Commun. Math. Phys. 135, 615–640 (1991)

Communications in Mathematical Physics

2-*d* Physics and 3-*d* Topology

L. Crane

Department of Mathematics, Yale University, New Haven, CT 06520, USA

Received March 2, 1990

Abstract. Invariants of three dimensional manifolds and of framed oriented labeled links in them are rigorously defined using any solution to the Moore-Seiberg axioms for a Rational Conformal field theory. These invariants are generalizations of Witten's Chern-Simons path integrals. Connections are explored with supersymmetry, four dimensional manifolds, and quantum gravity.

I. Introduction

In [1] E. Witten studied a new object which he called a Chern-Simons path integral. By means of this formal expression, he was able to make a connection between a family of physical systems in two dimensions, the WZW models (which are the classical examples of rational conformal field theories), and a new set of invariants for three dimensional configurations, such as closed manifolds or closed manifolds containing framed links.

Unfortunately, the Chern-Simons path integral does not have a rigorous mathematical definition. Thus Witten's work does not constitute a proof that the expressions he produces are in fact topological invariants.

In [2], I outlined a proof of this in the simplest geometrical situation: a compact oriented 3-manifold with no boundary and no link. My proof avoided any consideration of three dimensional path integrals; and used the axiomatic description of a rational conformal field theory of Moore and Seiberg [3]. Thus, my result is mathematically rigorous, and more general than Witten's statement since it uses any RCFT, but deals with a less general topological configuration.

The key concept which connects rational conformal field theory to topology in three dimensions is duality. Duality is the physical principle which states that a physical process on a surface should be independent of the choice of decomposition for the surface. Changing decompositions of a surface is a process which can describe surface maps and braiding on surfaces; hence it can give information on links and three manifolds.

Since writing [2], it has become clear that the methods of [2] could be extended to cover more general topological situations. This is because the combinatorial topological description of closed 3-manifolds using Heegaard splittings, has close analogs for links and for manifolds with boundary. There is every reason to believe that a combinatorial reformulation can rigorously reproduce *everything* Witten deduces from his path integrals,. i.e., what is called a "topological quantum field theory," from any RCFT.

There is another advantage, beside generality, to proving as much as possible from the axioms for an RCFT. It is potentially very important, both for mathematical and physical purposes, to classify all RCFT's. The existence of three dimensional invariants, including invariants of ordinary knots and links, may prove to be a very powerful tool for attacking that problem. It has been conjectured ($3\frac{1}{2}$) that all RCFT's are related to WZW models, and hence to quantum groups. It may be possible to recover such a structure from a link invariant, via the connection with statistical mechanics ($3\frac{3}{4}$).

The combinatoric picture sheds light in another direction as well. There is a description of 3 manifolds, called Kirby calculus, which can also be used to construct 4-manifolds with boundary. Thus, we are able to reinterpret our invariants as invariants of 4-manifolds. This is potentially very significant, since the original inspiration for Witten's work was a talk by Atiyah [4] which suggested a connection between new developments in dimensions 2, 3 and 4. Reshetikhin and Turaev [8] discovered some of the same invariants by a very different method, involving quantum groups and Kirby Calculus.

Finally, there is a profoundly suggestive connection between this work and the reformulation of general relativity in the Ashtekar variables [5] as reworked by Smolin and Ravelli [6].

The outline of this paper is as follows: In Sect. 2, we outline the full results of Witten's paper, and define a topological quantum field theory. In Sect. 3, we review the Moore-Seiberg description of a rational conformal field theory. Section 4 reviews the classification of Heegaard splittings, and restates the proof of existence of invariants for closed oriented 3-folds. Section 5 deals with the problem of including links. Section 6 deals with manifolds with boundary. In Sect. 7, we discuss reinterpreting the invariant as four-dimensional; finally, in Sect. 8, we note the connections with the problem of quantizing general relativity.

II. Chern Simons Path Integrals and Topological Quantum Field Theory

The attitude of a pure mathematician to path integrals is reminiscent of the attitude of a religious Jew of the fifth century BC to his deity: it is imperative to try to approach it, but to gaze upon its face is death. Accordingly, I shall abstain from writing or formally manipulating any Chern Simons path integrals, and state only the results.

The construction of a CSPI involves the choice of a compact semisimple lie group G and a positive integer coupling constant k. If the topological configuration to be studied includes a link, each component of the link must be labeled with a representation of G. In CSPI it is *sufficient*, and in the combinatorial RCFT picture it is *necessary*, to choose the representations from a certain finite list $R_1 \dots R_n$. It is also necessary to frame each component of the link.

The CSPI can be performed on a closed oriented 3-manifold or a closed 3-manifold with a framed, labeled link in it. In that case, the result is a number. The number is slightly ambiguous, since it is only defined up to a power of the root of unity $e^{\frac{2\pi ic}{24}}$, where c is a rational number depending on G and k.

The path integral can also be performed on a manifold with boundary. If the manifold is to contain a labeled link, the link may contain open components which end on the boundary. In either case, we obtain a function of the boundary conditions. Constraints and symmetry conditions conspire to reduce the vector space of boundary conditions to a finite dimensional one. Thus, the information in the path integral can be summarized as producing what we call a *topological, quantum field theory*, in dimension 3.

Definition. A Topological Quantum Field Theory (in dim 3) is the following data:
[1] A finite dimensional space V_Σ for each oriented, closed surface or labeled punctured surface Σ (the pictures are labeled or "colored" from a finite index set which depends on the TQFT).
[2] A vector $v_m \in V_\Sigma$ for each M^3 an oriented manifold with boundary Σ. M^3 may also be a manifold containing a colored oriented, framed link. Some of the components of the link can be open and end on Σ, provided their ends coincide with punctures of Σ and the colors match. The initial and terminal pictures of an open colored link component may require different colors.

These data must satisfy the following conditions:

[A] If $\bar{\Sigma}$ denotes Σ with opposite orientation, then $V_{\bar{\Sigma}}$ is canonically identified with $\bar{V}_\Sigma$.

[B] If $\Sigma = \Sigma_1 \cup \Sigma_2$ (disjoint union) then $V_\Sigma = V_{\Sigma_1} \otimes V_{\Sigma_2}$.

[C] $V_\phi = \mathbf{C}$.

[D] V_Σ carries a representation of M_Σ, the mapping class group of Σ.

[E] (gluing) If two manifolds with boundary are joined along some components of their boundaries to form a new manifold with boundary, the vector associated to the new manifold can be calculated from the two old vectors by "tracing out" the components in the joined boundary components via the dual pairing from (A). (If the two manifolds contain links with components which end at the boundaries to be jointed, we join the open link components which meet at punctures on the joined boundary components).

We say a TQFT is "slightly ambiguous" if the vector in (2) is only defined up to an overall multiplication by some power of a root of unity $e^{\frac{2\pi ic}{24}}$, where c is a rational number associated to the TQFT. All TQFT's in this paper are "slightly ambiguous," and we will suppress mentioning this.

We note that (C) implies that a TQFT assigns a numerical invariant to a closed 3-manifold (possibly containing a link). Note also that if a manifold with boundary has its boundary written as $\Sigma_1 \cup \bar{\Sigma}_2$, i.e. is regarded as a cobordism, then the TQFT produces a morphism from V_{Σ_1} to $V_{\bar{\Sigma}_2}$, since $\mathrm{Hom}(A,B) \cong A \otimes B^*$. Hence, we can think of a TQFT as a representation of categories; (7) i.e. as a functor from a "cobordism category," to the category of vector spaces.

The main task of this paper is to reproduce the data of a TQFT from a RCFT. Let us list a sequence of tasks to be performed.

[1] Assign a vector space (of finite dimension) to a closed oriented surface Σ.

[$1\frac{1}{2}$] Assign a finite dimensional vector space to an oriented surface with "colored" punctures.

[2] Assign a numerical invariant to a closed oriented 3-fold.

[$2\frac{1}{2}$] Assign a numerical invariant to a closed 3-fold containing a framed colored link.

[3] Assign a vector $v_m \in v_\Sigma$ to any oriented compact 3-fold M with boundary Σ.

[$3\frac{1}{2}$] Assign a vector $v_m \in V_\Sigma$ to any oriented compact 3-fold M with boundary containing a framed colored link whose open components end precisely at the punctures of Σ (with matching colors). We may require that the open components of the links be oriented, with initial and terminal punctures colored differently.

[4] Check axioms A–E.

As we shall see, 1 and $1\frac{1}{2}$ follow from the MS axioms for a RCFT, as do axioms A–D. Problem (2) is the result of [1]. We shall recapitulate (2) in chapter 4, and treat $2\frac{1}{2}$ in Sect. 5, and 3 in Sect. 6. We do not consider $3\frac{1}{2}$ here, but believe it can be handled similarly. The proof of the gluing axiom should follow from finding a decomposition of the "big" manifold which extends the two "little" decompositions, but we do not complete it here.

III. Rational Conformal Field Theory. The Moore-Seiberg Formulation

The MS approach to RCFT is more explicit and computational than other approaches (8). This actually makes it much more easily applicable to topological questions.

The connection between 2-d physics and 3-d topology is via maps of surfaces to themselves. On the one hand, combinatorial topology tells us how to describe 3-folds and links in terms of surface maps; on the other, the MS axioms allow us to construct projective representations of the groups of isotopy classes of diffeomorphisms of surfaces. A peculiarly tidy coincidence between the details of combinatorial topology and those of conformal physics produces the invariants.

For the sake of completeness, we recapitulate the MS description. For motivation and details, see [3]. The definition which follows is not, strictly speaking, that of an RCFT, but of a slightly simpler object called a "modular tensor category." The fact that we can construct TQFT's from modular tensor categories may be very important since a class of Hopf algebras, the modular quasitriangular ones, also generate modular tensor categories. This may mean that many theorems also include quantum groups, if "modular quasitriangular Hopf algebra" is the right definition of quantum group. Frenkel is responsible for this observation.

The basic data for an RCFT include an algebra A and a finite set of representations $R_1 \dots R_n$. There is a distinguished member of the set $R_1 = 1$, and an involution $\hat{R}_i = R_j$ with $\hat{\hat{R}}_i = R_i$ and $\hat{1} = 1$. Associated to each triple of representations R_i, R_j, R_k, we have a finite dimensional vector space $V^{R_i}_{R_j,R_k}$ (which is a space of intertwining operators of a new sort) which we endow with a fixed choice of basis.

We assume that the space $V^1_{R_i R_j}$ has dimension 1 if $R_i = \hat{R}_j$ and 0 otherwise. In particular, $V^1_{1,1}$ is one dimensional with a fixed generator denoted $\begin{pmatrix} & 1 & \\ 1 & & 1 \end{pmatrix}$.

[The distinction between a RCFT and a MTC is that in a RCFT the algebra A, representations R_i and intertwining operators $V^{R_i}_{R_j R_k}$ must be explicitly given and satisfy some further properties, whereas in the MTC the R's and V's are just formal symbols which generate finite dimensional vector spaces on which the duality matrices (see below) act. Thus, we can easily extract an MTC from an RCFT.

The converse problem, finding RCFT's from MTC's, is much conjectured about, but not well understood. We should note that since the equations for the duality matrices of an MTC include a generalization of the Yang-Baxter equation, which is overdetermined; it is surprising that they exist at all. Other than extracting an MTC from an RCFT or a quantum group, we have no idea how to obtain one.]

These data enable us to associate a vector space V_Σ either to a closed oriented surface or to a surface with "colored" punctures, i.e., punctures labeled by representations. The procedure for constructing the space V_Σ is as follows. Pick any decomposition of Σ into trinions. If Σ has punctures, each puncture must be an end of a trinion. The trinions should be thought of as having stripes to keep track of twists when they are put together; decompositions with a different number of twists are not equivalent. Also, it is necessary to label the three ends of the trinion 0, 1, ∞. Decompositions with the ends of trinions differently labeled are inequivalent. Isotopic decompositions *are* equivalent. We then pick any labeling of the internal circles of the decomposition from the set $R_1 \dots R_n$ (circles corresponding to punctures are labeled consistently with the punctures.) We then form the vector spaces $V^{R_i}_{R_j R_k}$ for each trinion, tensor them together, and take a direct sum over all possible labelings. Tensoring the bases of the V's produces a basis for V_Σ.

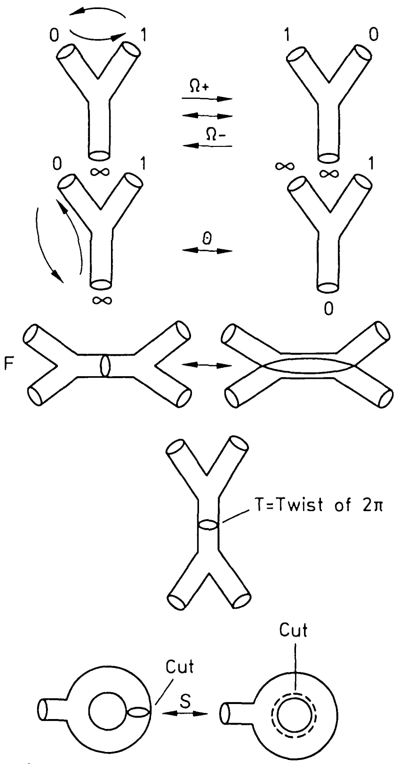
This construction does not yet appear to have any invariance with respect to changes in the trinion decomposition. In order to insure such invariance, we need to produce matrices which produce the change of basis resulting from a change in decomposition; and check that the matrices satisfy the same algebraic identities as the changes of decomposition themselves. Another way of saying this is that we have a representation of the "duality groupoid."

The duality groupoid is the concept which really forms the bridge between RCFT and topology. On the one hand, it is the most general mathematical expression of the "duality" of the antedeluvian dual model from which string theory arose; on the other, it contains both surface maps and braiding. Thus RCFT's, which possess "duality" produce projective representations of mapping class groups and braid groups.

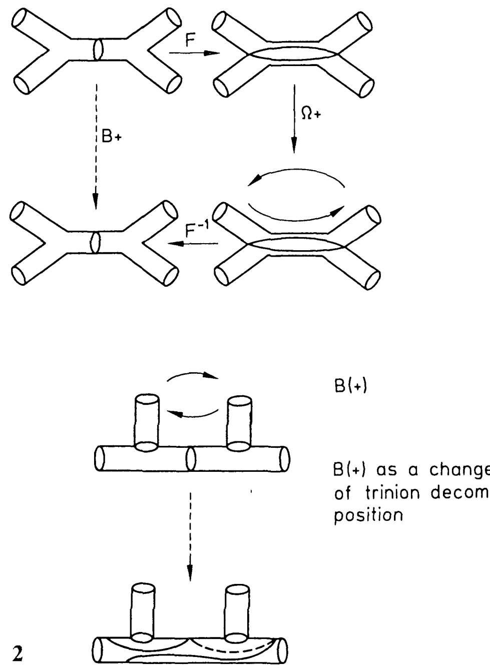
The duality groupoid of a surface Σ is the small category whose objects are the marked trinion decompositions of Σ described above. Its morphisms are words in a certain set of elementary changes in a decomposition or "moves." Moore and Seiberg describe a set of 5 elementary moves, labeled F, S, T, $\Omega^{\pm}$, Θ; which suffice to change any decomposition to any other (see Fig. 1). Thus the duality groupoid is connected. Next, it is necessary to find a list of identities which F, S, T, Ω, Θ must satisfy, in order to obey the same algebraic relations as the duality groupoid itself.

The list of identities which MS describe can be divided into 3 parts.

The first is the set of relations which come from the Yang-Baxter equation, which MS call "hexagons." The point here is that the duality groupoid contains



1



2

Fig. 1. The 5 duality moves

Fig. 2. Braiding $B+ = F + F^{-1}$

an operation called braiding $B\pm = F\Omega(\pm)F^{-1}$ which must obey the ordinary Reidemeister relations $B_{12}B_{23}B_{12} = B_{23}B_{12}B_{23}$ (Fig. 3).

The second relations, called the "pentagons" have to do with the relationship between braiding and trivalent vertices (Fig. 4). (The pentagon relations actually can be used to demonstrate that we have invariants of linked trivalent graphs, not only of ordinary links.)

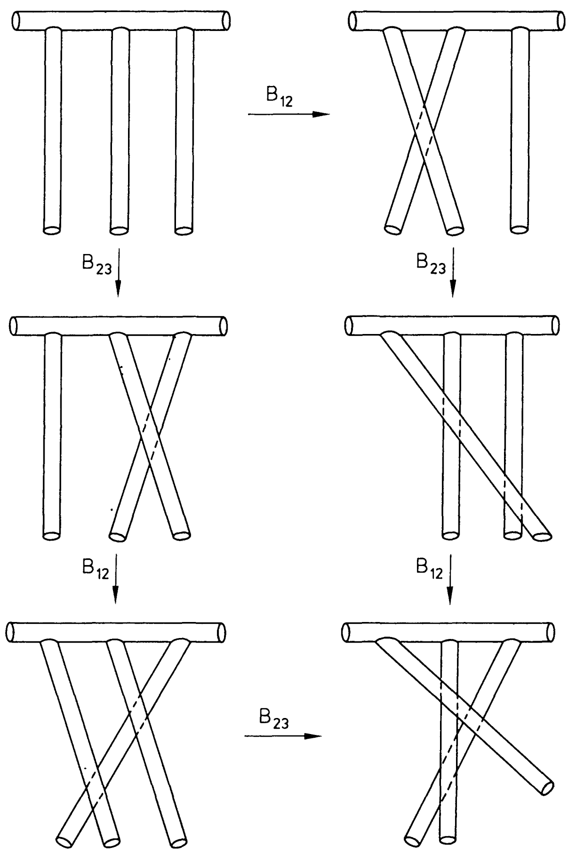
The third relations called the "genus 1 relations", insure that S and T generate a (slightly projective) representation of $PSL/2, Z)$.

In addition, MS assumes that Ω, Θ, T are diagonal matrices of phases in an appropriate basis. The number $e^{\frac{2\pi ic}{24}}$ appears in their formulae.

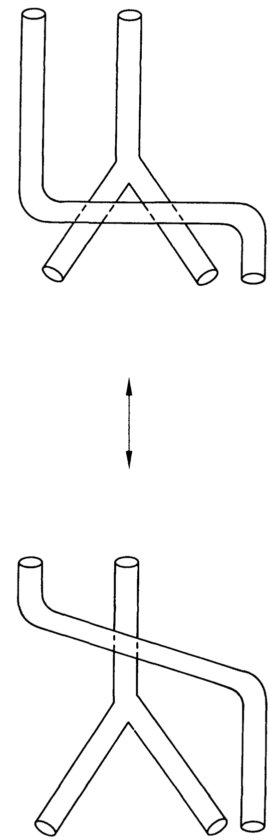
The main point of [3] is to prove that matrices which satisfy the identities listed above generate slightly projective representations of the duality groupoid. Thus, any set of solutions gives rise to a functor which assigns an invariant vector space to any orientable surface (possibly with colored punctures.)

Let us summarize by giving the definition of a Modular Tensor Category:

Definition. A modular tensor category consists of the following data:
[1] a set of labels $R_1 \dots R_n$,
[2] an involution on the set,
[3] a distinguished label in the set $R_1 = I$,



3



4

Fig. 3. The Yang Baxter equation. Note: The crossings in this diagram are shorthand for the braiding operation in Fig. 2

Fig. 4. The pentagon. Note: Crossing are shorthand for braiding as in Fig. 2 while the triple vertex denotes a fusion coefficient

[4] vector spaces $V^{R_i}_{R_j R_k}$ of finite dimension,
[5] matrices

$$
\begin{aligned}
T &: V^{R_i}_{R_j R_k} \to V^{R_i}_{R_j R_k}, \\
S &: V^{R_i}_{R_j \hat{R}_j} \to \oplus_t V^{R_i}_{R_t \hat{R}_t}, \\
\Omega^{\pm} &: V^{R_i}_{R_j R_k} \to V^{R_i}_{R_k R_j}, \\
\Theta &: V^{R_i}_{R_j R_k} \to V^{R_i}_{R_j R_k}, \\
F &: V^{R_i}_{R_j R_k} \otimes V^{R_k}_{R_l R_m} \to \oplus_t V^{R_j}_{R_l R_t} \otimes V^{R_t}_{R_j R_m}\,.
\end{aligned}
$$

Such that
(A) $\hat{1} = 1$.
(B) The MS identities are satisfied.
(C) $V^{\hat{R}_i}_{\hat{R}_j \hat{R}_k} = \hat{V}^{R_i}_{R_j R_k}$ (dual space).

(D) $V^{1}_{R_i R_j} = \mathbf{C}$ if $\hat{R}_i = R; = \phi$ otherwise.

(E) $V_{\hat{\Sigma}} = \hat{V}_{\Sigma}$ (reversing orientation dualizes).

We fix a choice of basis for each $V^{R_i}_{R_j R_k}$.

This is not a "minimal" definition, and assumes the theorem that V_Σ has an invariant definition, but it is good for our purposes.

IV. Topology of Oriented Three Manifolds and Invariants

In this section I will prove:

Theorem 1. *Any* MTC *gives rise to an invariant of compact oriented 3-manifolds.*

In order to prove this theorem, I recapitulate some well known facts about Heegaard splittings.

If H_1 and H_2 are solid handle bodies of the same genus, and ϕ is a diffeomorphism of $\partial H_1 = \Sigma_1$ onto $\partial H_2 = \Sigma_2$, then $H_1 \cap_\psi H_2 = M$ is a compact oriented 3-manifold. We call such a decomposition a Heegaard splitting of M. Without loss of generality, we can think of $\Sigma_1 = \Sigma_2 =$ some standard surface in R^3_1, $H_1 = \operatorname{int}\Sigma$, $H_2 = \operatorname{exterior}\Sigma \cup \{\infty\}$. Thus any Heegaard splitting can be described by a self map of a standard surface of genus g.

We recite the following result.

Fact 1. *Every compact oriented 3-fold admits a Heegaard splitting* [9].

There are three processes by which a Heegaard splitting can be changed into another of the same manifold. The first is isotopy of the surface map φ. This allows us to describe 3-folds via elements of the group of isotopy classes of diffeomorphisms of a compact oriented surface Σ to itself. This group is called the mapping class group M_g, where g is the genus of our standard surface Σ.

The second process is ambient diffeomorphism of the manifold M. This reduces [10] to action on the left and right of φ by the subgroups of M_g which extend to the solid handlebodies. We denote the subgroup of M_g which extends to a standard solid handlebody as N_g. Thus ambient diffeomorphism reduces oriented, compact 3-manifolds to double cosets in $N_g \backslash M_g / N_g$ for some g.

The third process is stabilization.

Stabilization relates Heegaard splittings along surfaces of different genus. To stabilize, we add a new handle to the surface, and extend φ by the identity on the new handle.

It is easy to see that these processes do not change the 3-fold. The converse is much harder. A classic result of Singer [11] implies:

Fact 2. *The equivalence relation between mappings $\varphi : \Sigma_g \to \Sigma_g$ (for all genuses) which create homeomorphic 3-folds is generated by isotopy, left and right action by copies of N_g, and stabilization.*

Thus, in order to create an invariant of 3-manifolds, it suffices to find an invariant of mapping class group elements which is preserved under stabilization and action of N_g.

Fortunately for us, the subgroup $N_g \subset M_g$ is well understood. Suzuki [12] has written a set of generators for N_g, which we shall describe below.

Our method for constructing an invariant is as follows. Given a compact oriented surface Σ which bounds a handlebody H, we construct, for any MTC, an element $v_H \in V_\Sigma$, which is canonical up to the usual phase ambiguity $\left(e^{\frac{2\pi i c}{24}}\right)^n$. We then show that this vector is invariant under the action of Suzuki's generators for N_g.

If we have a Heegaard splitting $M = H_1 \cup_\varphi H_2$, then H_1 and H_2 have as boundaries the same surface with opposite orientations. Thus v_{H_1} and v_{H_2} live in dual vector spaces V_Σ and $V_{\hat{\Sigma}} = \hat{V}_\Sigma$; and we can form their inner product $\langle v_{H_1} \mid v_{H_2}\rangle$. The invariances of the two vectors give our inner product the invariance under double cosets of Ng. The invariance under stabilization follows from suitable normalization.

In order to make this argument more precise and computable, we take the point of view that our two handlebodies are the interior and exterior of a standard surface in $S^3 = R^3 \cup \{\infty\}$. Thus our construction of V_H will give us two standard vectors $v_H \in V_\Sigma$ and $v_{\hat{\sigma}} \in V_{\hat{\Sigma}}$, and our invariant must involve v_h, V_H, and the isotopy class of our surface map φ in M_g.

At this point it is crucial to understand the relationship between the duality groupoid and the mapping class group. If we take a marked trinion decomposition and push it forward by a diffeomorphism of the surface to itself, we obtain another marked trinion decomposition. Conversely, if, two marked trinion decompositions have the same combinatoric structure as marked trivalent graphs (including twists), we can piece together maps of corresponding trinions, and obtain a surface map. Thus the duality groupoid contains copies of the mapping class group, all conjugate to one another. This implies that the vector space V_Σ admits a slightly projective representation of M_g. (The projective representation of M_g is historically very important in Conformal field theory) (see [13]). If we denote this representation by ϱ we are now able to write our prospective invariant as $I(M) = \langle v_{H_1} | \varrho\{\varphi\} | v_{H_2}\rangle$ for $M = H_1 \cup_\varphi H_2$, H_i standard.

It will be useful to normalize

$$\tilde{I}(M) = \frac{\langle v_{H_1} | \varrho\{\varphi\} | v_{H_2}\rangle}{\langle v_{H_1} \mid v_{H_2}\rangle}\,.$$

As we shall see, this normalization assures invariance under stabilization.

Now we must define v_H.

Definition. If $\partial H = \Sigma$, a solid trinion decomposition of H is a trinion decomposition of Σ all of whose cut circles bound disks in H.

Definition. If τ is a solid trinion decomposition of (H, Σ), the local vacuum $v_{H,\tau} \in V_\Sigma$, is the vector in V_Σ produced by labeling all the cut circles of τ with 1, choosing the basis element $\begin{pmatrix} & 1 & \\ 1 & & 1\end{pmatrix}$ of $V_{11}^1 = \mathbf{C}$ for each trinion in τ and tensoring them together.

We now need to prove:

Lemma 1. *$v_{H,\tau}$ is independent of τ up to a phase factor* $\left(e^{\frac{2\pi i c}{24}}\right)^n$.

Corollary. *v_H is invariant under the copy of N_g fixing H.*

Proof of Corollary. N_g takes solid trinion decompositions to solid trinion decompositions.

The proof of the lemma is the result of the following series of lemmas.

Lemma 2. *Of the Moore-Seiberg moves F, S, T, Ω, Θ; F, T, Ω, Θ preserve $v_{H,\tau}$ up to $\left(e^{\frac{2\pi i c}{24}}\right)^n$.*

Proof. In the MS axioms, T, Ω, Θ are diagonal matrices of phases, and the phase for 1 is $e^{\frac{2\pi i c}{24}}$. The matrix F can only change the labeling of one cut, but the uniqueness of the couplings of 1 precludes that.

Corollary. *Braiding preserves $v_{H,\tau}$.*

Proof. $B = F\Omega F^{-1}$.

Lemma 3. *The Suzuki generators preserve $v_{H,\tau}$.*

To prove this, we need to describe the Suzuki generators. In the following, it is helpful to think of a standard solid handlebody as a ball in R^3 with some number of small solid tubes attached to it at both ends to form handles. We can make a slight reformulation of Suzuki's work as follows:

Fact. *Ng is generated by the following five moves:*
[1] *Knob twists – twist a region containing both ends of a solid handle so that the handle maps into itself in reverse direction.*
[2] *Handle twists – cut thru a handle, twist by 2π, and rejoin.*
[3] *Handle exchanges – move 2 solid handles into each other.*
[4] *Handle braidings – twist a region in the ball so that one end of one handle is dragged around one end of another.*
[5] *Handle slides – drag one end of one handle up along and down another handle.*

We are reminded that we are to think of these as maps of the surface to itself which extend to its interior.

The easiest way to prove Lemma 3 is to first prove the following, which we also need separately.

Lemma 4. *It is possible to move from any orbit of Mg in the duality groupoid to any other, while staying in the set of solid trinion decomposition of a particular handlebody, without changing $v_{H,\tau}$.*

Proof. The orbits of Mg are just the marked trivalent graphs of genus g. Any such graph can be thickened to give a solid trinion decomposition of H. Thus we are reduced to proving that we can get from one graph to another via repeated operations of F, T, Ω and Θ. The proof is a simple induction on the number of external lines on each ring. (An application of F can reduce it by 1.) Moore [14] informs me that he has written out the proof in another context.

Proof of Lemma 3. The first four moves on Suzuki's list can be written as products of Dehn twists [9] each of which is on a loop in Σ which bounds a disk in H. Using the technique in Lemma 4, we can rearrange the decomposition τ, so that the Dehn twists occur along cuts of the decomposition, without changing the state $v_{H,\tau}$. A Dehn twist on a cut only changes $v_{H,\tau}$ by a factor $\left(e^{\frac{2\pi i c}{24}}\right)^n$.

The fifth move requires a more subtle argument, since if written as a product of Dehn twists, one must use loops which do not bound in H. However, it is possible to accomplish a handle slide by successively braiding one end of one handle past the exterior links of another. Thus the corollary to Lemma 2 finishes the proof.

Proof of Lemma 1. If two solid trinion decompositions of H are in the same orbit of M_g, then piecing together maps of solid trinions would show that the element of M_g extends to H; hence it is in N_g. Suzuki's result, together with Lemma 3, shows that they determine the same $v_{H,\tau}$. Lemma 4 now shows that all orbits determine the same $v_{H,\tau}$.

Proof of Theorem 1. Let $M = H_1 \cup_\varphi H_2$, where H_1 and H_2 are the interior and exterior of a fixed $\Sigma \subset R^3 \subset S^3$. Form the expression $\tilde{I}(M) = \frac{\langle v_{H_1} | \varrho\{\varphi\} | v_{H_2} \rangle}{\langle v_{H_1} | v_{H_2} \rangle}$. This clearly depends only on the class of φ in M_g. If $\varphi' = n_1 \varphi n_2$, where n_i are in the appropriate copies of N_g, then the corollary to Lemma 1 shows that $\tilde{I}$ is unchanged. To show that $\tilde{I}$ is topological, we now need only show that it is invariant under stabilization. Stabilization implies tensoring V_Σ with some other vector space, on which φ acts as the identity. Thus, our normalization will cancel it.

It has become clear from calculations made by other authors that these new invariants are non-trivial. What they will tell us about 3-manifolds, I cannot say at this point.

V. Links, Plats, and Physical Invariants

We now address the problem of including framed links in the 3-manifold.

The method we shall adopt uses the plat representation of links due to Birman [15]. This representation bears a deep similarity to Heegaard splittings, of which Birman was completely aware. This similarity allows us to imitate the argument of Sect. 4. I regard this as an aesthetic feature of this treatment, in that it is tempting to think of one half of a Heegaard splitting as a "subject," and the other as an "observer." Later we shall discuss whether this picture has physical significance.

In order to construct an invariant for links in a general 3-manifold, it is necessary to generalize J. Birman's plat concept to links in a 3-manifold; hence I include here a theory of "generalized plats."

I begin by recapitulating the plat description of ordinary links in $S^3 = R^3 \cup \{\infty\}$. For details, see [15].

If we picture an ordinary link before us in space, grasp it between two hands with large numbers of (very flexible) fingers by looping one finger of the (right) hand thru each local left (right) extremum of each little loop of the link, then pull the two hands apart; we will have a plat representation of the link. More specifically, we will have a braid between our two hands, with each side of the braid closed by a set of arcs (one around each finger). This reduces links to braids. Since braid groups are closely related to mapping class groups for punctured surfaces, the apparatus of Sect. 4 applies.

To formalize:

Definition. A plat is a subset of R^3 formed by the union of two sets of n semicircles on the farther sides of two parallel planes (with endpoints on the planes) with a $2n$ stranded geometric braid in the region between the planes, meeting the $2n$ endpoints of the semicircles on each plane.

Note. A plat is a link. Without loss of generality it is possible to line up the endpoints of the semicircles in a straight line.

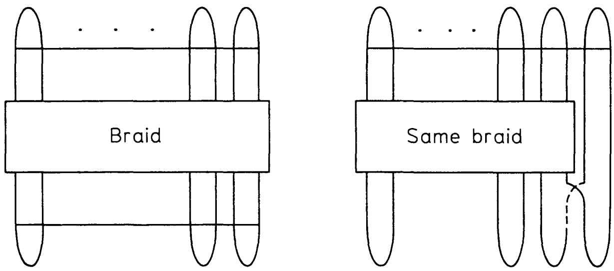


Fig. 5. Stabilizing a plat

We reproduce the following results from [15].

Fact 3. *Any link in S^3 has a plat representation.*

Fact 4. *The equivalence relation of plats which give rise to isotopic links in generated by two moves:*
1) *Stabilization,*
2) *Action on the left and right by a subgroup $K_{2n} \subset B_{2n}$ of the braid group.*

The analogy with Heegaard splittings is now visible. We shall see that it is deeper than it appears.

Let us describe the moves listed in Fact 4. Stabilization consists of making a break in a strand of the braid, extending the left end to the right plane and vice versa, attaching a new little semicircle on each side, and connecting the free ends of the new semicircles, so that the new loop formed does not encircle any part of the link (Fig. 5). It is sufficient to stabilize close to the left plane, i.e. past any crossings of the braid.

In order to describe the subgroup K of the braid group which appears in Fact 4, it is useful to remind ourselves of what the braid group B_{2n} looks like. It has generators $\sigma_{i,i+1}$ which cross the i^{th} strand over the $i+1^{\text{st}}$. There are two relations:

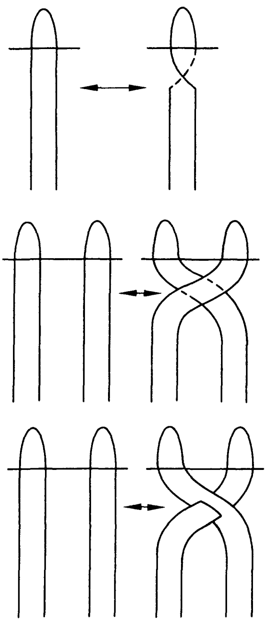
$$[\text{R1}] \quad \sigma_{i,i+1}\sigma_{j,j+1} = \sigma_{j,j+1}\sigma_{i,i+1} \ \text{ if } \ |i-j| \geq 2\,,$$

$$[\text{R2}] \quad \sigma_{i,i+1}\sigma_{i+1,i+2}\sigma_{i,i+1} = \sigma_{i+1,i+2}\sigma_{i,i+1}\sigma_{i+1,i+2}\,.$$

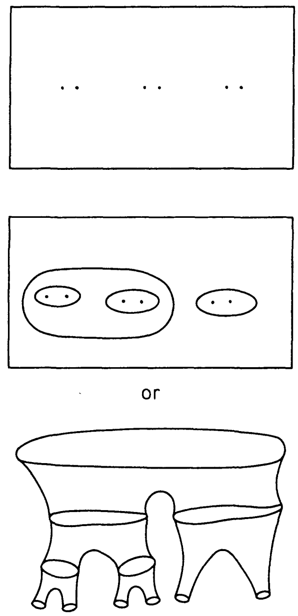
[The hexagon relation tells us that the Moore-Seiberg B matrix obeys [R2], while [R1] is trivial for MS matrices, since the two B matrices would act on disjoint parts of a tensor product.]

The subgroup K_{2n} has 3 generators, $\sigma_{2j-1,2j}$; $\sigma_{2j+2,2j+3}\ \sigma_{2j+1,2j+2}\ \sigma_{2j+3,2j+4}\ \sigma_{2j+2,2j+3}$; and $\sigma^{-1}_{2j+2,2j+3}\ \sigma^{-1}_{2j+1,2j+2}\ \sigma_{2j+3,2j+4}\ \sigma_{2j+2,2j+3}$ (Fig. 6). This subgroup can be characterized as the elements of the braid group which can be extended to the complement of the set of semicircles of the plat in the half space formed by the plane on which they terminate. Since that space is essentially a handlebody, we see that the analogy with Heegaard splittings is growing stronger. We shall see that in a reasonable sense $K_{2n} = B_{2n} \cap N_n$.

Before tackling the general case, let us construct the invariants of classical links. We already expect from Witten's work that each MTC will produce many invariants of links, one for each labeling of the components of the link with labels from the set $R_1 \ldots R_n$.



6



7

Fig. 6. The generators of Kan
Fig. 7. Trinion decomposition of punctured plane

Thus, we need to construct, not a single invariant vector, but a family of them, which can be permuted by elements of Mg.

The setting for applying an MTC to a link in R^3 represented as a plat is the sphere with an even number of punctures. We think of the parallel planes in the plat construction as copies of S^2 completed at ∞, with $2n$ punctures, and opposite orientations. Since the braid group is exactly the mapping class group of the punctured surface, [15] the machinery of the last two chapters gives us a representation of B_{2n} on the vector space V_{Σ_1}, where $\Sigma_{1,2}$ are the two punctured planes. Fact 4 then allows us to find an invariant of classical links, provided we can find states $V_{L,R}$ with suitable invariance under K_{2n}. Specifically, we form the expression

$$I(\mathscr{L}) = \langle V_L \,|\varrho(b)|\, V_R \rangle ,$$

where $b \in B_{2n}$ defines the link $\mathscr{L}$.

We now define a family of states $V_{L,R}$. As we shall see, these states do not only keep track of the coloring of the link, but also of the orientation of its components, and even of a framing of it in R^3. We end up with a set of invariants of colored, framed, oriented knots and links. All of this flows naturally from the machinery we have described.

To describe one of the states we need, we choose a trinion decomposition of the punctured sphere, and label the cuts.

Since the punctures of the planes of a plat are linked in pairs by the semicircles, the states we choose need to contain the information of the pairing. Furthermore, the state needs to have a lot of invariance, which we get in the case of 3-folds by labeling with 1.

The solution to these problems is straightforward. We decompose the punctured plane into trinions by drawing circles around the paired points, then drawing larger circles in whatever pattern we like. We label all the large circles with 1, and the pairs of punctures with R_i, $\hat{R}_i$, i.e. with pairs of dual representations. The MS axioms now assure us that each of the trinions has associated to it a vector space of dimension 1, so we pick the fixed basis elements and tensor them together. This gives us a state $V_{C,\tau} \in V_\Sigma$, where C is the coloring of the punctures and τ is the rest of the trinion decomposition of the punctured plane.

Lemma 5. *$V_{C,\tau}$ is independent of τ.*

Proof. τ is a trivalent tree graph. Any two such objects can be exchanged by a string of the operations F, Ω, Θ, T. Such a string of operations will not affect $V_{C,\tau}$; since it would not act on the trinions containing the pairs of punctures, while the other circles are all labeled 1. Thus, the lemmas in Chapter 4 apply.

Lemma 6. *The subgroup $K_{2n} \subset B_{2n}$ permutes the states V_c among themselves, the permutation on the colors being simply the permutation which a braid induces on the braided points.*

Proof. It will be helpful to remind ourselves how B_{2n} acts on V_Σ. Each σ is just the MS B matrix acting on the tensor factor of Σ corresponding to the trinions being braided.

Now of the three generators of K, (Fig. 6) the first just braids two paired punctures. This can equally well be achieved by applying Ω to a trinion in the basis of Lemma 5. This simply trades the vector space $V^1_{R\hat{R}}$ for the space $V^1_{\hat{R}R}$. If the basis vectors have been chosen sensibly, they are carried into one another with at worst a factor of $e^{\frac{2\pi ic}{24}}$, which we are tolerating as usual. This is because the R and $\hat{R}$ are coupled to 1, so that Ω^2 can only pick up a twist around 1.

The other two generators braid both of a pair of punctures around another. By the pentagon (Fig. 4), this is equivalent to braiding around 1, which is trivial. Q.E.D.

We now need to specify the process by which the link invariants are calculated. We first color, frame, and orient a link. We next represent it as a plat. This forces us to pick up a *particular* coloring of the plane punctures of the plat as well as a particular choice of where to put the "stripes" on the trinions. (Push all the twists off the braid and onto the semicircles). A different plat representation of the same colored framed oriented link would have to choose a different V_C. It is important to note that a change of framing twists a circle labeled with R, not 1, so that it contributes a *significant* phase, not merely the ubiquitous $e^{\frac{2\pi ic}{24}}$. (It may be useful to remark that the phrase "oriented colored link," is slightly misleading. The invariant is preserved if we simultaneously reverse orientation of a component and switch R for $\hat{R}$. Our coloring process would not detect such a switch, so we do not, strictly speaking, specify an orientation.)

If we change a labeled plat by multiplying on the left by an element k of K_{2n}, we get a labeled plat with precisely the change of coloring on the left semicircles induced by k. Thus, if the original plat is generated by $b \in B_{2n}$, the expressions for the invariant would be

$$\langle V_C \,|\varrho(b)|\, V_{C^1}\rangle$$

and

$$\langle V_{\tilde{C}} \,|\varrho(kb)|\, V_{C^1}\rangle\,,$$

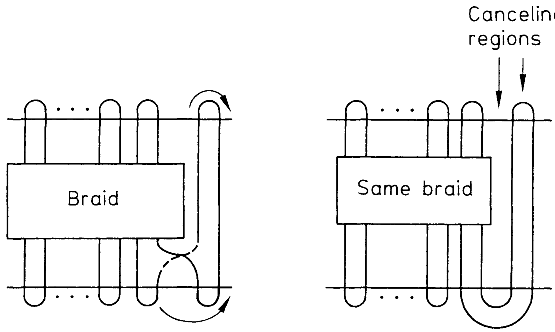


Fig. 8. The twist in a stabilization

where $\langle V_C | \varrho(b) = \langle V_{C^1}|$ by our last lemma. However, by the *canonical* identification of $V_{\hat{\Sigma}} = \hat{V}_{\Sigma}$, the left and right actions of $\varrho(k)$ must be dual. Hence the two conditions are equal.

In order to complete the proof of Theorem 2, in the light of Fact 4, we now need only consider stabilization. Since stabilization relates plats with different numbers of punctures, and since the configurations on the trinions near the new punctures are the same for any stabilization; we could always absorb it into a normalization. In fact, it is better not to. A simple twist (Fig. 8) turns a stabilization into two symmetric pieces which cancel because they are matrix elements of dual matrices. What remains is the phase from the twist itself, which keeps track of the change of framing which a stabilization induces by putting an extra loop into the link.

We have now proven the following theorem.

Theorem 2. *The above expression* $\tilde{I}(\mathscr{L})$ *applied in any* MTC *gives an invariant of framed labeled, oriented links. A change in the framing of one component changes the invariant by a phase depending on the label of the component. Simultaneously reversing one component's orientation and dualizing its label leaves* $\tilde{I}(\mathscr{L})$ *unchanged.*

This theorem was proven by J. Fröhlich and C. King in another context, using a different combinatorial technique [28]. Jones [29] knew that his invariant could be calculated from the plat representation of a knot.

Now we wish to extend our result to links in arbitrary closed oriented 3-folds. In order to do so, we need a synthesis of plats and Heegaard splittings which I call "generalized plats." As far as I know, this work has not been done before.

A solid handlebody can be sliced into a one parameter family of copies of its boundary Σ, with a union of circles and intervals, called the *spine*, left over at one end. [Think of dipping a piece of wire with loops into candlewax over and over (Fig. 9).]

If we think of a closed oriented 3-fold as the union of two solid handlebodies, i.e. pick a Heegaard splitting for it, we can think of it as $\Sigma \times (0,1)$, with two spines attached at 0 and at 1.

Any link in M can be slightly deformed so that it misses both spines, and crosses the foliation by copies of Σ transversely, except at isolated points where the $(0,1)$ coordinate attains a local extremum. The $(0,1)$ function on M now assumes a role analogous to the left-right axis in the description of classical plats.

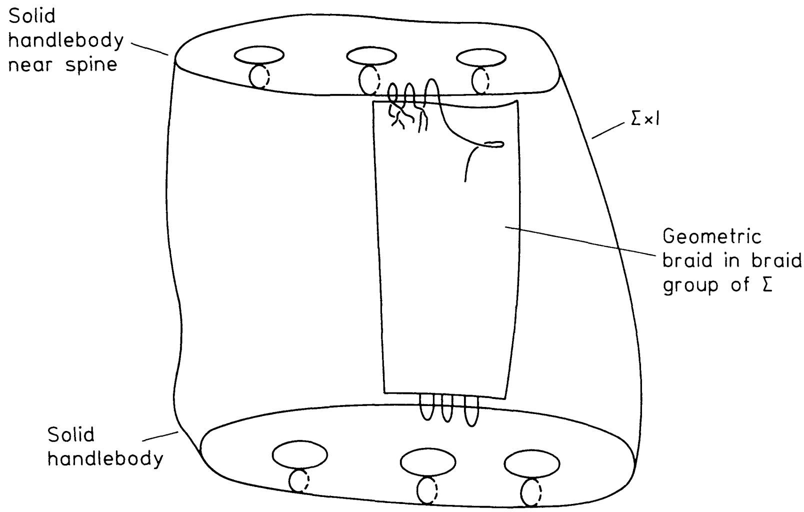


Fig. 9. The spine of a handlebody

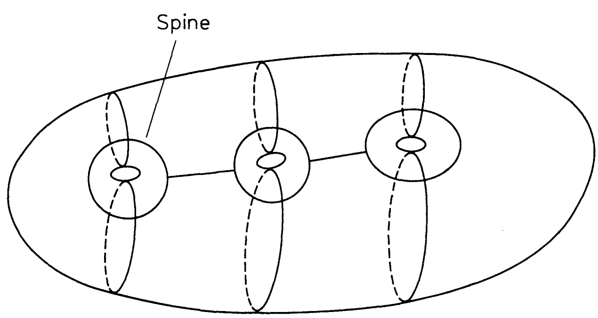


Fig. 10. A generalized plat (a link in a 3-manifold with a heegaard splitting

We can now pull the local minima toward the spine at 0 unti they all end in small arcs in $(0, \varepsilon) \times \Sigma$, i.e. a "thin" solid handlebody. Doing similarly near 1, we end up with a generalized plat. This motivates the following definition.

Definition. A generalized plat in $M = H_1 \cup_\varphi H_2$ is the union of a set of n small semicircles in a smaller handlebody $H_{\varepsilon_1} \subset H_1$ [where $H_1 = \Sigma \times (0,1) \cup S$ (S is the spine of H_1) and $H_{\varepsilon_1} = \Sigma \times (0, \varepsilon) \cup S$] which all are contained in a small ball and are not intertwined, a similar set in H_2, and a $2n$ stranded geometric braid in $M - (H_{\varepsilon_1} \cup H_{\varepsilon_2}) \cong \Sigma \times I$ (where I is an interval), which joins the endpoints of the two sets of semicircles (Fig. 10).

The result of the discussion leading up to this definition is to prove:

Proposition 1. *Every link in a compact oriented 3-fold admits a generalized plat representation.*

Thus we are able to describe links in 3-folds in terms of elements of braid groups of higher genus surfaces. (For a discussion of general braid groups see [15].)

The description of 3-folds by Heegaard splittings and of links by plats admits a further unification. The mapping class group of a punctured surface is a sort of product of the braid group of the compact surface with its mapping class group. More precisely, there is an exact sequence [15]

$$* \to B(g,n) \to M(g,n) \to M(g,0) \to 1\,,$$

where $B(g,n)$ is the n-point braid group on the surface of genus g, $M(g,n)$ and $M(g,0)$ are the mapping class groups of the punctured and unpunctured surface. The kernel $*$ is usually trivial, and always small.

It seems tempting to construct both the 3-fold and the link at once from an element of $M(g,n)$ and prove the invariance all at once from a suitable synthesis of Facts 2 and 4. We do not achieve that here. Nevertheless, the fact that links and 3-folds are both described by surface maps (a braid can be thought of as a map of the punctured surface to itself) is a strong hint that the most natural object to apply MTC's to is the *unified* object manifold-with-link. This is not as strongly indicated in Witten's picture, and, as we shall see, it is very suggestive for quantum gravity.

We shall now construct the invariant for a link in a general 3-fold, and prove its invariance in a somewhat piecemeal way. Namely, we shall first consider isotopies of a link described plat while fixing the Heegaard splitting, then discuss the problem of changing the Heegaard splitting. We do not present a complete proof, but the omitted points do not seem likely to cause trouble.

We need to prove an analog of Birman's theorem on isotopies of plats for generalized plats. The statement of Birman's theorem is stronger than our Fact 4. It says that if two plats are isotopic then after stabilizing both sufficiently often they are in the same double coset of K_{2n} in B_{2n} for some n.

Birman's proof has several stages. First, she restricts attention to piecewise linear links and simple isotopies which exchange one leg of a triangle for the other two legs. Next, she shows that such moves can be written as products of only two types of moves on plats, abandoning the linear structure and working only on isotopy classes. Finally, she checks that the algebra of these moves allows them to be written in a special order.

One of the moves is stabilization or its inverse. The other is a move which pulls a small loop which ends as a semicircle of the plat in some distance into the braid, then pushes it back out on some other path. The special order in which the moves can be written is stabilizations; loop moves on left $\cdot$ loop moves on right $\cdot$ destabilizations. The loop moves are contained in the subgroup K_{2n}, so the theorem is proved.

A direct imitation of Birman's proof in a 3-manifold would be difficult, since no natural linear structure is available. However, isotopies are products of local isotopies, so we can break our manifold up into simple pieces, and apply her result to each piece.

Let us first consider only isotopies of generalized plats which avoid the spines of the two handlebodies. Thus, we are considering isotopies of plats in $\Sigma \times (0,1)$. (When we remove the spines, the surface map of a Heegaard splitting becomes irrelevant.) Let us further choose a set of cuts in Σ which reduces it to a disk. Now, if we consider any isotopy which avoids the cuts, Birman's theorem tells us that it is a product of her two types of moves. If we wish to consider an isotopy which crosses a cut, we can arrange to have it live in another disk in Σ. Since stabilizations can be moved around freely by products of the other move, it is not necessary to add stabilizations belonging to the region around a cut. Thus we find that in order to include all isotopies which miss the spines, we need add to K_{2n} only a move which drags a semicircle around a cut (i.e. around a non-trivial cycle in Σ). There remains to consider isometries which cross the spines. There can be described by small triangles which intersect the spine at a single interior point. In the generalized plat picture, this translates into dragging a single puncture around a non-trivial cycle in Σ which bounds a disk in H_Σ [half of $H^1(\Sigma)$ does].

It is easy to see that Birman's arguments getting these moves in the right order go over.

We have shown the following:

Proposition 2. *The equivalence relation between elements of the braid group $\tilde{B}_{2n}$ of Σ which generate isotopic links is generated by stabilization, and double cosets of the subgroup $\tilde{K}_{2n}$ generated by Birman's 3 moves and the two moves described above: dragging one puncture around a loop in Σ which bounds in H_Σ or two paired punctures around any loop.*

In order to construct the invariant for a link in a 3-fold, it is necessary to describe a set of states $V_C \in V_\Sigma$ for Σ a higher genus surface with $2n$ punctures and C a coloring of the punctures with the label of a MTC. We then need to check that they are permuted appropriately by $\tilde{k}_{2n}$ and the generators of $N_g \subset M_g \subset M(g, 2n)$.

The formula for the invariant of a colored framed oriented link in a 3 fold would then be as follows: construct the element of $M(g, 2n)$ corresponding to the surface map of a Heegaard splitting for the manifold and the braid group elements of a plat description of the link under the exact sequence mentioned above. Then sandwich the representation matrix of this element of $M(g, 2n)$ between the appropriate V_C's, and normalize as we would just for the unpunctured surface map.

The definition of the V_C's is not difficult. We pick trinions for each pair of punctures as before, label the outside circles with 1, then extend to any decomposition of Σ, using only circles which bound discs in H_ε, and label all the cuts 1. Tensoring together the basis vectors for the $V^1_{R\hat{R}}$'s and V^1_{11}'s yields the states V_C.

We now need to check several things. The generators of $\tilde{k}$ must permute the V_C's. It is easy to see that this is so, since the generators of $K_{2n} \subset \tilde{K}_{2n}$ have already been studied, while the two new generators both amount to braiding two states, one of which is labeled 1. In one case, the outer circle around two paired pictures is labeled 1. (The pentagon allows us to braid the outer circle instead of both punctures.) In the other case, we drag a puncture around a circle which bounds, and is therefore labeled 1. (The proof that *all* bounding circles are labeled 1 is as in Sect. 4.)

The action of the generators of N_{2g} on the colored vacuua is as in Sect. 4, provided we are sensible enough to situate the outer circles around the punctures away from the cuts.

To complete the proof of invariance, we need to examine the action of N_g on generalized plats, which boils down to studying the algebra of $M(g, 2n)$. Since I have not yet done that, my proof is not quite complete. The algebra of $M(g, 2n)$ is fairly straightforward, so I expect no surprises.

Let us therefore state:

Almost Proved Theorem I. *Any MTC gives rise to an invariant of framed oriented, colored links in a closed oriented 3-manifold.*

The analogy between 3-folds and links goes farther yet. Consider a 3-ball with some simple arcs removed. Call such a space a preplat. A link represented as a plat is two preplats glues by a surface map. A branched cover of a preplat over the removed arcs gives rise to a handlebody. Thus Heegaard splittings for 3 folds can be constructed as covering spaces of plats.

If we had an MTC with two colors, one of which was related to the other on a branched cover over the puncture, then we would have a symmetry relating invariants of links in 3-folds to invariants of branched covers of the 3-folds over components of the links.

Such an MTC seems very plausible. If we started with a super symmetric RCFT, the fermionic states look very like the bosonic states on branched covers. I conjecture that the symmetry described above can be realized as *topological supersymmetry.*

An example of topological supersymmetry may already be known. The Alexander polynomial [9] can either be calculated from the Reidemeister torsion of a link, or from the cohomology of an infinite cyclic cover. (Since parastatistics exist in $2-D$, we are not bound to use only double covers.) Further evidence that the Alexander polynomial is related to supersymmetry has been discovered by Reshetikhin (personal conversation) who can calculate it from the quantum version of the supergroup $SU(1\,|\,1)$. This observation has also been made by L. Kauffman [16].

Finally, we wish to observe that everything we have said about links is still true if links are generalized to knotted trivalent graphs. L. Kauffman has assigned invariants to knotted trivalent graphs in a slightly different context [17] see also [8]. In doing so, he used a combinatoric description for knotted graphs completely analogous to one for links. The description is slightly different from the plat one, but it would be very surprising if it did not translate. The only new move needed is to slide a braiding past a vertex, but that is just the pentagon.

Let us state.

Almost Proved Theorem II. *Any MTC gives rise to an invariant of oriented labeled framed trivalent knotted graphs in an oriented closed 3-manifold.*

We should remark that in labeling a graph edges receive labels from $R_1 \ldots R_n$, and vertices are labeled by vectors in $V^{R_1}_{R_2 \cdot R_3}$, where the R_i are the labels of the edges at the vertex.

(Proofs of the gaps in this chapter will be completed elsewhere.)

VI. Manifolds with Boundary

Witten applied his path integral also to manifolds with boundary, in which case it gave him a state in V_Σ, for Σ the boundary. We are actually able to prove a slightly stronger result.

Theorem 3. *Let* (M^3, ∂) *be a compact oriented 3-fold with boundary. Let the boundary* ∂ *be separated into two disjoint sets of components in any way* $\partial = \Sigma_1 \cup \hat{\Sigma}_2$. *Then an* MTC *gives rise in a natural way to a homomorphism* $V_{\Sigma_1} \to V_{\Sigma_2}$. *(Note that the orientation of* Σ_2 *is reversed.)*

Note that a homomorphism $V_{\Sigma_1} \to V_{\Sigma_2}$ is the same as an element of $V_{\Sigma_1} \otimes \hat{V}_{\Sigma_2} = V_\partial$. Thus, we are producing the same object as Witten.

As usual, the map is ambiguous by an overall $\left(e^{\frac{2\pi i c}{24}}\right)^n$. It is really a stroke of luck which allows us to prove this theorem. Michael Motto's thesis (1988) [18] developed a combinatorial description of 3-folds with boundary which parallels Heegaard splittings and interfaces nicely with conformal field theory.

To explain Motto's work, we begin by describing a space called a *compression body*. A compression body is a handlebody with part of a neighborhood of the spine removed. Another way to describe it is to take a thickened neighborhood of an oriented surface in R^3, and add in some thickened neighborhoods of some disks in the interior which bound some circles on the surface. (The circles then become "compressible," whence the name.)

A compression body is a 3-fold with boundary. The boundary comes in two parts: the *outer boundary*, which is just the original surface; and the *inner boundary*, which is a union of disjoint surfaces of lower genus. [We arrange the contracted disks to be non-intersecting.]

The structure which Motto develops for 3-folds with boundary is called a generalized Heegaard splitting.

Definition. If a 3-fold with boundary (M, ∂) is written as a union of two compression bodies with their outer boundaries identified, then we have a generalized Heegaard splitting for (M, ∂).

Motto proves

Fact 5. *Every 3 fold with boundary (M, ∂), with an arbitrary division of its boundary components into two disjoint sets $\partial = \Sigma_1 \cup \Sigma_2$, admits a generalized Heegaard splitting, where Σ_i is the inner boundary of compression body H_i.*

As usual, we may choose Σ standard, so that (M, ∂) is essentially specified by an element of M_g for some g. Once again we need to know the equivalence relation.

Fact 6. *The equivalence relation between elements of mapping class groups which give homeomorphic 3-manifolds with boundary is generated by stabilization, and by left and right multiplication by the subgroups which extend to the two solid handlebodies.*

(We slightly restate Motto's result.)

Now we need homomorphisms, rather than elements of vector spaces.

Lemma 7. *If $\tilde{H}$ is a compression body with inner boundary Σ_I and outer boundary Σ_0, then for any* MTC *there is a natural homomorphism $h : V_{\Sigma_1} \to V_{\Sigma_0}$. This map is invariant under the subgroup $N_{\tilde{H}} \subset M_g$ of surface maps of Σ_0 which extend to $\tilde{H}$.*

We construct the map as follows. On the outer boundary Σ_0 draw copies of the circles which are compressed in $\tilde{H}$. Contract each circle down to a point. We are now left with a copy of Σ_I with some extra punctures glued together in pairs. If we label all the punctures with 1, it follows from the MS axioms that the vector space associated to this new surface Σ_M is isomorphic to V_{Σ_I}. This is because adding a puncture labeled with 1 to a surface can be done by inserting a trinion with labels ${}^{1}_{R\hat{R}}$, and $V^{1}_{R\hat{R}} \cong \mathbf{C}$.

Taking any state in our new vector space V_{Σ_M} and pulling it back to V_{Σ_0} gives a map $V_{S_M} \to V_{\Sigma_0}$.

We now have two worries: the identification of Σ_m and Σ_I, and the invariance under $N_{\hat{H}}$. Fortunately, these problems cancel one another.

The reason for this is that the extension of self-maps of Σ_0 to $\tilde{H}$ is a very rigid business. If we first think only of a self map of a thickened shell $\Sigma \times [0,1]$, we would find that any self map of such a shell would give a homotopy between two self maps of Σ. However, for surfaces, homotopy implies isotopy [19]. Thus the two end maps on the thickened shell give the same element of M_g.

Now let us suppose we have a homeomorphism $\varphi : \Sigma_0 \to \Sigma_0$ which extends to $\tilde{H}$. It must permute the circles on Σ_0 which contract in $\tilde{H}$ among themselves (up to an isotopy). A small rearrangement of the extension of φ allows us to assume that it permutes the disks. Cutting the space H_Σ along the disks cuts it into thickened shells on lower genus surfaces. The extension of φ must be the same, up to isotopy, on the inside and outside of each. Since the machinery of an MTC only depends on isotopy classes of trinion decompositions, we are left with an invariant map $V_{\Sigma_I} \to V_{\Sigma_0}$ as we wished.

In order to prove Theorem 3, we note that reversing orientation gives a map $\tilde{h} : V_{\hat{\Sigma}_0} \to V_{\hat{\Sigma}_I}$.

If we have a generalized Heegaard splitting for a 3 fold with boundary $(M, \partial) = \tilde{H}_I \cup_\varphi \tilde{H}_2$; then the two outer boundaries have reversed orientation. Thus φ is naturally a surface map from $\Sigma_{0,1}$ to $\hat{\Sigma}_{0,2}$. We write $h_1 \circ \varrho(\varrho) \circ \hat{h}_2$ as our map.

Proof of Theorem 3. The double coset invariance follows from the invariances of h_i. The invariance under stabilization is trivial. Hence Theorem 3 follows from Fact 6.

The invariance of these maps tells us that components of the boundary of a 3-fold have some sort of "intertwinedness" which we can now measure.

In order to complete our reconstruction of Witten's topological conformal field theory, we need to prove an extension of all our theorems up to this point which would study a 3-fold with boundary containing a knotted labeled framed oriented trivalent graph, with some open components which end on the boundary. This seems straightforward, but tedious. Perhaps the existing results would cover the most general case with a little manipulation.

The other omission in completing our construction of a TQFT is a proof of the gluing axiom. That could be accomplished by finding a generalized Heegaard splitting for the "big" space which included the 2 "little" splittings and not much else. I will attack this elsewhere, but it appears straightforward.

VII. Extensions to Four Dimensions

The contact we have developed between $3-d$ topology and $2-d$ physics is strong and elegant, but there is one irritating point; namely, the ubiquitous phase ambiguity. One would like to remove this in a geometrical way. Witten [20] suggests including a framing of the 3-fold, but this has the unappealing feature that we are constrained in the end to resort to "weak framings." In any case, regularizing the Chern Simons path integral seems very challenging.

I wish to propose another solution: regarding the 3-dimensional objects as boundaries of 4-folds with boundary. As I shall argue, the relationship between the combinatorial topology of 3-folds and 4-folds strongly suggests that this approach should be fruitful.

We shall first consider the case of a 3-fold, then briefly speculate on including links. Every oriented 3-fold is the boundary of a simply connected 4-fold. This simple fact can be proved many ways. The combinatoric proof, which we recite here, actually shows us how to construct many inequivalent simply connected 4-folds with a given 3-fold as boundary. It is easy to assign these 4-folds invariants from a MTC. The invariants of the 4-folds do not share the ambiguity of the 3-fold.

In order to construct our 4-folds, we must convert our Heegaard splitting description of a 3-fold into another type of combinatoric description, called a Dehn surgery [9]. A Dehn surgery is a construction of an oriented 3-fold which begins with a link in S^3, each component of which is labeled by a rational number. Small tubular neighborhoods of each component are removed, then reattached by some non-trivial maps of the tori which are their boundaries. The rational numbers specify the surface maps. It is a classical result that every oriented 3-fold M has a Dehn surgery presentation.

A standard proof of that fact can be obtained, beginning with a Heegaard splitting for the 3-fold. The Heegaard splitting can be described by the isotopy class of a surface map $\varphi : \Sigma \to \Sigma$. The class of φ in M_g can be written as a product of simple surface maps called Dehn twists [21]. A Dehn twist is a map of Σ to itself which is the identity outside of a small neighborhood of a circle in Σ. Near the circle, the map can be described by cutting Σ at the circle, twisting one side by 2π, and regluing. The theorem of [21] actually shows that any element of M_g is a product of Dehn twists around a fixed set of circles on Σ, which can be taken as the cuts for a trinion decomposition of Σ, plus others.

Given a factorization of $\{\varphi\} = D_1 \dots D_n$, where the D_i are Dehn twists, we can produce a Dehn surgery as follows. The first D can be thought of as changing the manifold only in a small tubular neighborhood around the cut. The effect of this change is a Dehn surgery around the cut with index ± 1. We can pull this small solid torus away from Σ without changing the manifold. If we continue in this way, we end up with a Dehn surgery, and change φ so the $\{\varphi\} = 1$. However, a Heegaard splitting with $\{\varphi\} = 1$ produces S^3, so we have a Dehn surgery which changes S^3 to our M. This proof actually proves more. The Dehn surgery we have constructed has only integer coefficients, hence it belongs to the class of "honest surgeries."

Honest surgeries are so called because they give instructions for attaching 2-handles to the 4-ball so that the 3-fold produced is the boundary of a 4-dimensional 2-handlebody. Thus a particular decomposition of φ into a product of a certain fixed set of generators gives rise to a particular simply connected 4-fold with M as boundary.

Now remember that the phase ambiguity of $\tilde{I}(M)$ comes from the fact that different products of MS matrices which produce the same element of Mg can fail to be equal by a phase factor. Thus, choosing a particular factorization of φ removes the ambiguity. We have proven the following:

Proposition 3. *Any oriented closed 3-fold M is the boundary of a family of simply connected 4 folds. Any* MTC *assigns invariants to these 4-folds which are unambiguous, and equal to $\tilde{I}(M)$ up to a power of $e^{\frac{2\pi i c}{24}}$.*

We are naturally led to the following:

Conjecture. *Any* MTC *gives an invariant of simply connected 4-folds with boundary. This invariant is an invariant of relative cobordism wrt the boundary.*

Plausibility argument. The relative cobordism group of 4-folds with boundary M is Z. This Z can be taken as giving the Z in the universal central extension $\tilde{M}$, of M_g [22], if M is taken as coming from a Heegaard splitting. Thus, we are now using elements of $\tilde{M}_g$, which is represented, rather than projectively represented, in conformal field theory [13].

Second Plausibility Argument. The Witten path integral uses the Chern Simons form integrated on a 3-fold. The Chern Simons form is the boundary contribution to the integral of the first Pontrjagin form. Thus it is awfully tempting to rewrite the Witten integral as a 4-d path integral over some M^4 which bounds M^3.

It is provocative to speculate on the 4-d significance of links in M^3. The abelian version of Witten's integral can be interpreted as a sum of linking numbers [23]. The linking number of two loops in S^3 is the same as the intersection number of 2 disks in B^4 for which they are the boudaries.

It is tempting to think of a link invariant as an invariant of surfaces bounding the link components.

The picture here is reminiscent of Donaldson's invariant polynomials, but extended to 4-folds with boundary. Those polynomials are also "non-abelian intersection numbers," in a different sense.

There is actually a combinatoric way to relate choice of surfaces, bounding a link to factorizations in $M(g,n)$. The choice of factorization would tell us when to pull a link thru a 2-handle in M^4. This approach will be investigated further.

VIII. RCFT and Quantum Gravity. Spin Networks

The most general geometric setting for the invariants we have been discussing is a 3-fold with boundary which contains a link or trivalent graph. My interest in finding invariants for such a configuration antedates Witten's paper by six months, and derives from a completely separate source: the quantization of general relativity, as attempted in the Ashtekhar variables [5].

Although there has been a great deal of work recently studying $2+1$ dimensional gravity as a topological QFT, [24] there has been a perception that in $3+1$ dimensions, this approach would be impossible. This is because $3+1-D$ gravity has local excitations which propagate from place to place; which surely does not have a "topological" feel.

Nevertheless, I want to argue that $3+1-D$ gravity can be approached by the methods of MTC's and TQFT. The price one has to pay for the propagating modes is precisely the one we can afford to pay: inclusion of links or graphs in the 3-fold.

Let us recapitulate the line of development which led from general relativity to link invariants. Ashtekhar showed that the constraint equations for general relativity became more tractable if a certain connection, A, on a 2-plane complex bundle over a 3-fold was substituted for the 3-metric as configuration space variable. (The definition of A uses a half spin bundle, so it is chiral.) Thus the states of our quantum theory are to be functions $\psi(A)$ which must be both gauge and diffeomorphism invariant.

The next step in our development is the loop space formulation of Smolin and Ravelli [6]. They proposed substituting traced holonomies around closed loops as variables, since these are automatically gauge invariant and suffice to distinguish gauge classes of A's. The change from states $\psi(A)$ to states $\Psi(L)$ involved a dualization, which enormously simplified the constraint algebra. Thus, we are led to consider "measures" $\mu(A)$, [they are *not* measures in the precise mathematical sense, but elements in the dual of some restricted class of functions $\psi(A)$] which could give rise to diffeomorphism invariant states $\Psi(L)$. A diffeomorphism invariant state $\Psi(L)$ *is* an invariant of link classes. (Link classes are equivalence classes of embeddings under ambient diffeomorphism.) Thus Smolin, Ravelli and I were led to the following problem: *Do diffeomorphism invariant "measures" exist on the space of connections over a 3-fold, which, when integrated against the traced holonomy of a set of curves, give a non-trivial link invariant?* This could be taken, in one interpretation, as the problem of finding states for quantum gravity.

Witten's paper came barely a month after we asked this question. Formally, his path integral provided precisely what we were trying to find.

Further work by Smolin [25] showed that it was in fact necessary to include trivalent graphs as well.

In all this development there was a gaping hole in this theory; namely, how to make a physical interpretation of it. I took the point of view that in order to have physical observables, it is necessary to divide the universe into two parts, system and observer. Hence a physical Hilbert space could only be assigned to a part of a universe, which I wanted to interpret as a manifold with boundary.

Thus the problem arose to find vectors associated to manifolds with boundary containing links or graphs.

By now the parallels between this approach to quantizing gravity and the TQFT's constructed from conformal field theory should be evident. It is tempting to regard the gluing axiom of a TQFT as a consistency relation between different observers. To sum up, we could say that $3 + 1$ gravity is a theory whose states correspond to TQFT's. (An *arbitrary* state is just a functional of link classes, which is much flabbier than a TQFT, but only states corresponding to TQFT's would admit a physical interpretation, by "factorizing" over subdomains corresponding to experimental systems.) The TQFT would give the "state of the universe," while subsystems would have state vectors in the vector spaces provided by TQFT.

We are thus led to try to explain physics as we know it by choosing the right conformal field theory. This is intriguing, especially when we remember that CFT was invented to describe ground states for string theory, which, it was hoped, would explain the physics we see. Is it too much to hope that choosing a WZW model as RCFT would even couple gravity to gauge theory?

It is still very hard to see how to interpret a TQFT as a physical theory. There is some hope of achieving this by means of the relationship with a third field of mathematical physics–namely, spin networks.

If we state the problem we face in its simplest terms it is this: given an evaluation machine which assigns numbers to trivalent graphs which are closed and vectors to ones with loose ends, how do we achieve a picture of a physical space-time with some sort of matter in it?

Incredibly enough, this is exactly the problem which Penrose [26] attacked in his work on spin networks. A spin network is a labeled trivalent graph. In the original case, the labels are representations of $SU(2)$; there is a fairly direct generalization to other lie groups. The evaluation of trivalent graphs is very close to the invariant for the WZW models. In the language of Moore and Seiberg, it is the "classical" conformal field theory which they approach in the classical limit.

Penrose gives a line of argument in which certain combinations of evaluations are interpreted as angles between macroscopic physical vectors. He finds that for "macroscopic" objects the ordinary laws of flat geometry emerge. Subsequently [27] this was interpreted as a path integral approach to euclidean signature quantum general relativity in dimension 3. The equations of motion in that dimension imply flatness.

When larger groups than $SU(2)$ are used, the matter is less well understood, but we seem to get a picture of a 3-fold mapped into some higher dimensional symmetric space.

The most striking difference between the evaluation for spin networks and the one for RCFT is that in spin networks the holonomy of one edge around another is trivial. In spin networks, it is not necessary to distinguish over from under crossings when a diagram is projected on a plane. Otherwise, the rules for evaluating spin notworks closely parallel the skein rules for the Jones polynomial (which is the invariant from the $SU(2)$ level k WZW models). I owe this observation to L. Kauffman.

It is very appealing now to translate Penrose's interpretation to invariants derived from RCFT's. The addition of holonomies into the situation looks optimistic for producing a curved space-time picture. This direction will be further investigated.

The picture which seems to be possible here is that the universe initially chooses a state, which selects a RCFT, which tells us what sort of matter fields we must include. Anyone familiar with string theory must find this picture familiar.

We are left with the problem of picking the right RCFT (or perhaps only the right MTC). In string theory, the potential solutions to this involved numerical coincidences around $c = 24$. In our current setting, RCFT's with $c = 24$ are special, in that they give rise to invariants without ambiguity in $d = 3$. J. Lepowsky informs me that he may be able to produce MTC's of $c = 24$ from constructions on the Leech lattice.

In the 3-dimensional setting, the numerical coincidences around 24 pick up another aspect, since framed cobordism of 3-folds is Z_{24}. (R. Kirby pointed this out to me.) Thus, any RCFT with integer c, such as a free boson, would produce an *unambiguous* framed cobordism invariant of framed 3-folds. This is another suggestion that thinking of 3-folds as boundaries may be fruitful.

Acknowledgements. I wish to thank J. Birman, I. Frenkel, H. Fröhlich, C. Gordon, V. Jones, L. Kauffman, R. Kirby, J. Lepowsky, W. Massey, G. Moore, N. Reshetikhin, L. Smolin, E. Witten, and G. Zuckerman for helpful conversations. C. Gordon was very kind to send me excerpts from M. Motto's thesis. The staff of the mathematics department at Yale has astonished me at every point with their efficiency and kindness.

References

1. Witten, E.: Quantum field theory and the Jones polynomial. IAS preprint HEP 88/33
2. Crane, L.: Topology of 3-manifolds and conformal field theory. Yale University CTP preprint YCTP-P8-P9, April, 1989
3. Moore, G., Seiberg, N.: Classical and quantum conformal field theory. Commun. Math. Phys. **123**, 177–254 (1989)

$3\frac{1}{2}$. Moore, G., Seiberg, N.: Taming the conformal zoo. IAS preprint HE 89-6

$3\frac{3}{4}$. Witten, E.: Gauge theories and integrable lattice models. IAS preprint HE 89-11; Kauffman, L.: Knot theory and statistical mechanics. Preprint

4. Atiyah, M.: New invariants for three and four manifolds. In: The mathematics Heritage of Herman Weyl. Providence, RI: Ams 1988
5. Ashtekar, A.: Phys. Rev. Lett. **57**, 24 (1986), Phys. Rev. D **36**, 1587–1603 (1987)
6. Ravelli, C., Smolin, L.: Loop Space Representation of quantum general relativity, preprint IV 20, Phys. Dept., V. di Roma "La Sapienza" 1988
7. Frenkel, I.: Personal communication
8. Reshetikhin, N., Turaev, V.: Ribbon graphs and their invariants derived from quantum groups and invariants of 3-manifolds via link polynomials and quantum groups. Preprint
9. Rolfson, D.: Knots and links. Wilmington, De: Publish or Perish, 1976
10. Birman, J.: Poincaré's conjecture and the homotopy group of a closed orientable 2-manifold. J. Aust. Math. Soc. XVII part 2, pp. 214–221
11. Singer, J.: Three dimensional manifolds and their Heegaard diagrams. Trans. Am. Math. Soc. **35**, 88–111 (1933)
12. Suzuki, S.: On homeomorphisms of a three dimensional handlebook. Canad. J. Math. **29**, No 1, 111–124 (1977)
13. Friedan, D., Shenker, S.: Nucl. Phys. B **281**, 509 (1987)
14. Moore, G.: Personal communication
15. Birman, J.: Braids, links and Mapping class groups, No 82, Ann. Math. Studies PUP 1975
16. Kauffman, L.: Spin networks and knot polynomials. Preprint
17. Kauffman, L.: Invariants of graphs in 3-space. Trans. Am. Math. Soc. **311**, 697–710 (1989)
18. Motto, M.: Thesis, University of Texas at Austin, Mathematics Department, 1988
19. Massey, W.: Personal communication
20. Dijkgraaf, R., Witten, E.: Topological gauge theories and group cohomology. IAS preprint HEP 89-93
21. Lickorish, W.B.R.: A finite set of generators for the homeotopy group of a 2 manifold. Proc. Comb. Phil. Soc. **60**, 769–778 (1964)
22. Gordon, C.: Personal communication
23. Polyakov, A.M.: Fermi-Bose transitions induced by gauge fields. Mod. Phys. Lett. A **3**, 325 (1988)
24. Witten, E.: $2+1$ Dimensional gravity as an exactly solvable system. IAS preprint, HE 88-132
25. Smolin, L.: Personal communication
26. Penrose, R.: Angular momentum: an approach to combinatorial space-time in quantum theory and beyond. Bastin, T. (ed.), CUP 1971
27. Ponzano, G., Regge, T.: Semi-classical limit of Racah coefficients in spectroscopic and group-theoretical methods in physics. Bloch, F. (ed.). Amsterdam: North-Holland 1968
28. Fröhlich, J., King, C.: The Chern-Simons theory and knot polynomials. ETM 89-9 and Two-Dimensional Conformal field Theory and Three Dimensional Topology ETM 89-10, preprints
29. Jones, V.: Hecke algebra Representations of Braid groups and link polynomials, preprint

Communicated by S.-T. Yau