatic moves implies isotopy by moves of the links is quite easy: each diagrammatic move becomes an isotopy-by-moves of the following form—use moves across triangles perpendicular to the plane of projection to adjust heights until the diagrammatic move can be realized as a single

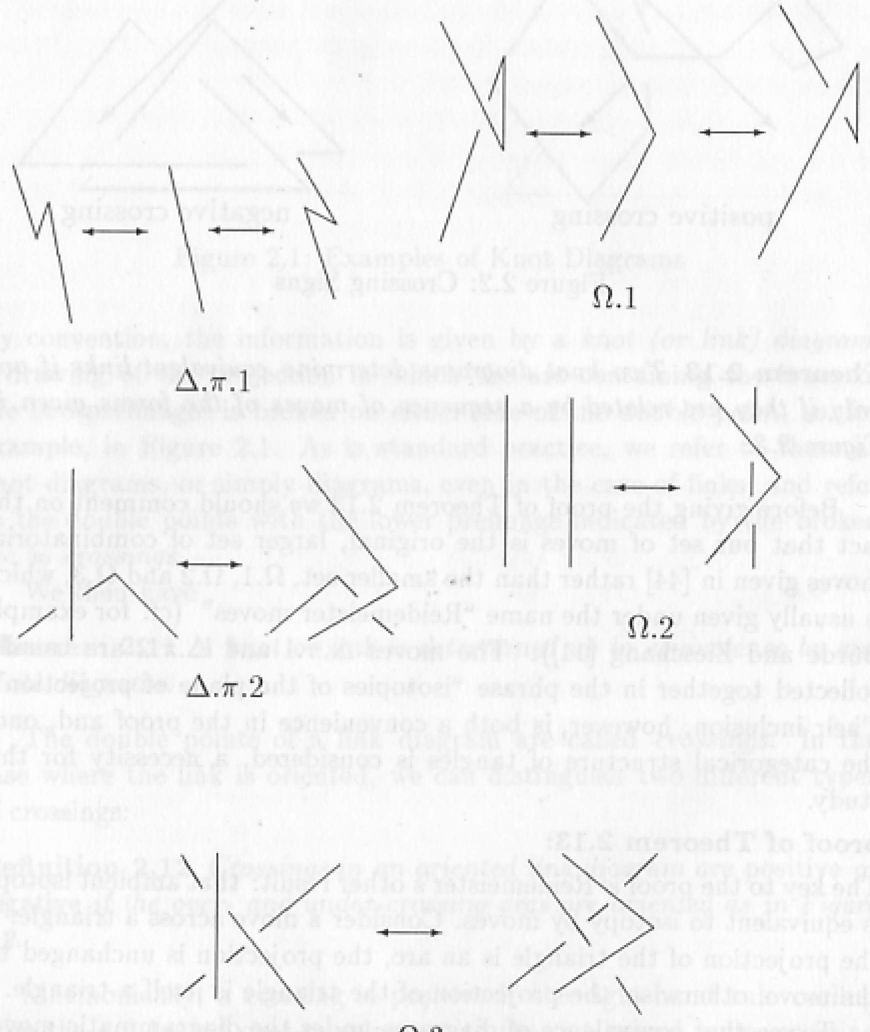

Figure 2.3: Reidemeister's Moves

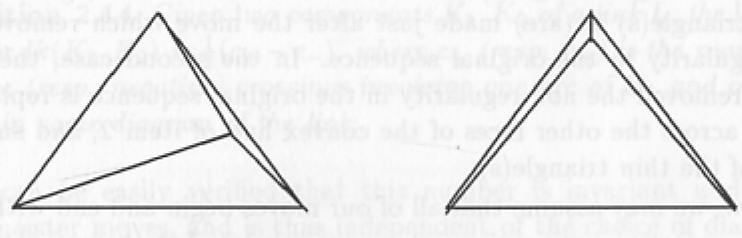

Figure 2.4: Subdivisions Useful to Avoid Non-regular Projections

move across a triangle parallel to the plane of projection.

For the converse, we would like to proceed by simply considering the effect of isotopy by a single move across a triangle on the projection. However, before doing so, we must show that we may assume, without loss of generality, that each move not only begins, but ends, with a regular projection.

Now, if we subdivide any triangle into smaller triangles, the move across the triangle can be realized instead as a sequence of moves across the smaller triangles. This observation is the key both to the remainder of the proof, and to solving the difficulty just mentioned.

If a move results in a non-regular projection, we can replace it with three moves across smaller triangles as in Figure 2.4. The subdivision point must be chosen so that the move across the large triangle(s) results in regular projections, and near enough to the new arc. Near enough, here, means

1. within a neighborhood bounding the new arc away from the triangles of later moves, if the non-regularity is removed by moves not involving the new arc, or

2. so that the convex hull of the triangle of the move removing the non-regularity and the image of the nearest-neighbor projection of its starting arc across the thin triangle(s) does not intersect the remainder of the link, if the non-regularity is removed by a move involving the new arc.

In either case, we replace the sequence of moves with a sequence in which the move introducing the non-regularity is replaced by the move(s) across the large triangle(s). In the first case, the move(s) across the small triangle(s) is (are) made just after the move which removed the non-regularity in the original sequence. In the second case, the move which removed the non-regularity in the original sequence is replace by moves across the other faces of the convex hull of item 2, and subdivisions of the thin triangle(s).

Now, we may assume that all of our moves begin and end with links whose projection onto a given plane are regular. Let the complexity of a move to be given by the number of edges, vertices and crossings of the link whose projection intersect the projection of the interior triangle of the move. If the move has a complexity greater than three, or if there are no vertices or crossings whose projection lies in the interior of the triangle and the move has a complexity greater than one, we can replace the move with a sequence of less complex moves across a subdivision of the triangle.

It therefore suffices to show that the result holds for moves of minimal complexity: those of complexity 0 and 1 with no vertices or crossings in the projection of the interior of the triangle, and those of complexity 3 involving a vertex or crossing.

Now, a move of complexity 0 is immediately seen to be one of type $\Delta.\pi.1$. A move of complexity 1 is of type $\Omega.1$ in the case where the edge whose projection is interior to the projected triangle is incident with the arc being moved, of type $\Delta.\pi.2$ in the case where it crosses the arc being moved on the boundary, and of type $\Omega.2$ otherwise.

A move of complexity 3 involving a vertex is of type $\Omega.2$ if the arc including the vertex does not cross the edge being moved, and of type $\Delta.\pi.2$ if it does.

Finally, a move of complexity 3 involving a crossing is plainly of type $\Omega.3$. □

It is easy to incorporate orientation data into a knot diagram: one need only equip the projection of each component of the link with an arrow on one of its arcs

Using crossing signs, itt is now possible to give a combinatorial definition of linking number:

Definition 2.14 *Given two components K_1, K_2 of a link L, the* linking number *$lk(K_1, K_2)$ is $\frac{1}{2}(c_+ - c_-)$, where c_+ (resp. c_-) is the number of positive (resp. negative) crossings involving one arc of K_1 and one arc of K_2 in some diagram of the link.*

It can be easily verified that this number is invariant under the Reidemeister moves, and is thus independent of the choice of diagram.

What is slightly less clear is that one can incorporate the framing information for an oriented framed link in the knot diagram as well: perform an ambient isotopy which is trivial outside of a tubular neighborhood of the link to make the ribbon parallel to the plane of projection, and pointing right with respect to the orientation vectors. In doing this, one may have to introduce kinks into the diagram (by moves of the form $\Omega.1$).

The ambient isotopy class of the oriented framed link can then be recovered from the resulting knot diagram by mapping the ribbon in such a way that it lies to the right of the curve when traversing it in the direction determined by the orientation. The framing determined in this way from a diagram is called the *blackboard framing* (cf. [36]). This process of introducing kinks to "flatten" the ribbon makes clear that the move $\Omega.1$ does not preserve the ambient isotopy type of the framed link which is recovered from the diagram.

All of the other Reidemeister moves may readily be seen to preserve the equivalence class of oriented links with the blackboard framing. Omitting $\Omega.1$ from the Reidemeister moves give a combinatorial notion of equivalence called *"regular isotopy"* which was used by Kauffman [32] in his formulation of the Jones polynomial, the so-called "Kauffman bracket" (cf. also [29]).

For our purposes, this combinatorial notion is less useful than a reduction to diagrams of ambient isotopy of framed oriented links. For this, we need to replace $\Omega.1$ with a substitute move which does respect the framing. To do this, we need to examine how the various cases of $\Omega.1$ change the blackboard framing. Observe that those which introduce

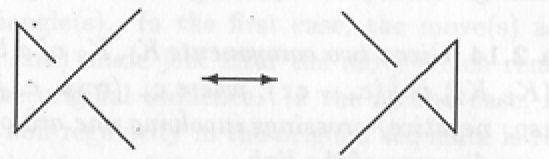

Figure 2.5: The Framed First Reidemeister Move

positive crossings change the framing (thought of as an integer) by +1, while those with negative crossings change it by −1. It therefore follows that any combinations of moves of type Ω.1 which change the framing by 0 must be admitted as moves.

Now any such sequence of Reidemeister moves which preserves the framing can be modified by moves of the types other than Ω.1 (by sliding curls along the component of the link) in such a way that moves of type Ω.1 which increase the framing are paired with moves of type Ω.1 which decrease the framing in small balls (or disks in the projection). By use of the simplest "Whitney trick"—the fact that moves of types Ω.2 and Ω.3 suffice to remove a pair of loops, provided they have opposite crossings, and lie on opposite sides of the arc in the projection, all of the various cases can be reduced to the single move in Figure 2.5.

2.2 Categories, Functors, Natural Transformations

We now turn to the basic notions from category theory needed for this study. The reader interested in a more thorough treatment is referred to Mac Lane [40], which contains most of the standard elementary definitions and theorems. We repeat those of particular importance for this study in this section and the next chapter.

Definition 2.15 (objects-and-arrows) *A category C consists of two collections $Ob(C)$ and $Arr(C)$, whose elements are called, respectively,* the objects *and* arrows *of C together with assignments of objects* target(f) *and* source(f) *to each arrow f; of an arrow Id_X to each object X; and*

of an arrow denoted fg or $g(f)$, called the composition *of f and g, to each pair of arrows f, g for which* target(f) = source(g), *and satisfying*

$$\begin{aligned} \mathrm{source}(Id_X) &= X \\ \mathrm{target}(Id_X) &= X \\ Id_{\mathrm{source}(f)} f &= f \\ f Id_{\mathrm{target}(f)} &= f \\ h(g(f)) &= h(g)(f). \end{aligned}$$

The arrows of a category are also, particularly in concrete settings, referred to at *morphisms* or *maps*. At first, we will adhere to calling them "arrows", but as we move to setting where the other names are common, we will begin to use them interchangeably.

The coyness of not describing source(−), target(−), Id_{-} and composition as functions is traditional (and to some minds necessary) because the collections involved are often proper classes.[2] The reader who dislikes bothering about the niceties of set theory may proceed safely: all of the categories which will occur in this book, outside of some illustrative examples in this section, are either *small* (that is, both $Ob(C)$ and $Arr(C)$ are sets) or *essentially small* (that is *equivalent*–as defined below–to a small category). One other notion connected with size in the set-theoretic sense should be mentioned: if for every pair of objects X, Y, the collection of arrows with source(f) = X and target(f) = Y is a set, we say the category is *locally small.* All categories considered herein are locally small.

[2]Many categorists object to the habit of mind which tries to place all of mathematics on a set-theoretic foundation. After all, when is the last time anyone ever actually cared about the ϵ-tree defining an element of a smooth manifold? For the insistently set-theory minded, we will dispense with the problems usually raised concerning sizes of categories by using a sufficiently strong large-cardinal axiom—Grothenieck's Axiom of Universe. Those size problems which do not collapse in the face of this axiom, and there are some, do not arise in this study.

Definition 2.16 *For a locally small category $\mathcal{C}$, the set of all arrows with* source$(f) = X$ *and* target$(f) = Y$ *is denoted $Hom_{\mathcal{C}}(X,Y)$ or simply $\mathcal{C}(X,Y)$, and called the* hom-set from X to Y.

If we do write the structure given in Definition 2.15 in terms of sets and functions we have

$$\text{source} : Arr(\mathcal{C}) \longrightarrow Ob(\mathcal{C})$$

$$\text{target} : Arr(\mathcal{C}) \longrightarrow Ob(\mathcal{C})$$

$$Id : Ob(\mathcal{C}) \longrightarrow Arr(\mathcal{C})$$

$$-(-) : Arr(\mathcal{C}) \times_{Ob(\mathcal{C})} Arr(\mathcal{C}) \longrightarrow Ob(\mathcal{C})$$

satisfying the functional equations given element-wise in the definition.

It will be observed that both source and target split Id, and thus Id is a bijection between $Ob(\mathcal{C})$ and its image. As we are concerned only with the structure of the category, not with the identity of its objects or arrow in some external ideal universe, this bijection allows us to forget the objects entirely: we can consider the identity maps themselves as the objects. Doing so gives an alternative definition of category which is sometimes more convenient:

Definition 2.17 (arrows-only) *A* category $\mathcal{C}$ *is a collection $\mathcal{C}$ whose elements are called* "arrows", *equipped with two unary operations* **source** *and* **target** *and a partially defined operation denoted by the null infix, with the property that fg is defined if and only if* **target**$(f) =$ **source**(g), *and satisfying*

$$\begin{aligned}
\mathbf{source}(\mathbf{source}(f)) &= \mathbf{source}(f)\\
\mathbf{target}(\mathbf{source}(f)) &= \mathbf{source}(f)\\
\mathbf{source}(\mathbf{target}(f)) &= \mathbf{target}(f)\\
\mathbf{target}(\mathbf{target}(f)) &= \mathbf{target}(f)\\
\mathbf{source}(f)f &= f
\end{aligned}$$

$$\begin{aligned}
f\mathbf{target}(f) &= f\\
[fg]h &= f[gh]
\end{aligned}$$

Example 2.18 Sets: *Objects are all sets in your favorite model of your favorite set-theory; arrows are all set-functions; source is domain; target is codomain; Id_X for any set X is the identity function on X; and composition is composition of set-functions.*

Example 2.19 Esp: *Objects are all topological spaces; arrows are all continuous maps; source is domain; target is codomain; Id_X for any space X is the identity function on X; and composition is composition of continuous maps.*

Example 2.20 K **− mod**: *Fix a ring K. Objects are all K modules; arrows are all K-linear maps; source is domain; target is codomain; Id_X is the identity map on X; and composition is composition of K-linear maps.*

Examples of this sort can be multiplied *ad infinitum*: take as objects all examples of some mathematical structure, and as arrows all maps preserving (some part of) the structure, ... In these cases it is most convenient to use the objects-and-arrows definition. This is not always the case. Consider

Example 2.21 G: *Fix a group (or monoid) G. Consider its elements as arrows with composition defined by the group law, and* **source** *and* **target** *given by the the constant map to e, the identity element.*

More important for this study are:

Example 2.22 Tang *(resp.* **Otang, Frtang***): Consider as arrows all equivalence classes of tangles (resp. oriented tangles, framed oriented* *tangles).* **source**(T) *(resp.* **target**(T)*) is the linear embedding of a disjoint union of copies of* $\mathbb{I}$ *which is constant in the first two coördinates and intersects* T *at each point of* $\mathbb{I}^2 \times \{0\}$ *(resp.* $\mathbb{I}^2 \times \{1\}$*) in the same set of points as* T *with (resp. the same set with the same orientation, the same set with the same orientation and framing). The composition of two tangles* T_1, T_2 *has as underlying 1-manifold the union of the underlying 1-manifolds of* T_1 *and* T_2 *with the points of the boundary lying in the face containing the common source/target identified. The composition* T_1T_2 *is then defined by the map on this underlying 1-manifold given as a composition of* $T_1 \coprod T_2$*, with the map* $\gamma_3 : \mathbb{I}^3 \coprod \mathbb{I}^3 \to \mathbb{I}^3$ *given by*

$$(x, y, z) \mapsto (x, y, \frac{z}{2}) \text{ for elements of the first summand}$$

$$(x, y, z) \mapsto (x, y, \frac{z+1}{2}) \text{ for elements of the second summand,}$$

with the connected components PL homeomorphic to $\mathbb{I}$ *reparameterized to preserve the condition at the boundary.*

It requires a little work to verify that this actually gives rise to a category. The conditions involving only **source** and **target**, but not composition, are immediate. To verify the other conditions, observe first that the two sides of the equations are certainly not equal by construction until we pass to equivalence classes. It is necessary to construct a PL (smooth) ambient isotopy rel boundary to verify the equations.

The required isotopies are constant in the first two coördinates of $\mathbb{I}^3$ and in all coördinates in a neighborhood of $\partial\mathbb{I}^3$. In the third coördinate they are given in a set F of the form $[\epsilon, 1-\epsilon]^3$ by (smoothings of) the PL maps shown schematically in Figure 2.6. The extension of this isotopy given by Lemma 2.5 then gives an isotopy which preserves the condition on the boundary.

Example 2.23 n**-Cobord***: As objects, take oriented smooth* $(n-1)$*-manifolds. As arrows, let* $Hom_{n\text{-}\mathbf{Cobord}}(M, N)$ *be the set of all equivalence classes of oriented* n*-manifolds with boundary* X *equipped with*

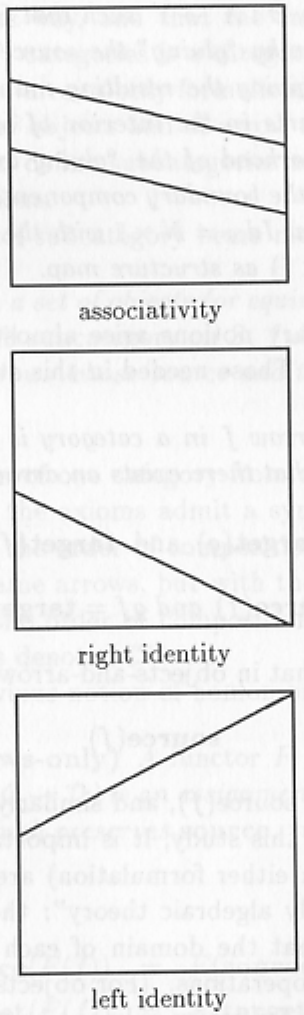

Figure 2.6: Isotopies Giving Identity and Associativity Conditions in Categories of Tangles