Quantum Gravity

John C. Baez Miguel Carrión Álvarez

June 10, 2005

Part I

Spin networks/Field theory

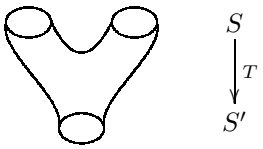
Introduction

One of the things that makes mathematics fun is its relation to physics. It's not surprising that one can build beautiful self-consistent mathematical structures and prove theorems about them. What is surprising and mysterious is that some of these structures are well suited to describing aspects of the world we live in. We call these aspects the 'laws of physics'. Why does the universe have mathematical laws? Nobody really knows. Lots of people have thought about this question, but they didn't get very far. Perhaps it is too soon to answer this question. After all, we don't even fully know what the laws of physics *are* yet. So we should probably start by figuring out *what* they are, and then think more about *why* they exist.

From this, we are inevitably led to quantum gravity. After all, one of the big problems in figuring out the laws of physics is that right now there are two sets of laws, general relativity and quantum theory, which do not seem to get along well. Quantum gravity is an attempt at reconciling them.

To really understand the latest ideas about quantum gravity one must first know general relativity and quantum theory. So this course really should have an introduction explaining these subjects before we go into quantum gravity. Unfortunately, this introduction would need to be very long! To get around this problem, we will take two complementary approaches. In Track 1, we will do things that do not assume any knowledge of general relativity or quantum field theory. In Track 2, we will assume the reader is already rather familiar with both these subjects. Eventually the two tracks will merge.

In general relativity there is a thing called space-time, and we can think of it as made of slices which we call "space" evolving into one another as time passes.



Spacetime and space are smooth manifolds in this theory, so the fundamental mathematics one uses in general relativity is differential geometry.

In quantum mechanics, on the other hand, one uses completely different mathematics, namely Hilbert spaces (roughly speaking, vector spaces with an inner product). A unit vector in the Hilbert space $\psi \in H$ is taken to describe a "state" that the world can be in. There are also linear operators

$$\begin{array}{c} \psi \in H \\ \Big\downarrow T \\ T(\psi) \in H' \end{array}$$

which describe how things can change (a note on terminology: the terms "linear map", "linear operator" and "linear function" will be used interchangeably throughout this seminar).

Quantum mechanics is therefore mainly based on algebra, which looks nothing like the geometry of smooth manifolds on which general relativity is based, and so quantum gravity is like trying to mix oil and water. Just about the only thing these theories have in common is the way in which both talk about states that undergo some transformation. This analogy is best displayed diagrammatically—just look at the above diagrams—and with this motivation we can plunge right into track 1.

Chapter 1

Diagrammatic Methods for Linear Algebra (I)

We are going to study a diagrammatic notation for doing linear algebra. The amazing thing about it is that, if one takes the diagrams we will be using really literally, one starts to see how space-time might be built of just these kinds of diagrams and nothing else.

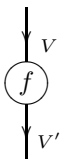
The basic objects in our theory will be vector spaces (which we will usually take to be finite-dimensional and complex). Let us now exhibit how different operations of linear algebra can be represented diagrammatically:

1.1 Linear Maps

A linear map is a function

$$f\colon V \to V' \qquad \text{such that} \quad f(\alpha v + \beta w) = \alpha f(v) + \beta f(w) \qquad (\alpha, \beta \in \mathbb{C}).$$

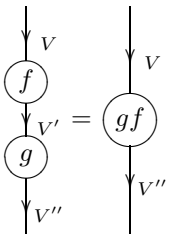
One well-known way to represent linear maps is with matrices, but we will introduce diagrams for that purpose.



A linear map is represented by the name of the map surrounded by a "blob" with arrows sticking out at the top and bottom of the blob. Arrows are labeled by the name of the vector space they represent. The downward direction represents the passage of a "metaphorical time", in other words, from top to bottom one draws the domain, the function and the codomain.

1.2 Composition of Maps

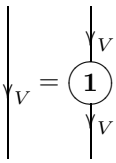
Given linear maps $f\colon V \to V'$ and $g\colon V' \to V''$ we can compose them to obtain $gf\colon V \to V''$ and we draw the composition by sticking the diagrams for f and g one on top of the other.



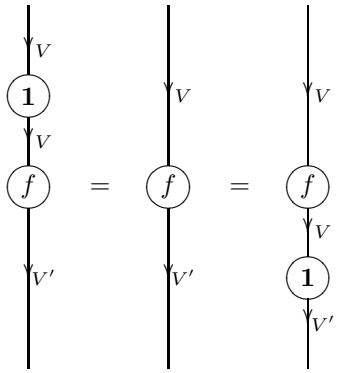
If you have a set, it always comes with an identity function "at no extra cost". Similarly, every vector space is equipped with an identity linear map

$$\begin{array}{rccc} \mathbf{1}_V\colon & V & \to & V \\ & v & \mapsto & v \end{array}$$

which we draw as just an arrow labeled by V.



This is a good notation in that the identity map is the identity for the operation of composition of maps, and attaching an arrow to another arrow does not change the diagram. Note that, if we have $f\colon V \to V'$, then $f\mathbf{1}_V = f = \mathbf{1}_{V'} f$. Diagrammatically,



1.3 Tensor Products (a crash course)

If V, W are (finite-dimensional) vector spaces, $V \otimes W$ is a vector space which can be defined thus: pick bases $\{e_i\} \subset V$ and $\{f_j\} \subset W$ and let $V \otimes W$ be such that $\{e_i \otimes f_j\}$ is a formal basis for $V \otimes W$. We have that $\dim(V \otimes W) = \dim(V)\dim(W)$.

We now define a tensor product of vectors in V and W as a bilinear map $\otimes\colon V \times W \to V \otimes W$, so that if $v = v^i e_i \in V$ and $w = w^j f_j \in W$, their tensor product is $v \otimes w = v^i w^j (e_i \otimes f_j)$. Here we use for the first time Einstein's summation convention, which is that when an index appears twice, once as a subscript and one as a superscript, a summation over the range of the index is understood implicitly; thus, we have

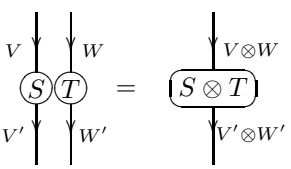
$$v \otimes w = v^i w^j (e_i \otimes f_j) := \sum_{i=1}^{n} \sum_{j=1}^{m} v^i w^j (e_i \otimes f_j).$$

This notation is arguably Einstein's most important contribution to human thought.

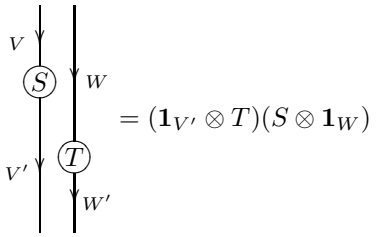
Given linear maps $S\colon V \to V'$ and $T\colon W \to W'$, we can construct another linear map

$$\begin{array}{rccc} S \otimes T\colon & V \otimes V & \to & V' \otimes W' \\ & e_i \otimes f_j & \mapsto & S(e_i) \otimes T(f_j) \end{array}$$

We can draw this as follows:

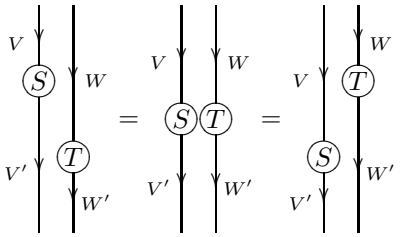


Now consider the following diagram:



$= (\mathbf{1}_{V'} \otimes T)(S \otimes \mathbf{1}_W)$

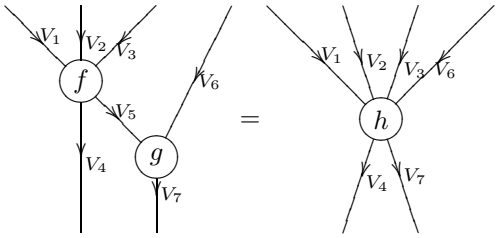
(in the future we will feel free to equate diagrams and formulae). The following "identity" suggests itself:



We will call this operation shifting. It is a special case of the principle that deforming the diagram (there is quite a bit of topology lurking here) does not change the answer. To explain what "deforming the diagram" means, picture the diagram drawn on a framed surface with the endpoints of free lines glued to the frame, and allow any smooth one-to-one deformation of the surface (and hence of the diagram).

Exercise 1 *Prove algebraically that shifting works.*

Finally, consider the following example: given linear maps $f\colon V_1 \otimes V_2 \otimes V_3 \to V_4 \otimes V_5$ and $g\colon V_5 \otimes V_6 \to V_7$, we can combine them is a unique way, which we draw as follows:



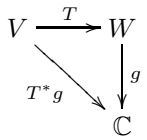
and we obtain a new map $h\colon V_1 \otimes V_2 \otimes V_3 \otimes V_6 \to V_4 \otimes V_7$ by means of a weird combination of tensoring and composition for which there is essentially no good coordinate-free notation other than the diagram![1]

1.4 Duality (I)

Given a vector space V over $\mathbb{C}$ one has the dual vector space defined as

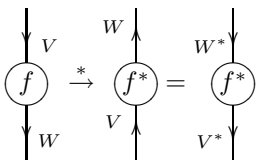
$$V^* = \{\text{linear maps} \quad f\colon V \to \mathbb{C}\}.$$

Moreover, given a linear map $T\colon V \to W$ one can define its adjoint, which is the linear map $T^*\colon W^* \to V^*$ defined by $(T^*g)v = g(Tv)$ for all $g \in W^*$ and $v \in V$:

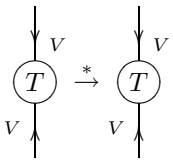


[1] Do we need a section on abstract index notation? In abstract index notation, we have $h^{ij}_{klmn} = f^{ip}_{klm} g^{j}_{pn}$.

There is a nice way to draw adjoints, which is to rotate the diagram by 180°:



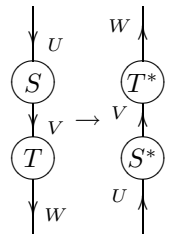
Remember that "time" always flows downwards so, when we introduce duals, arrows pointing downstream represent vector spaces and arrows pointing upstream represent their duals. Note that we do not need to write the asterisk on the label to denote the dual space because the direction of the arrow does this for us automatically. We do need to write the asterisk on the name of the operator because the direction of the arrows may not be sufficient to tell T apart from T^*. Consider an operator $T\colon V \to V^*$. Then the adjoint is $T^*\colon V \to V^*$ and, because an operator need not be self-adjoint, not writing the asterisk would lead to an ambiguous diagram



However, at this point we might decide that "blobs" should not be drawn as circles but as some other shape that is not symmetric, in which case we could drop all asterisks without ambiguities.

The idea to represent adjoints by drawing the diagrams "backwards in time" arose in particle physics, in which taking the adjoint is equivalent to exchanging particles and antiparticles. Richard Feynman was the first to think of antiparticles as "particles going backwards in time", and represented them by reversing the arrows on what we now call Feynman diagrams.

Exercise 2 *Consider the following ambiguous diagram:*



Check that "rotate-then-compose" is the same operation as "compose-then-rotate", therefore showing that the diagrammatic notation is unambiguous. (Hint: translate the diagram into symbols in two ways which will be the left- and right-hand sides of an identity; then prove the identity.)

Chapter 2

Lagrangians for Field Theories (I)

2.1 Framework and Notations

These are the "stars of the show":

- A Lie group denoted by G, which physicists call the "gauge group" and is not to be confused with the "group of gauge transformations". For simplicity, we will assume that the group is a group of matrices like $\mathrm{SO}(n)$, $\mathrm{SU}(n)$, $\mathrm{Sp}(n)$, so it will be a submanifold of the linear space $\mathrm{End}(V)$ for some V. The whole theory can be carried through without assuming that G is a group of matrices.

- The Lie algebra of G, denoted by $\mathfrak{g}$. The names of the lie algebras are obtained from the group names by transmogrifying them into low-case gothic script, for example $\mathfrak{so}(n)$, $\mathfrak{su}(n)$, $\mathfrak{sp}(n)$, etc. A Lie algebra is a vector space, but we will assume that it is a space of $n \times n$ matrices so they can be multiplied, although strictly speaking[1] one is not allowed to do that.

- The trace of an $n \times n$ matrix, denoted Tr. This actually represents two functions: $\mathrm{Tr}\colon G \to \mathbb{C}$ and $\mathrm{Tr}\colon \mathfrak{g} \to \mathbb{C}$. These operations can be defined without reference to matrices, but they are still denoted Tr for convenience.

- An n-dimensional (smooth, paracompact, Hausdorff) manifold representing "space-time" and denoted by M. We will require that M be oriented (to be able to integrate functions) and (usually) compact. M will usually be boundaryless, but sometimes we will consider manifolds with a boundary.

- A principal G-bundle over M, denoted $\pi\colon P \to M$. Since *Gauge Fields, Knots and Gravity* does not cover principal bundles—the principal flaw of that book—we will give a definition of these sometime. (For now, we'll assume you either know it or can fake it.)

- Fields (functions) on M, especially

  – A connection A on the principal bundle P, which physicists call "gauge field" or "vector potential". Locally (or on a trivial G-bundle), A is a $\mathfrak{g}$-valued 1-form, so in a coordinate patch we can write $A = A_\mu \mathrm{d}x^\mu$. The connection is associated to an exterior covariant derivative

  $$\mathrm{d}_A = \mathrm{d} + A = \begin{cases} \mathrm{d} + A\wedge & \text{acting on the fundamental representation of } G \\ \mathrm{d} + [A,\ ] & \text{acting on the adjoint representation of } G. \end{cases}$$

  – Gauge transformations, which are locally G-valued functions on M. The action of a gauge transformation g on the connection is required to satisfy $\mathrm{d}_{A'} g = g\mathrm{d}_A$, which implies

  $$A \mapsto A' = gAg^{-1} + g(\mathrm{d}g^{-1}) = gAg^{-1} - (\mathrm{d}g)g^{-1}.$$

[1] Strictly speaking, when the lie algebra is not an algebra of matrices, one defines a "universal enveloping algebra" with an associative product such that the original Lie bracket equals the commutator of the enveloping algebra.

– The curvature F of the connection A, which physicists call "field strength". Locally, F is a $\mathfrak{g}$-valued 2-form, and it is a function of A:

$$F = \mathrm{d}_A^2 = \begin{cases} \mathrm{d}A + A \wedge A & \text{on the fundamental representation} \\ \mathrm{d}A + \frac{1}{2}[A, A] & \text{on the adjoint representation} \end{cases}$$

When we apply a gauge transformation to A, the curvature changes as $F \mapsto F' = gFg^{-1}$. We therefore say that F is $Ad(P)$-valued.

Exercise 3 *Show that this is the case. [Hints:* $\mathrm{d}(AB) = (\mathrm{d}A)B + A(\mathrm{d}B)$ *for matrix-valued functions; when d jumps over an n-form it picks up a factor of* $(-1)^n$*; and* $gg^{-1} = 1$*. Alternatively, show that* $F = \mathrm{d}_A^2$ *and use* $\mathrm{d}'_A = g\mathrm{d}_A g^{-1}$*.]*

Now, to do field theory we need to concoct Lagrangians from these ingredients (the connection and other fields at hand). Mathematically, a Lagrangian $\mathcal{L}$ is just a scalar-valued n-form which is a function of the fields. It has to be an n-form so that it can be integrated over the manifold M to get a number.

If $\mathcal{L}$ is a Lagrangian and M is a manifold, the integral

$$S = \int_M \mathcal{L}$$

is called "the action of the field configuration". For "nice" theories, the action should be invariant under gauge transformations. One way to do this is to require that the Lagrangian itself be invariant under gauge transformations.

2.2 Gauge-invariant Lagrangians for Gauge Theories (I)

So far the only field we have is the connection A, so let us see if we can build any gauge-invariant n-forms from it.

The most simple-minded n-form we can obtain from A is simply A itself which, as a $\mathfrak{g}$-valued 1-form, would work on 1-dimensional manifolds. To obtain a scalar 1-form, we take the trace. Unfortunately, $\mathcal{L} = \mathrm{Tr}A$ is not gauge-invariant, as

$$\mathrm{Tr}A' = \mathrm{Tr}(gAg^{-1} + gdg^{-1}) = \mathrm{Tr}(gAg^{-1}) + \mathrm{Tr}(gdg^{-1}) = \mathrm{Tr}(A) + \mathrm{Tr}(gdg^{-1}) = \mathrm{Tr}(A) - d\log\det g,$$

and the last term vanishes only if $\det g$ is constant, which will only be the case if G is a special group, in which case $\mathfrak{g}$ is an algebra of traceless matrices and $\mathrm{Tr}A$ is zero in the first place. We have used the cyclic property of the trace, $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, which will prove very useful in the following.

2.2.1 The First Chern Theory

In two dimensions we can do $\mathcal{L} = \mathrm{Tr}F$, which turns out to be gauge-invariant, again by the cyclic property of the trace:

$$\mathrm{Tr}F' = \mathrm{Tr}(gFg^{-1}) = \mathrm{Tr}(F).$$

Two-dimensional field theories are interesting, among other reasons, because in string theory the fundamental objects are the two-dimensional world sheets of one-dimensional strings, and all dynamical variables, including the coordinates of space-time, are fields defined on this world sheet.

Mathematicians call the integral

$$S = \int_M \mathrm{Tr}F$$

the first Chern class, so a natural name for a theory with this action would be "the first Chern theory". It turns out, for example, that when $G = \mathrm{SO}(2)$ the first Chern theory is two-dimensional general relativity!

The nth Chern class is

$$\int_M \mathrm{Tr}(\underbrace{F \wedge \cdots \wedge F}_{n \text{ times}}),$$

which, when taken as an action, gives rise to a perfectly sensible gauge-invariant theory on $2n$-dimensional manifolds which we may well call the nth Chern theory.

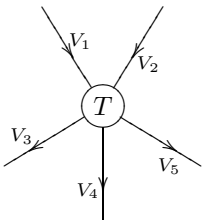
The special case of the second Chern theory is interesting because it works in four dimensions and that's what we think our space-time is! This is not general relativity, though, for any choice of the gauge group G. Sometimes the second Chern theory with $G = \mathrm{SO}(3,1)$ is called "topological gravity". It is similar to general relativity but much simpler—a bit like GR's baby brother.

In fact, as we shall see, there is a way to obtain general relativity from an Lagrangian of the form $e \wedge e \wedge F$, where e is an additional 1-form (variously called "cotetrad", "Vierbein" or "soldering form") independent of the connection A.

Chapter 3

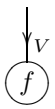
Diagrammatic Methods for Linear Algebra (II)

The basic building blocks in the theory we are developing are linear mappings between tensor products of vector spaces. For example:



3.1 Degenerate Cases

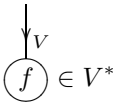
What does this diagram stand for?



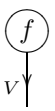
The key to answering this question is to give a meaning to "the tensor product of no spaces". The only possibility is that it is the base field, $\mathbb{C}$ in this case, and that makes sense because the base field is the identity for the operation "tensor product of vector spaces"! That is, we have the following canonical linear isomorphism:

$$\begin{array}{ccc} \mathbb{C} \otimes V & \to & V \\ 1 \otimes v & \mapsto & v \end{array}$$

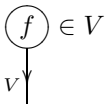
Therefore, we may write



Now, what is this?



It is a linear mapping $f\colon \mathbb{C} \to V$, which can be characterised by giving $f(1) = v \in V$, so we conclude that



Despite the fact that our formalism involves only vector spaces and linear maps, we now have a way to talk about individual elements of vector spaces.

Finally, we should have no problem interpreting this:

This is a linear mapping from $\mathbb{C}$ to $\mathbb{C}$, clearly a complex number! This last example will prove especially useful in the sequel, since it means that a diagram with no arrows sticking into or out of it represents a complex number.

3.2 Duality (II)

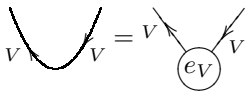
Given a vector space V we can construct its dual V^*, and it comes along with two bilinear maps: the unit and the counit.

3.2.1 The Counit

This is a map from $V^* \otimes V$ to $\mathbb{C}$ defined in the obvious way:

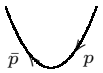
$$\begin{array}{rccc} e_V\colon & V^* \otimes V & \to & \mathbb{C} \\ & f \otimes v & \mapsto & f(v) \end{array}$$

This mapping is called "evaluation" or "dual pairing" and we draw it as

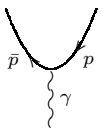


We do not draw a blob in it because the dual pairing is canonical, just like we did not draw a blob for the identity map on V. Another name for this operation is "cup", for obvious reasons. To understand the name "counit" we'll have to wait until the "unit" has been introduced.

The cup operation was introduced into physics by Feynman, in connection with his idea that antiparticles could be interpreted as particles moving backwards in time. In this language, the cup diagram is interpreted as the annihilation of a particle/anti-particle pair.



The reader may be wondering what ever happened to the photons that are produced by such an annihilation event. The answer is that in quantum field theory particles are described by representations of a group including the Poincaré group and hence time-translation, But time-translation symmetry implies energy conservation, so a particle/anti-particle pair cannot annihilate into nothing.



With the tools at our disposal we still cannot talk about time-translation invariance and so we are allowed to draw diagrams such as the cup with nothing coming out at the bottom.

3.2.2 The Unit

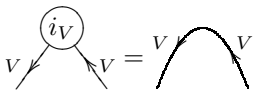
The unit is a map $i_V\colon \mathbb{C} \to V \otimes V^*$. Now, $V \otimes V^* \equiv \{\text{linear } T\colon V \to V\} = \text{End}(V)$, since to any $v \otimes f \in V \otimes V^*$ we can assign a unique endomorphism

$$\begin{array}{rccc} T_{v\otimes f}\colon & V & \to & V \\ & w & \mapsto & vf(w) \end{array}$$

Now, the simplest endomorphism is the identity, and so we define

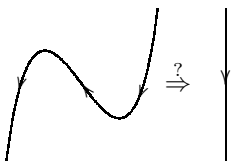
$$\begin{array}{rccc} i_V\colon & \mathbb{C} & \to & V\otimes V^* \\ & 1 & \mapsto & \mathbf{1}_V \end{array}$$

The unit is drawn as

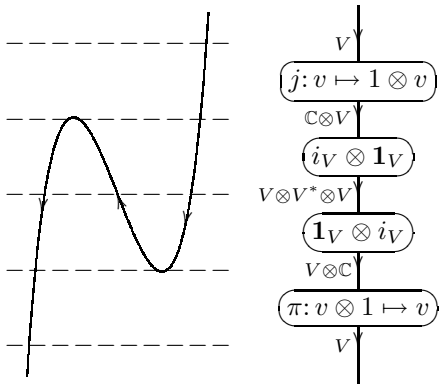


and is called "cup". Note that this is not the dual of the counit, as they are obtained from one another by reflection and not by rotation.

Now that we have the unit and counit we can combine them in different ways. For example, we have

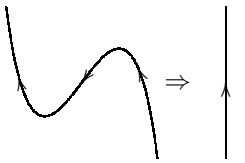


which we would like to equate to the identity on V by "straightening out" the diagram. Let us prove that we can actually do that:



We want to prove that $\pi(\mathbf{1}_V\otimes e_V)(i_V\otimes\mathbf{1}_V)j=\mathbf{1}_V$. Pick $v\in V$. Then $j(v)=1\otimes v\in\mathbb{C}\otimes V$. Now, $i_V(1)$ is the identity endomorphism on V, which we can write as $e_i\otimes f^i$, where $f^i\in V^*$ is defined by $f^i(e_j)=\delta^i_j$. Then $(i_V\otimes\mathbf{1}_V)(1\otimes V)=e_i\otimes f^i\otimes v$. Now, $e_V(f^i\otimes v)=f^i(v)$, so $(\mathbf{1}_V\otimes e_V)(e_i\otimes f^i\otimes v)=e_i\otimes v^i$, where $v^i=f^i(v)\in\mathbb{C}$ are the components of V with respect to the basis $\{e_i\}$, as we now show: $f^j(e_iv^i)=f^j(e_i)v^i=\delta^j_iv^i=v^j$. Finally, $\pi(e_i\otimes v^i)=e_iv^i=v$ and we are done.

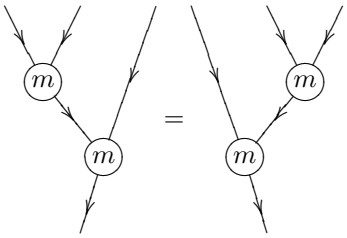
Exercise 4 *Rigorously straighten out the following diagram:*



3.3 Matrix Algebra

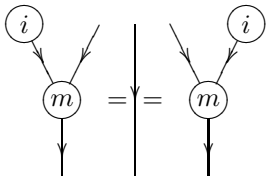
An (associative) algebra is a (complex) vector space A with a bilinear "multiplication" $m\colon A\otimes A\to A$ and a unit element $i\colon\mathbb{C}\to A$ satisfying

- the associative law: $m\big(m(a,b),c\big) = m\big(a,m(b,c)\big)$, drawn as



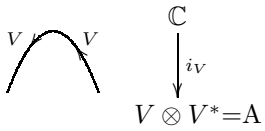
and

- the left and right unit laws, $m(a,i) = a = m(i,a)$, which we draw as

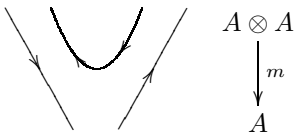


For the associative law to be true, we must extend the notion of equivalence of diagrams to include the new operation "sliding one multiplication blob over another". The unit laws can be interpreted as "unit and multiplication blobs cancel each other out". Now, in the case of matrix algebra we do not need to introduce these new rules, but they follow from the ones we know already.

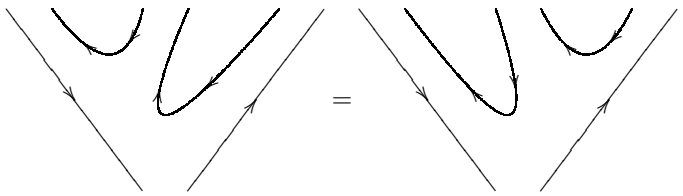
Given a vector space V, we can get an algebra by considering $\mathrm{End}(V)$ with the operation of composition of maps. Now, we have $\mathrm{End}(V) = V \otimes V^*$, so the unit is a certain element of $V \otimes V^*$



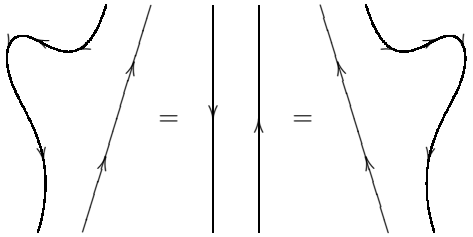
and matrix multiplication is based on the counit:



The associative property can be drawn as



where equality follows by "shifting". Similarly, the unit laws can be drawn as



which is true because we can "straighten out the bends".

Chapter 4

Lagrangians for Field Theories (II)

4.1 Gauge-invariant Lagrangians for Gauge Theories (II)

So far we have seen that, starting from just A on a $2n$-dimensional manifold, we can obtain the gauge-invariant Lagrangian $\mathrm{Tr}(F^{\wedge n})$. This seems to be the end of the story unless we endow the manifold with additional structure.

4.1.1 Yang–Mills Theory

If M is equipped with a metric (i.e. a symmetric, non-degenerate 2-form[1]), we can construct a Hodge $*$ operator which turns p-forms into $(n-p)$-forms, so $*F$ is an $Ad(P)$-valued $(n-2)$-form—locally a $\mathfrak{g}$-valued $(n-2)$-form. Then, $F \wedge *F = *F \wedge F$ is an n-form whose trace is gauge-invariant!

$$\mathrm{Tr}(F' \wedge *F') = \mathrm{Tr}(gFg^{-1} \wedge *gFg^{-1}) = \mathrm{Tr}(gFg^{-1} \wedge g * Fg^{-1}) = \mathrm{Tr}(gF \wedge *Fg^{-1}) = \mathrm{Tr}(F \wedge *F),$$

where $\wedge$ and $*$ act on the differential form part only and g acts on the $Ad(P)$ part only, so g can jump over the other operators. This is the so-called Yang–Mills Lagrangian, and as far as we know it describes all known forces except gravity. (Note: because both $*$ and the measure change sign under a change of orientation, the Yang–Mills action is independent of the orientation of the manifold.) If we want to construct a theory of gravity we don't want the metric to be fixed *a priori*, but the Yang–Mills Lagrangian hints at a way to obtain new gauge-invariant lagrangians by adding more fields.

4.1.2 EF Theory

The properties of $*F$ that we needed in order to prove that the Yang–Mills Lagrangian is gauge-invariant are that: 1) it is an $(n-2)$-form so $\mathcal{L}$ can be integrated over an n-dimensional manifold; and 2) it is $Ad(P)$ valued, so it has the proper behaviour under gauge transformations. Accordingly, we will assume that the manifold M is equipped not only with a connection, but also with a new field E which is an $Ad(P)$-valued $(n-2)$-form. Now we can form the lagrangian $\mathrm{Tr}(E \wedge F)$, which gives rise to the ill-named "BF theory". Originally, E was "wrongly" named B because it was supposed to be analogous to the magnetic field, when in reality it is closer to the electric field, as we will see later on. We will use the name "EF theory" throughout.

EF theory leads to general relativity in $3D$ if we choose $G = \mathrm{SO}(3)$ (Riemannian gravity) or $G = \mathrm{SO}(2,1)$ (Lorentzian gravity). Now we observe that, in $3D$, E is a 1-form, so $\mathrm{Tr}(E \wedge E \wedge E)$ is also a valid Lagrangian.

[1] Do we need to mention that the signature need not be Euclidean?

Let us now list all the EF Lagrangians in various dimensions:

$2D$	$\mathrm{Tr}(F)$	$\mathrm{Tr}(E \wedge F)$	$\mathrm{Tr}(E \wedge E \wedge F)$	$\mathrm{Tr}(E \wedge E \wedge E \wedge F)$	$\ldots$
$3D$		$\mathrm{Tr}(E \wedge F)$		$\mathrm{Tr}(E \wedge E \wedge E)$	
$4D$	$\mathrm{Tr}(F \wedge F)$	$\mathrm{Tr}(E \wedge F)$	$\mathrm{Tr}(E \wedge E)$		
$(2m-1)D$		$\mathrm{Tr}(E \wedge F)$			
$(2m)D$	$\mathrm{Tr}(\wedge^m F)$	$\mathrm{Tr}(E \wedge F)$			

E is an $(n-2)$-form in nD; $m \geq 3$.

In all these cases $\wedge$ is taken to mean "product in $Ad(P)$" and "wedging of the differential forms"[2]. We observe that in $2D$ there is an infinite collection of linearly independent EF Lagrangians, since E is, in this case, a 0-form (an $Ad(P)$-valued function) and can be wedged with F arbitrarily many times to give a 2-form. In $3D$ and $4D$ there are terms depending only on E. These terms are called "cosmological" because $3D$ general relativity with cosmological constant λ follows from the Lagrangian $\mathrm{Tr}(E \wedge F) + \lambda \mathrm{Tr}(E \wedge E \wedge E)$. In $5D$ or higher dimension there cannot be a cosmological term.

Incidentally, the cosmological constant λ of $3D$ general relativity is closely related to the "q parameter" appearing in the theory of quantum groups, so maybe "quantum" is also an inappropriate adjective and these groups should be called "cosmological groups" instead! I don't expect this terminology to catch on, but we'll see it makes sense.

4.1.3 $4D$ General Relativity

$4D$ EF theory is not equivalent to general relativity for any choice of the gauge group, so for this purpose we need something more complicated. The Lagrangian for $4D$ general relativity follows from the following trick, which works for all the $\mathrm{SO}(p,q)$ groups[3].

For simplicity, let's consider $\mathrm{SO}(n)$ and let $V = \mathbb{R}^n$ with its usual inner product. The lie algebra $\mathfrak{so}(V)$ consists of the linear transformations of V that are skew-adjoint with respect to the inner product. This is a Lie subalgebra of $\mathrm{End}(V) = V \otimes V^*$—i.e., it is a subspace that is closed under commutators (Lie brackets). Now, the inner product on V provides a canonical isomorphism between V and V^*, so we can consider $\mathfrak{so}(V)$ imbedded in $V \otimes V$ as $\Lambda^2 V$ (skew-symmetric 2-tensors, or "bivectors").

The trick is to use this identification in the opposite direction. Suppose we have a V-valued 1-form e on spacetime. Then we can define $e \wedge e$ to be a $\Lambda^2 V$-valued 2-form in the following way: if $e = e_i dx^i$, where $\{e_i\}$ are vectors in V, $e \wedge f = (e_i \wedge f_j) dx^i \wedge dx^j$, which is symmetric in e, f so that $e \wedge e = (e_i \wedge e_j) dx^i \wedge dx^j$ does not necessarily vanish. By the above identification, we can reinterpret this as an $\mathfrak{so}(V)$-valued 2-form. This is very much like the E field in 4-dimensional EF theory with gauge group $\mathrm{SO}(n)$!

This allows us to write down an EF-like Lagrangian for $4D$ Riemannian general relativity in terms of two basic fields:

- an $\mathrm{SO}(4)$ connection A, and
- the *cotetrad* field e, which is locally an $\mathbb{R}^4$-valued 1-form.

The Lagrangian looks like this: $\mathrm{Tr}(e \wedge e \wedge F)$, where Tr is the "trace" on $Ad(P)$

[2] This is a place where the abstract index notation for internal indices resolves ambiguities

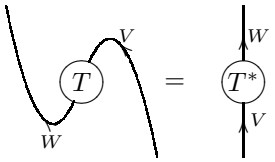
[3] In fact, with a slight modification it works for all symplectic and unitary groups, too.

Chapter 5

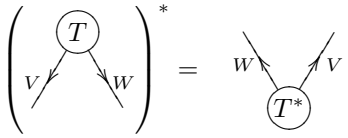
Diagrammatic Methods for Linear Algebra (III)

We are now ready to answer the question why dualisation is represented by a 180° rotation rather than a reflection:

Exercise 5 *Given $T\colon V \to W$, show that*



We note in passing that with our notation it becomes obvious that the dual of a tensor product is the tensor product of the duals in the opposite order: $(V \otimes W)^* = W^* \otimes V^*$.



5.1 Braidings

There is an isomorphism

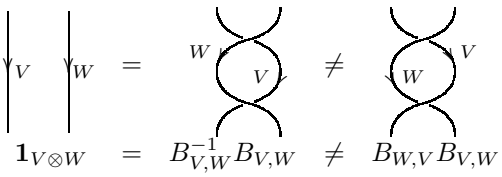
$$\begin{array}{rccc} B_{V,W}\colon & V \otimes W & \to & W \otimes V \\ & v \otimes w & \mapsto & w \otimes v \end{array}$$

called "braiding" which we draw as

The braiding has an inverse

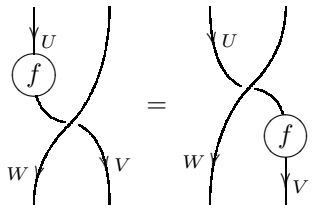
$$\begin{array}{rccc} B^{-1}_{V,W}\colon & W \otimes V & \to & V \otimes W \\ & w \otimes v & \mapsto & v \otimes w \end{array} =$$

Note that, diagrammatically, $B^{-1}_{V,W} \neq B_{W,V}$ because

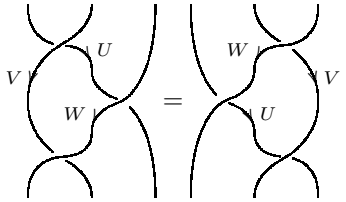


The first equality is called "the second Reidemeister move" in knot theory. This is an example of how to become famous by being the first mathematician to state a trivality.

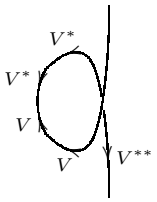
Now, for any linear operator $f: U \to V$ we have $B_{V,W}(f \otimes 1_W) = (1_W \otimes f)B_{U,W}$. This is drawn



Replacing f by $B_{U,V}: U \otimes V \to V \otimes U$, we obtain the identity known as "the third Reidemeister move", which relates two different ways to go from $U \otimes V \otimes W$ to $W \otimes V \otimes U$ by repeated braiding:

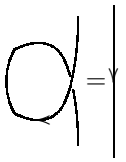


Finally, the first Reidemeister move applies to just one space, and we introduce it last because it involves the unit and counit:



If V is reflexive (as is the case for the finite-dimensional spaces we are considering) this represents the canonical isomorphism between V and V^{**}.

Exercise 6 *Prove the first Reidemeister move:*

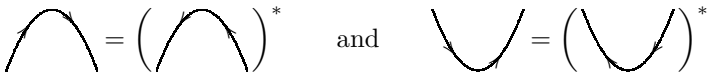


The non-trivial thing that Reidemeister did was to prove the following theorem:

Theorem 1 *Given two two-dimensional projections of the same knot, one can be obtained from the other by a composition of one-parameter diffeomorphisms of the plane and the three Reidemeister moves.*

As an aside, we have defined the cap and cup operations

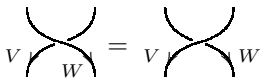
but there are two other operations which can be drawn as a cap and a cup. They can be defined from the usual ones by duality



or by braiding

Exercise 7 *Check that the two definitions coincide.*

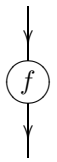
In linear algebra, the "four-dimensional" fact that $B_{V,W}^{-1} = B_{W,V}$, drawn as



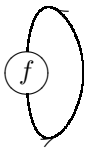
is true because both map $w \otimes v$ to $v \otimes w$. This means the category of vector spaces is a symmetric monoidal category, and we say symmetry is "four-dimensional" because only in more than three dimensions is the above diagram true.

This means that vector spaces and linear maps are nicely suited to (the generally rather boring) four-dimensional knot theory. To have an interesting knot-theoretical application of our formalism we would have to modify the identity above; in other words, we would have to modify the braiding so that it is not symmetrical. This was realised in the 1980's, when representations of quantum groups were used to obtain invariants of knots.

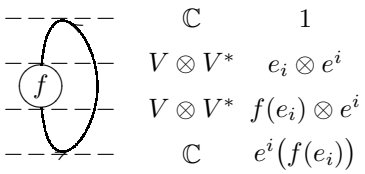
Now, consider an endomorphism of V



and stick a unit (cap) at the top and a "pseudocup" at the bottom:



This, having no free lines, is obviously a complex number that we can obtain from T in a canonical way. The only likely candidate for this is the trace of T, which we now check is true.



And to end, a puzzle. What is this?

If we interpret this as the trace of the identity operator, we conclude that it is nothing other than the dimension of V!

Chapter 6

Physics from Lagrangians (I)

6.1 Lagrangians in Particle Mechanics

To illustrate the Lagrangian formulation of mechanics we will use the simplest case of all, that of a particle moving in one dimension. The trajectory of the particle will be a function $q\colon \mathbb{R} \to \mathbb{R}$ where the domain represents time and the codomain represents one-dimensional space. A Lagrangian is then a function $L(q, \dot{q}; t)$. In most applications the Lagrangian can be split into two parts, $L = T(\dot{q}) - U(q; t)$, where T is called kinetic energy and U potential energy. For one-dimensional particle motion $L(q, \dot{q}; t) = \frac{m}{2}\dot{q}^2 - V(q)$ for some $V\colon \mathbb{R} \to \mathbb{R}$.

Given a Lagrangian, the action for the trajectory is the functional

$$S[q] = \int dt\, L(q, \dot{q}; t).$$

To ensure that the action is finite we uaually restrict time to a closed interval, so $q\colon [t_0, t_1] \to \mathbb{R}$ and

$$S[q] = \int_{t_0}^{t_1} dt\, L(q, \dot{q}; t).$$

Now, the physical trajectory going from an initial position q_1 at $t = t_1$ to the final position q_2 at $t = t_2$ is somewhat misteriously determined by the condition that it be a stationary point of the action S given the endpoints.

To formalise this, we consider "variations of the trajectory" of the form $\tilde{q} = q + \delta q$, where $\delta q\colon [t_1.t_2] \to \mathbb{R}$ is the "variational field" and is required to vanish at the endpoints. We let $q_\epsilon = q_0 + \epsilon\delta q$. The condition that q_0 be a stationary point of S is that

$$\delta S[q_0, \delta q] \colon = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} S[q_\epsilon] = 0$$

for all variational fields δq.

The variation of the action is

$$\begin{aligned}
\delta S[q_0, \delta q]\colon &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_{t_0}^{t_1} dt\, \left[\frac{m}{2}\dot{q}_\epsilon^2 - V(q_\epsilon)\right] = \int_{t_0}^{t_1} dt\, \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \left[\frac{m}{2}\dot{q}_\epsilon^2 - V(q_\epsilon)\right] = \\
&= \int_{t_0}^{t_1} dt \left[m\dot{q}_0\delta\dot{q} - \left.\frac{dV}{dq}\right|_{q_0} \delta q\right] = \underbrace{m q_0 \delta q|_{t_0}^{t_1}}_{=0} - \int_{t_0}^{t_1} dt \left[m\ddot{q}_0 + \left.\frac{dV}{dq}\right|_{q_0}\right]\delta q.
\end{aligned}$$

For this to be zero for all δq we must have

$$m\ddot{q}_0 + V'(q_0) = 0,$$

which is Newton's law with force $F = -V'$.

6.2 Lagrangians in Field Theory (I)

Let us now apply these ideas to the field theory Lagrangians we introduced previously.

6.2.1 The First Chern Theory

Let M be a $2D$ orientable manifold. From the lagrangian $\mathcal{L} = \mathrm{Tr}F$, define the action

$$S = \int_M \mathrm{Tr}F.$$

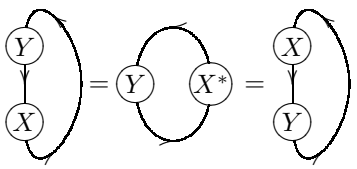
Then we do $A \to \tilde{A} = A + \epsilon\delta A$, where δA, being the difference between two connections, is an $Ad(P)$-valued 1-form. Now, from $F = dA + \frac{1}{2}[A, A]$ we get

$$\delta S[A, \delta A] = \int_M \mathrm{Tr}(\delta F) = \int_M \mathrm{Tr}\Big\{ d(\delta A) + \frac{1}{2}\big([\delta A, A] + [A, \delta A]\big)\Big\}.$$

Now, Tr and d commute because the first acts on the $Ad(P)$-valued part and the second on the 2-form part. Also, $[A, \delta A] = [\delta A, A]$ because $[A, \delta A] = [A_i dx^i, \delta A_j dx^j] = [A_i, \delta A_j] dx^i \wedge dx^j$, which is symmetric. Therefore

$$\delta S[A, \delta A] = \int_M \big\{ d\mathrm{Tr}(\delta A) + \underbrace{\mathrm{Tr}[\delta A, A]}_{=0}\big\} = \int_{\partial M} \mathrm{Tr}(\delta A).$$

We have used $\mathrm{Tr}[A, \delta A] = 0$, which follows from the cyclic property of the trace. For matrices, we can prove the cyclic property of the trace as follows:



so $\mathrm{Tr}(XY) = \mathrm{Tr}(YX)$ and therefore $\mathrm{Tr}[X, Y] = 0$. One can see from the diagram that "cyclic" is a fitting name for this property.

In conclusion, for the first Chern theory the stationary action condition is

$$0 = \delta S[A, \delta A] = \int_{\partial M} \mathrm{Tr}(\delta A),$$

which does not depend on A at all! This means that either all conections are admissible solutions or none is. One way to ensure that there are stationary solutions is to impose the condition that $\delta A|_{\partial M} = 0$, which is analogous to the variational field $\dot{q}$ vanishing at the endpoints of the trajectory. This condition holds trivially if M does not have a boundary.

We anticipate once more that $2D$ general relativity is an instance of the first Chern theory, so that the Einstein equations in 2D are vacuous. 2DGR is therefore rather boring, at least until one quantizes it.

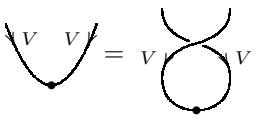
Chapter 7

Diagrammatic Methods for Linear Algebra(IV)

7.1 Orthogonal and Simplectic Vector Spaces

Definition 1 *An* **orthogonal vector space***q* (V, g) *is a vector space* V *equipped with a metric* $g\colon V \otimes V \to \mathbb{C}$ *which is symmetric and non-degenerate.*

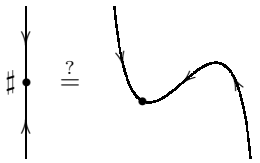
We draw symmetry as



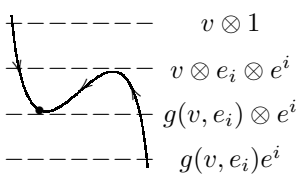
Now the question is, how does one draw the requirement that g be non-degenerate? To answer this question we observe that a metric induces a mapping called "sharp" because in traditional index notation it raises indices much like ♯ "raises" musical notes.

$$\begin{array}{rccc} \sharp\colon & V & \to & V^* \\ & v & \mapsto & g(v, \cdot) \end{array}$$

From g we can construct essentially one mapping from V to V^*, by attaching a "cap" to one of the ends of g, and we would hope that to be a definition of "sharp":

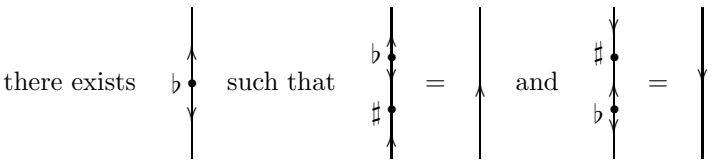


Let $\{e_i\}$ be a basis of V and $\{e^i\}$ its dual basis, such that $e^i(e_j) = \delta^i_j$. Then, $\sharp(e_i)$ is an element of V^* and $\sharp(e_i)(e_j) = g(e_i, e_j) = g(e_i, e_k)\delta^k_j = g(e_i, e_k)e^k(e_j)$, so $\sharp(e_i) = g(e_i, e_k)e^k$. To prove the equality we take a hapless vector and feed it into the infernal device above:

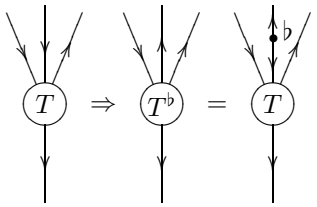


Now, g is non-degenerate if, and only if, ♯ is injective. In the finite-dimensional case, since $\dim V^* = \dim V$, this is equivalent to ♯ being an isomorphism, so it has an inverse $\flat\colon V^* \to V$ called, quite naturally, "flat".

The condition for non-degeneracy is, then,

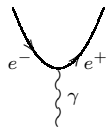


The maps ♯ and ♭ allow us to reverse the direction of arrows at will, for example:



Incidentally, we see that there is an ambiguity in the notations $T^\flat$ and $T^\sharp$ as soon as T has more than one input of the same type.

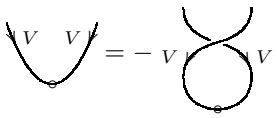
In fact, we can leave out all the arrows on edges labelled by (V, g). This is what is traditionally done in physics with Feynman diagrams such as



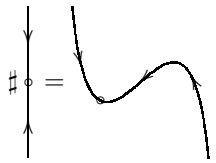
where the photon line does not carry an arrow.

Definition 2 *A* **symplectic vector space** (V, ω) *is a vector space V equipped with a symplectic structure $\omega: V \otimes V \to \mathbb{C}$ which is antisymmetric (also called skew-symmetric or alternating) and non-degenerate.*

Antisymmetry is drawn



As in the previous case we define the "sharp" operator by $\sharp(e_i) = \omega(e_i, \cdot)$. Note that in this case, unlike the orthogonal case, there is a difference in sign depending on which slot of ω is left empty. We choose:



Non-degeneracy is imposed by requiring that ♯ has an inverse ♭.

Exercise 8 *Suppose V is an orthogonal (resp. symplectic) vector space. Define*

Then calculate

Chapter 8

Physics from Lagrangians (II)

8.1 Calculating Variations

Suppose $P \to M$ is a principal G-bundle and A is a connection on it, and let $f[A]$ be a functional of A.

Definition 3 *The* **variational derivative** *of $f[A]$ at A is the linear functional of δA defined by*

$$\delta f[A, \delta A] := \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} f(A + \epsilon \delta A)$$

for all A and δA.

As an example, consider $F = \mathrm{d}A + \frac{1}{2}[A, A] = \mathrm{d}A + A \wedge A$. For the nth time, let us recall that $A \wedge A$ means "wedging the 1-forms and multiplying the $Ad(P)$ values", while $[A, A]$ means "wedging the 1-forms and bracketing the Lie Algebra values". Therefore, If $C = x \otimes \omega$ is a matrix-valued p-form and $D = y \otimes \eta$ is a matrix-valued q-form, we have $[C, D] = [x, y] \otimes (\omega \wedge \eta) = (xy) \otimes (\omega \wedge \eta) - (yx) \otimes (\omega \wedge \eta) = C \wedge D - (yx) \otimes \big((-1)^{pq} \eta \wedge \omega\big) = C \wedge D - (-1)^{pq} D \wedge C$, so

$$[C, D] = \begin{cases} C \wedge D + D \wedge C & \text{if } C, D \text{ are forms of odd degree,} \\ C \wedge D - D \wedge C & \text{else.} \end{cases}$$

When C and D are not matrix-valued, we can still use

$$[C, D] = (-1)^{pq+1} [D, C]$$

(graded commutativity). If, moreover, E is an r-form, we have the following identity

$$(-1)^{pr} \big[C, [D, E]\big] + (-1)^{qr} \big[E, [C, D]\big] + (-1)^{pq} \big[D, [E, C]\big] = 0$$

(graded Jacobi identity), which follows from graded commutativity in the case of matrix-valued forms

Exercise 9 *Prove the graded Jacobi identity*

$$(-1)^{pr} \big[C, [D, E]\big] + (-1)^{qr} \big[E, [C, D]\big] + (-1)^{pq} \big[D, [E, C]\big] = 0$$

where C, D and E are matrix-valued p-, q- and r-forms respectively. [Hint: prove the equivalent formula $\big[C, [D, E]\big] = \big[[C, D], E\big] + (-1)^{pq} \big[D, [C, E]\big]$.]

Since δ is a derivative, it can be completely characterised by linearity and the properties 1) $\delta(f) = 0$ if f is independent of A; and 2) $\delta(C \wedge D) = (\delta C) \wedge D + C \wedge \delta D$. Also, δ operates pointwise on M so it commutes with the ordinary operations of differential calculus on M. Armed with this knowledge, we can proceed to calculating $\delta F[A, \delta A]$:

$$\delta F = \delta\big(\mathrm{d}A + \frac{1}{2}[A, A]\big) = \mathrm{d}\delta A + \frac{1}{2}\big([\delta A, A] + [A, \delta A]\big) = \mathrm{d}\delta A + [A, \delta A],$$

which equals $\mathrm{d}_A\delta A$ by definition of the exterior covariant derivative. It is an ugly fact of life that, while $\delta F = \mathrm{d}_A\delta A$, $F \neq \mathrm{d}_A A$. However, it is still true that $F = \mathrm{d}_A\mathrm{d}_A$, since the following hold as operator equations:

$$\mathrm{d}_A\mathrm{d}_A = (\mathrm{d} + A\wedge)(\mathrm{d} + A\wedge) = \mathrm{d}^2 + \mathrm{d}(A\wedge) + A\wedge\mathrm{d} + A\wedge A\wedge = (\mathrm{d}A)\wedge + A\wedge A\wedge = F\wedge$$

acting on the fundamental representation, and

$$\mathrm{d}_A\mathrm{d}_A = (\mathrm{d} + [A,\ ])(\mathrm{d} + [A,\ ]) = \mathrm{d}^2 + \mathrm{d}[A,\ ] + [A, \mathrm{d}\ ] + \big[A,[A,\ ]\big] = [\mathrm{d}A,\ ] + \frac{1}{2}\big[[A,A],\ \big] = [F,\ ]$$

acting on the adjoint representation.

8.2 Lagrangians in Field Theory (II)

8.2.1 The Second Chern Theory

With the tools of the previous section it will now be easier to derive the equations of motion for the second Chern theory. Let M be a 4-dimensional manifold and F the curvature of a connection A on a principal G-bundle over M. The action of the second Chern theory is

$$S[A] = \int_M \mathrm{Tr}(F\wedge F).$$

We have

$$\delta S[A,\delta A] = \delta\int_M \mathrm{Tr}(F\wedge F) = \int_M \delta\mathrm{Tr}(F\wedge F) = \int_M \mathrm{Tr}\delta(F\wedge F)$$

since, by the linearity of the trace,

$$\delta\mathrm{Tr}\big(f(A)\big) = \mathrm{Tr}\big(f(A+\delta A)\big) - \mathrm{Tr}\big(f(A)\big) = \mathrm{Tr}\big(f(A+\delta A) - f(A)\big) = \mathrm{Tr}\big(\delta f(A)\big).$$

Now,

$$\delta S[A,\delta A] = \int_M \mathrm{Tr}\delta(F\wedge F) = \int_M \mathrm{Tr}(\delta F\wedge F + F\wedge\delta F) = 2\int_M \mathrm{Tr}(F\wedge\delta F),$$

because F is a 2-form. Now, since $\delta F = d_A\delta A$,

$$\delta S[A,\delta A] = 2\int_M \mathrm{Tr}(F\wedge\mathrm{d}_A\delta A) = 2\int_M \mathrm{Tr}\big(\mathrm{d}_A\big(F\wedge\delta A\big)\big) - 2\int_M \mathrm{Tr}(\mathrm{d}_A F\wedge\delta A),$$

where we have used the graded Leibniz law

$$\mathrm{d}_A(C\wedge D) = \mathrm{d}_A C\wedge D + (-1)^p C\wedge\mathrm{d}_A D.$$

Exercise 10 *Prove the graded Leibniz law*

$$\mathrm{d}_A(C\wedge D) = \mathrm{d}_A C\wedge D + (-1)^p C\wedge\mathrm{d}_A D.$$

for $Ad(P)$-valued forms C and D or order p and q.

Finally, if $C = x\otimes\omega$ is a p-form and $D = y\otimes\eta$ is a q-form,

$$\mathrm{Tr}(C\wedge D) = \mathrm{Tr}(xy)\otimes(\omega\wedge\eta) = \mathrm{Tr}(yx)(-1)^{pq}(\eta\otimes\omega) = (-1)^{pq}\mathrm{Tr}(D\wedge C),$$

(graded cyclic property of the trace) which, together with $[C,D] = -(-1)^{pq}[D,C]$ implies that $\mathrm{Tr}[C,D] = 0$ so that $\mathrm{Tr}(\mathrm{d}_A C) = \mathrm{Tr}(\mathrm{d}C)$ and

$$\begin{aligned}\delta S(A)(\delta A) &= 2\int_M \mathrm{Tr}\big(\mathrm{d}\big(F\wedge\delta A\big)\big) - 2\int_M \mathrm{Tr}(\mathrm{d}_A F\wedge\delta A) = 2\int_M \mathrm{d}\mathrm{Tr}(F\wedge\delta A) - 2\int_M \mathrm{Tr}(\mathrm{d}_A F\wedge\delta A) = \\ &= 2\int_{\partial M} \mathrm{Tr}(F\wedge\delta A) - 2\int_M \mathrm{Tr}(\mathrm{d}_A F\wedge\delta A)\end{aligned}$$

Imposing the usual condition that δA vanishes on ∂M (or if M is compact and has no boundary),

$$\delta S(A) = 0 \Leftrightarrow \mathrm{d}_A F = 0.$$

But this equation of motion is vacuous, since $\mathrm{d}_A F = 0$ is the Bianchi *identity* (i.e. it is satisfied by all connections).

We can give two proofs of the Bianchi identity. The first starts by noting that, as an operator equation, $[F, \mathrm{d}_A\ ] = \mathrm{d}_A \mathrm{d}_A \mathrm{d}_A = \mathrm{d}_A[F,\ ]$ because covariant derivatives are associative. But then

$$[F, \mathrm{d}_A\ ] = \mathrm{d}_A[F,\ ] = [\mathrm{d}_A F,\ ] + [F, \mathrm{d}_A\ ]$$

because F is a 2-form, so $[\mathrm{d}_A F,\ ] = 0$.

The second proof is by direct computation[1]:

$$\begin{aligned} \mathrm{d}_A F &= \mathrm{d}_A\big(\mathrm{d}A + \frac{1}{2}[A,A]\big) = \mathrm{d}\big(\mathrm{d}A + \frac{1}{2}[A,A]\big) + \big[A, \mathrm{d}A + \frac{1}{2}[A,A]\big] = \\ &= \frac{1}{2}\big([\mathrm{d}A, A] - [A, \mathrm{d}A]\big) + [A, \mathrm{d}A] + \frac{1}{2}\big[A,[A,A]\big] = \frac{1}{2}\big([\mathrm{d}A, A] + [A, \mathrm{d}A]\big) = 0 \end{aligned}$$

where $\big[A,[A,A]\big]$ vanishes by the graded Jacobi identity. Analogous proofs can be given using the action of d_A and F on the fundamental representation.

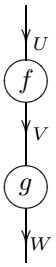
[1] This computation is wrong!

Chapter 9

Diagrammatic Methods for Linear Algebra (V)

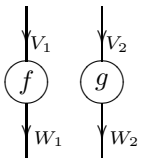
So far we have studied the category of finite-dimensional complex vector spaces with linear maps. Diagrammatically, the operations in the category give rise to structures which have natural diagrammatic representations of various dimensions:

- the basic operation in any **category** is the **composition** of morphisms, which we can draw as a 1-dimensional diagram with nothing to the side:



  Since linear maps can be composed we say that vector spaces with linear maps as morphisms form a category, which we call Vect.

- **products** in a category are represented by 2-dimensional diagrams like



  A category with a product is called **monoidal**, and since tensoring of vector spaces is a product, we say that Vect is a monoidal category.

  – If we have a monoidal category in which every object has morphisms to the tensor identity object, we say the category **has duals** and we can define "cup" and "cap" operators:

  The category Vect is a monoidal category with duals.

- The product in a monoidal category depends on the order of the factors, but if there may be a **braiding**, that is, an isomorphism between $V \otimes W$ and $W \otimes V$

  which is drawn as the 2-dimensional projection of a 3 dimensional structure. In this case the category is called **braided**. Since tensor products of vector spaces come equipped with a braiding, we say that Vect is a braided monoidal category with duals.

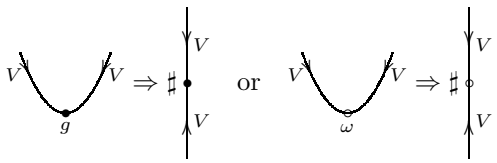
- a braided monoidal category is called **symmetric** if

  This diagram represents a Reidemeister move valid in more than three dimensions, so we can say that the diagram is 4-dimensional. Since this is the case for vector spaces, the category Vect is a symmetric monoidal category with duals.

It is generally true for any category that at the highest dimension we have a property (i.e. an identity that the structures of lower dimension have to satisfy) rather than a new structure. In the case of Vect we do not have a new structure in four dimensions but a property satisfied by the three-dimensional braiding.

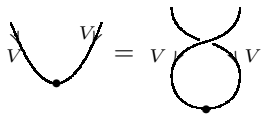
As promised at the beginning, one starts to see geometry arising from just algebraic diagrams. The fact that the category on which quantum mechanics is based is four-dimensional like space-time (at least macroscopically) may be just a coincidence, but we may also take it as an indication that quantum gravity will turn out to be four-dimensional.

We have also seen that some vector spaces (orthogonal and symplectic) have a canonical isomorphism from V to V^*

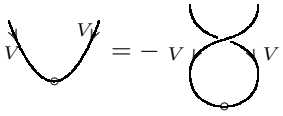


This blurs the distinction between V and V^*. In diagrammatic terms, it allows us to eschew the arrows on the lines representing the vector space V.

A metric has a symmetry property



which is the first Reidemeister move for two-dimensional projections of knots in three dimensions. A symplectic structure, on the other hand, gives rise to diagrams violating the first Reidemeister move:



We can resolve this discrepancy by representing symplectic vector spaces by ribbons rather than by strings; technically, we replace knots by framed knots. We have [1]

(diagram)

[1] Obtain the diagrams!

for ribbons imbedded in any dimension for which the diagram makes sense.

This would be a natural thing to do if we want to use Vect to represent spin-1/2 particles in quantum mechanics (e.g. electrons) since rotating one such particle by 360° results in multiplying the state by -1. We have

(diagram: twist)

In four dimensions this implies

(diagram: -1)

PROOF

To describe physical particles of spin 1/2 we want to use symplectic vector spaces. A vector space with a non-degenerate symplectic form must be even dimensional, so the simplest nontrivial complex vector space admitting a symplectic structure is $\mathbb{C}^2$. The symplectic structure $\alpha\colon \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}$ is determined by the quantity $\alpha(e_i, e_j) = \alpha_{ij} = -\alpha_{ji}$, so $\alpha = \alpha_{12}\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$.

We want to find now an operator $\beta\colon \mathbb{C} \to \mathbb{C}^2 \otimes \mathbb{C}^2$ such that the the following identities are satisfied:

1. antisymmetry

(diagram)

2.

(bubble diagram $= -2$)

3. "binor identity", due to Penrose and named by him by analogy with the term "spinor".

(diagram)

We know that any solution to 1) is of the form $\alpha = A\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. This will be a symplectic structure if and only if $A \neq 0$. Applying the binor identity to the β "cap", we obtain

(diagram)

so β is antisymmetric and it is of the form $\beta = B\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Finally, the bubble diagram evaluates to

$$-2 = (\text{bubble}) = 2AB,$$

so $AB = -1$ and α is non-degenerate and therefore a symplectic structure. As a consistency check, we can see that the binor identity is satisfied for any choice of $A \neq 0$.

To sum up, given a symplectic structure α on $\mathbb{C}^2$ we have a cap $\beta = \alpha^T$ such that the above equations are satisfied, and any solution to those equations is of this form.

Chapter 10

Physics from Lagrangians(III)

10.1 Lagrangians in Field Theory(III)

So far all our equations of motion have turned out to be vacuous, but this is not entirely bad news: we have seen that given a principal G-bundle $P \to M$ with a connection A on a $2n$-dimensional manifold, we can define the gauge-invariant action

$$S[A] = \int_M \mathrm{Tr}(F^{\wedge n}).$$

Then, $\delta S[A, \delta A] = 0$ for all A and δA, although we have only proved it for $n \geq 2$.

Exercise 11 *Derive the equations of motion for the n-th Chern theory and show that they reduce to the Bianchi identity for all $n \geq 1$.*

This is not entirely bad news, as it means that $S[A]$ depends only on the bundle and not on the connection. The numbers $S[A]$ are invariants called Chern numbers which can be used to classify principal G-bundles.

For example, if $G = U(1)$ and M is two-dimensional manifolds the first Chern number classifies principal bundles in the sense that, if $P \to M$ and $P' \to M$ have the same first Chern number, then they are isomorphic. If $G = SU(2)$ and M is four-dimensional, the second Chern number, which is related to the "θ angle" of the standard model of particle physics, classifies the principal G-bundles.

10.1.1 Maxwell Theory

Before taking up nonabelian Yang–Mills theory, we consider the abelian case (Maxwell's theory).

Assume that $G = U(1)$ and the manifold M is equipped with a metric. Since G is abelian, the adjoint representation coincides with the trivial (scalar) representation, so $Ad(P)$-valued objects are invariant. Therefore, $d_A = d$ on the adjoint representation; one can also argue that the term $[A, A]$ vanishes because the Lie algebra of $G = U(1)$ is $\mathfrak{g} \cong i\mathbb{R}$. On the fundamental representation we still have $d_A = d + A\wedge$, and the curvature 2-form is $F = dA$ because A is an ordinary 1-form and $A \wedge A = 0$. Also, in the presence of a metric we have a Hodge $*$ operator mapping p-forms to $(n-p)$-forms in such a way that, for any p-form F and $(n-p)$-form G,

$$F \wedge G = \langle *F, G \rangle \,\mathrm{vol},$$

where vol is the metric-induced volume n-form and $\langle\ ,\ \rangle$ acts on p-forms by

$$\langle \alpha_1 \wedge \cdots \wedge \alpha_p, \beta_1 \wedge \cdots \wedge \beta_p \rangle_p = \det \langle \alpha_i, \beta_j \rangle_1 .$$

So, $*F$ is an $(n-2)$-form and we can form the action

$$S[A] = \int_M \mathrm{Tr}(F \wedge *F) = \int_M F \wedge *F,$$

which describes a universe made purely of light. The variation of the action is

$$\begin{aligned}\delta S[A,\delta A] &= \delta\int_M F\wedge *F = \int_M \delta(F\wedge *F) = \int_M (\delta F)\wedge *F + F\wedge \delta *F = \int_M (\delta F)\wedge *F + F\wedge *\delta F \\ &= \int_M (\delta F)\wedge *F + F\wedge *\delta F\end{aligned}$$

The Hodge operator and wedge product have the following symmetry property[1]: if F, G are p-forms,

$$F\wedge *G = \langle *F, *G\rangle \,\mathrm{vol} = \langle *G, *F\rangle \,\mathrm{vol} = G\wedge *F.$$

Therefore,

$$\begin{aligned}\delta S[A,\delta A] &= \int_M (\delta F)\wedge *F + F\wedge *\delta F = 2\int_M (\delta F)\wedge *F = 2\int_M (d_A\delta A)\wedge *F = 2\int_M (d\delta A)\wedge *F = \\ &= 2\int_M d(\delta A\wedge *F) + \delta A\wedge d*F = 2\int_{\partial M}\delta A\wedge *F + 2\int_M \delta A\wedge d*F.\end{aligned}$$

As usual, we can force the first integral to vanish by requiring δA to vanish on ∂M, and

$$0 = \delta S[A,\delta A] = 2\int_M \delta A\wedge d*F$$

implies that the (vacuum) Maxwell's equations

$$d*F = 0$$

hold.

To see what this has to do with Maxwell's equations, assume that $M = \mathbb{R}\times S$, where S is an $(n-1)$-dimensional manifold (space). Then $F = dt\wedge E + B$, where E (resp. B) is a 1-form (resp. 2-form) on S. Now, if G is a p-form on S,

$$-\langle dt\wedge *_S G, dt\wedge *_S G\rangle \,\mathrm{vol} = \langle *_S G, *_S G\rangle_S \, dt\wedge \mathrm{vol}_S = dt\wedge G\wedge *_S G = (-1)^p G\wedge dt\wedge *_S G$$

Equating the first and last terms, we have

$$*G = (-1)^{p+1} dt\wedge *_S G.$$

Similarly, equating the middle terms one obtains

$$*(dt\wedge G) = *_S G.$$

Therefore, $*F = -dt\wedge *_S B + *_S E$, and $**F = -(dt\wedge *_S *_S E + *_S *_S B) = -F$ since $*_S *_S = \mathbf{1}$. Then, $0 = d*F = dt\wedge d_S *_S B + dt\wedge \partial_t *_S E + d_S *_S E$ so

$$*_S d_S *_S B + \partial_t E = 0 \qquad \text{and} \qquad *_S\, d_S *_S E = 0.$$

[1] This solves the exercise posed in the lecture: show that this does not depend on the signature of the metric

Chapter 11

Building Space from Spin (I)

We will now start constructing vector spaces and linear operators using only the symplectic space $(V = \mathbb{C}^2, \omega)$. It doesn't matter which symplectic structure we use on $\mathbb{C}^2$, since all are isomorphic and give rise to the same diagrammatic calculus, but in case we need an explicit choice we will take $\omega(e_1 \otimes e_2) = 1$ [Note that Carter, Flath and Saito use $i\omega$ as their symplectic structure].

Exercise 12 *Prove that a linear operator $g\colon \mathbb{C}^2 \to \mathbb{C}^2$ preserves ω if, and only if,* $\det g = 1$.

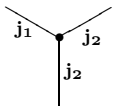
The group of 2×2 complex matrices with unit determinant is denoted $\mathrm{SL}(2, \mathbb{C})$. One of the great facts about this group is that it is the universal covering space of $\mathrm{SO}_0(3, 1)$ (the symmetry group of special relativity) and so, in a sense, it is the natural gauge group of general relativity. The group $\mathrm{SL}(2, \mathbb{C})$ is a double cover of $\mathrm{SO}_0(3, 1)$, and this means that, although a 360° rotation is the identity in $\mathrm{SO}_0(3, 1)$, it differs from the identity in $\mathrm{SL}(2, \mathbb{C})$. This is just the kind of behaviour observed in relativistic spin-$\frac{1}{7}2$ particles, so $\mathrm{SL}(2, \mathbb{C})$ is the symmetry group of relativity and $\mathbb{C}^2$ the state space of a spin-$\frac{1}{7}2$ particle. Since all the constructions that follow use only $(\mathbb{C}^2, \omega)$, all objects will have a built-in $\mathrm{SL}(2, \mathbb{C})$ symmetry.

11.1 Overview

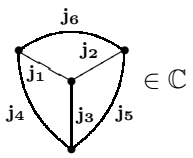
We will build a collection of self-dual vector spaces labelled by half-integers $\mathbf{j}$ which correspond to the $(2j)$th symmetric tensor powers of $V = \mathbb{C}^2$.

$$\left| \mathbf{j} \simeq \right| S^{2j}V \simeq \left| \mathbb{C}^{2j+1} \right.$$

Then, there will be intertwining operators for each triple of such spaces



and combining them we can interpret a tetrahedron with labelled edges, called "$6j$-symbol" in physics, as a complex number



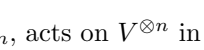
$\in \mathbb{C}$

which we can use to associate a complex amplitude to a triangulated 3-manifold with edges labelled by half-integer (quantised) lengths. This is the basis of some formulations of three- and four-dimensional quantum gravity.

11.2 Symmetrised Tensor Products

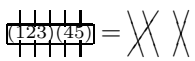
There is essentially only one way in which we can construct new vector spaces starting from a given vector space V, and that is by tensoring it with itself (a symplectic vector space is canonically isomorphic to its dual and we have not introduced direct sums so far). We denote the n-th tensor power of V by

$$V^{\otimes n} = V \otimes \cdots \otimes V$$



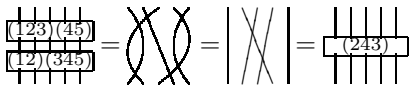
The symmetric group on n letters, S_n, acts on $V^{\otimes n}$ in the following way:

$$\sigma(v_1 \otimes \cdots \otimes v_n) = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(n)}, \qquad \text{e.g.,}$$



and

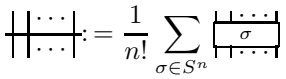
$$\tau\sigma(v_1 \otimes \cdots \otimes v_n) = v_{\sigma^{-1}\tau^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}\tau^{-1}(n)}, \qquad \text{e.g.,}$$



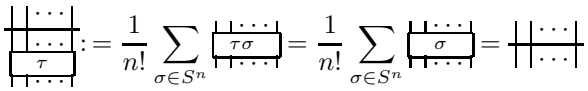
Sitting inside $V^{\otimes n}$ we have the symmetric tensors

$$S^n V = \{x \in V^{\otimes n} \mid \sigma x = x \quad \forall \sigma \in S_n\}.$$

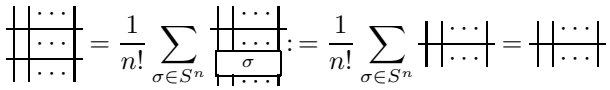
The projector on $S^n V$ is the operator $P\colon V^{\otimes n} \to V^{\otimes n}$ defined by



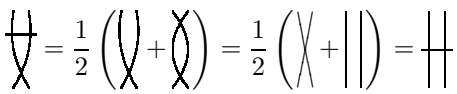
We can easily check that it projects onto $S^n V$



and that it is idempotent:

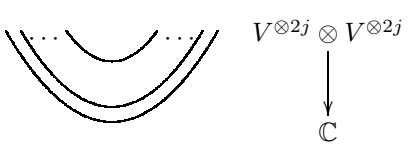


The crucial step, that $\{\tau\sigma\colon \sigma \in S_n\} = S_n$, can be proved diagrammatically. Let's do it for $n = 2$:

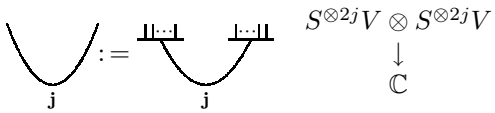


In the case of $V = \mathbb{C}^2$, we denote $S^n V$ by $\mathbf{\frac{n}{2}}$. We represent the projector $P\colon V^{\otimes 2j} \to S^{2j} V$ by

Now, the map

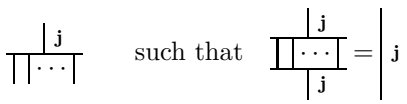


restricts to[1]

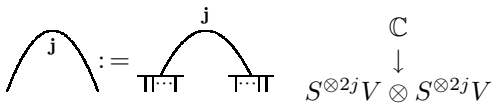


[1] Is this diagram correct?

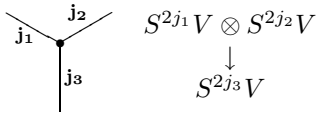
Similarly, there is a canonical injection $J\colon \mathbf{j} \to S^{2j}V \subset V^{2j}$



which we can use to define



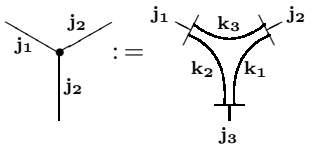
We end this section with a puzzle: can one define a natural mapping



(called an "intertwiner") and if so, how?

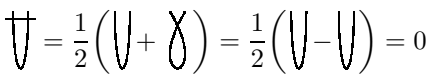
We start by observing that, if $i + j = k$, we can define operators $P\colon \mathbf{i} \otimes \mathbf{j} \to \mathbf{k}$ and $J\colon \mathbf{k} \to \mathbf{i} \otimes \mathbf{j}$

We can combine these operators to obtain



where $\begin{cases} 2k_1 = j_2 + j_3 - j_1 \\ 2k_2 = j_3 + j_1 - j_2 \\ 2k_3 = j_1 + j_2 - j_3 \end{cases}$

One might wonder whether it is possible to take two $\frac{\mathbf{1}}{\mathbf{2}}$ lines from the expansion of, say, $\mathbf{j_1}$ and "cap" them. However, that would make the operator defined above zero, because



Therefore, our definition of the intertwiner is unique up to normalisation.

Chapter 12

Physics from Lagrangians (IV)

12.1 Lagrangians in Field Theory (IV)

12.1.1 EF Theory

Recall that EF theory is defined in terms of two independent fields,

- a connection A on a G-bundle P over an n-dimensional manifold M, and
- an $Ad(P)$-valued $(n-2)$-form E.

The EF action is

$$S = \int_M \mathrm{Tr}(E \wedge F),$$

so

$$\begin{aligned}
\delta S &= \int_M \mathrm{Tr}(\delta E \wedge F + E \wedge \delta F) = \int_M \mathrm{Tr}(\delta E \wedge F + E \wedge d_A \delta A) = \\
&= (-1)^{n-2} \int_M \mathrm{Tr}\big(d_A(E \wedge \delta A)\big) + \int_M \mathrm{Tr}\big(\delta E \wedge F - (-1)^{n-2} d_A E \wedge \delta A\big) \\
&= (-1)^n \int_{\partial M} \mathrm{Tr}(E \wedge \delta A) + \int_M \mathrm{Tr}\big(\delta E \wedge F) + (-1)^{n-1} \int_M \mathrm{Tr}(d_A E \wedge \delta A\big)
\end{aligned}$$

If the variation is assumed to vanish on ∂M, the equations of motion are

$$\delta S = 0 \Leftrightarrow \begin{cases} F = 0 \\ d_A E = 0. \end{cases} \quad \text{and}$$

These equations say that A is a flat connection and that E is parallel with respect to A (since A is flat we can actually say that E is "constant").

These equations do not seem so trivial as others we have previously seen, but in fact all flat connections are equivalent up to gauge transformations. Even more is true: all solutions to the EF equations are equivalent[1] modulo "gauge symmetries" of the following two kinds:

1. "gauge transformations"[2]

$$\begin{cases} A \mapsto gAg^{-1} + (dg)g^{-1} \\ E \mapsto gEg^{-1} \end{cases}$$

and

[1] But $E = 0$ is a solution, so are all solutions of the form $E = d_A \eta$?

[2] Beware that, according to the standard usage in physics, the expressions "gauge symmetry" and "gauge transformation" are not synonims.

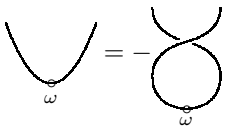
2.

$$\begin{cases} A \mapsto A \\ E \mapsto E + d_A \eta \quad \text{for some } Ad(P)\text{-valued } (n-1)\text{-form } \eta. \end{cases}$$

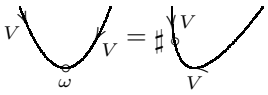
Chapter 13

Building Space from Spin (II)

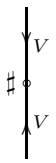
Recall that we have started from just $(V = \mathbb{C}^2, \omega)$,



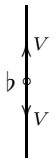
which we draw with no arrows because



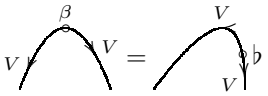
defines a canonical isomorphism:



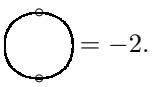
The inverse of ♯ is



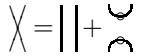
which yields a unique



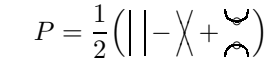
such that



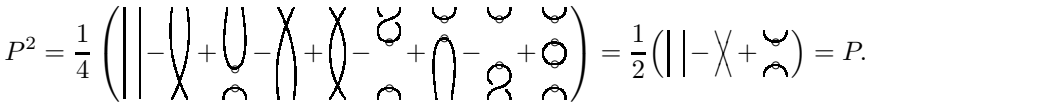
Now, this implies the spin-$\frac{1}{2}$ skein relation



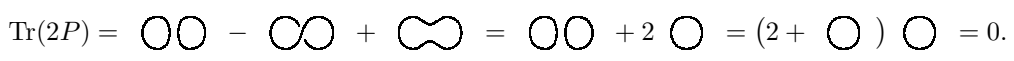
For the proof, we consider



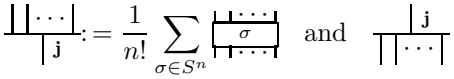
To prove $P = 0$, we first show it is a projector:



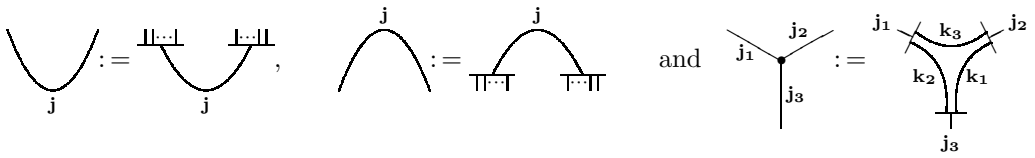
Then, we show that the dimension of its image is zero by taking its trace:



The space $S^{2j}V$ of symmetric tensors is called the "spin-j representation of $\mathrm{SL}(2,\mathbb{C})$" and we denote it by $\mathbf{j}$. We have defined a symmetriser operator on the $2j$-th tensor power of V which projects onto $\mathbf{j}$, and an injection from $\mathbf{j}$ to $V^{\otimes 2j}$:



We define



Recall that the vertex only exists if the following equivalent conditions hold:

$$\left.\begin{array}{l} j_1 = k_2 + k_3 \\ j_2 = k_3 + k_1 \\ j_3 = k_1 + k_2 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{c} 2k_1 = -j_1 + j_2 + j_3 \\ 2k_2 = j_1 - j_2 + j_3 \\ 2k_3 = j_1 + j_2 - j_3, \end{array}\right.$$

where the numbers $\{k_i\}$ are arbitrary nonnegative half-integers, which is equivalent to the conditions that

1. $\{j_i\}$ are the sides of a triangle, and
2. $j_1 + j_2 + j_3 \in \mathbb{N}$.

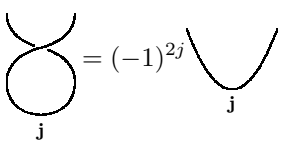
A triple of half-integers $\{j_i\}$ satisfying these conditions is called "admissible". Finally, our diagrammatic calculus allows us to uniquely define vertices with legs sticking out in any direction.

13.1 Spin Networks

With these ingredients we can construct spin networks, which are trivalent graphs with edges labelled by half-integers satisfying the admissibility conditions at every vertex. For example:

(diagram: spin network)

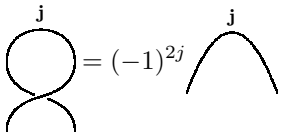
The j representation satisfies the following identities, analogous to those for the 1/2 (fundamental) representation:



so that the cup is an orthogonal structure for integer j and a symplectic structure for half-integer j.

(diagram: proof)

Similarly, we have



$= (-1)^{2j}$

and the diagrammatic calculus lets us show that

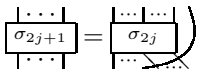
(diagram: cap–cup) = (diagram: identity) = (diagram: cup-cap)

Therefore, we have

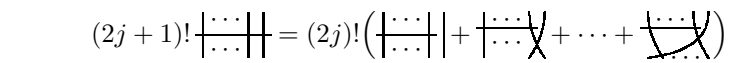
$= (-1)^{2j} \dim S^{2j}V.$

But it is well-known that $S^{2j}V$ is isomorphic to the space of homogeneous polynomials of degree $2j$ in two variables, and there are $2j+1$ of those.

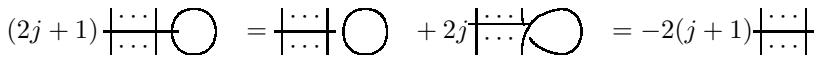
It is possible to calculate $\dim S^{2j}V$ directly using the diagrammatic calculus. First, any permutation can be decomposed in the following form:



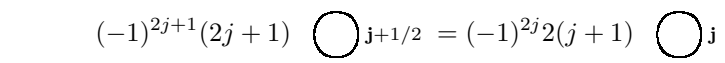
Summing over all permutations,



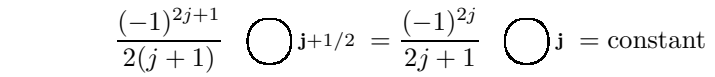
We now divide by $(2j)!$ and take the trace over the last line to obtain



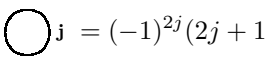
where we have used the fact that the first Reidemeister move produces a sign change. Taking the remaining traces and multiplying by $(-1)^{2j+1}$ we have



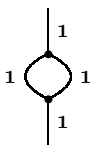
Equivalently,



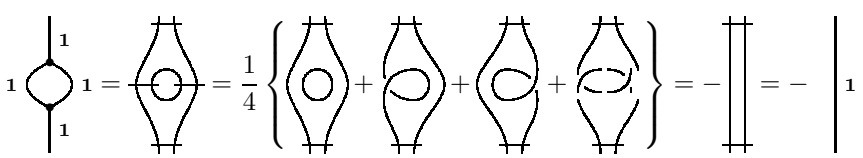
But the case $j = 1/2$ indicates that the constant is 1, so



As a further example, let's consider the operator



Expanding all **1**-lines into two $\frac{\mathbf{1}}{\mathbf{2}}$-lines, we have



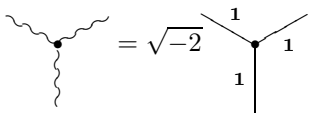
In general, any operator of this form is a multiple of the identity. This is a special case of Schur's lemma, which we will prove, along with some interesting generalisations, in the next lecture.

Chapter 14

Vector networks and the four-colour theorem

Just for fun, let's see what we can do with spin networks where all the edges are labelled by the same spin. Labelling them all with **0** is boring—every such spin network gives an identity operator. We can't label them all with $\frac{\mathbf{1}}{\mathbf{2}}$, since this would violate the admissibility conditions at the vertices. So the simplest interesting case is to label all the edges with **1**. This was studied extensively by Penrose, Kauffman and Bar-Natan.

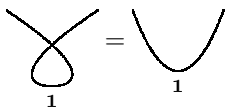
We start by introducing a simplified notation for these spin networks. Since all the edges are labelled by **1**, we leave out the edge labels. Also, it will be convenient to change our normalization of the trivalent vertex. For this purpose, we draw wiggly spin-1 lines, and we define a vertex with three wiggly lines coming out to be $\sqrt{-2}$ times the ordinary spin-1 vertex:



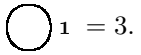
This has the advantage that we can maintain our convention that spin-$\frac{1}{2}$ lines do not carry a label, and on the few occasions when we need to espand a spin-1 line into a pair of symmetrised spin-$\frac{1}{2}$ lines there will be no ambiguity even though no lines are labelled.

Just as spin-$\frac{1}{2}$ particles are called "spinors", in physics spin-1 particles are called "vectors", for reasons soon to become clear[1]. Thus we call a spin network with only wiggly lines a "vector network".

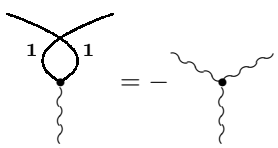
Now let's derive the basic indentities satisfied by vector networks. Since the **1** representation carries an orthogonal structure, we have



so we don't need to think of the edges as ribbons or keep track of twists. Also, a closed loop gives the dimension of the **1** representation:



Furthermore, the vertex is antisymetric[2]:

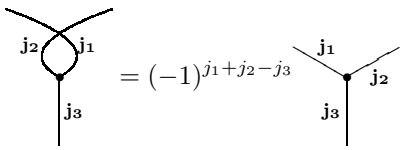


[1]The **1** representation of $\mathrm{SL}(2,\mathbb{C})$ is isomorphic to the fundamental representation of $\mathrm{SO}(3,1)$, just like the **1** representation of $\mathrm{SU}(2)$ is isomorphic to the fundamental representation of $\mathrm{SO}(3)$.

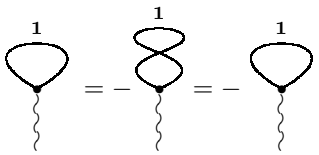
[2]Unfortunately, it is not possible to draw curved wiggly lines. We need to come up with a different representation.

This follows from a more general fact about spin networks, which we leave as an exercise, to be proved straight from the definitions.

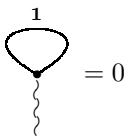
Exercise 13 *Prove that*



The antisymmetry of the vertex together with our ability to remove twists implies that

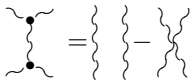


so

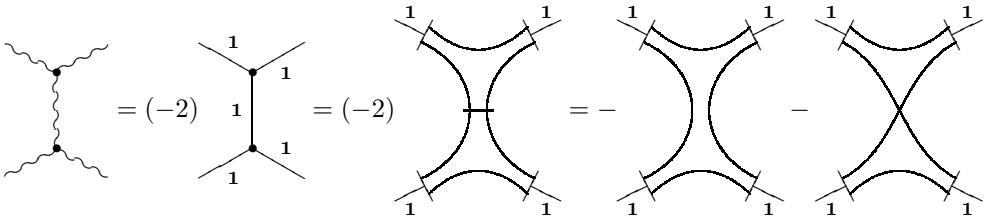


It follows that whenever a vector network has an edge connecting a vertex to itself, this network evaluates to zero.

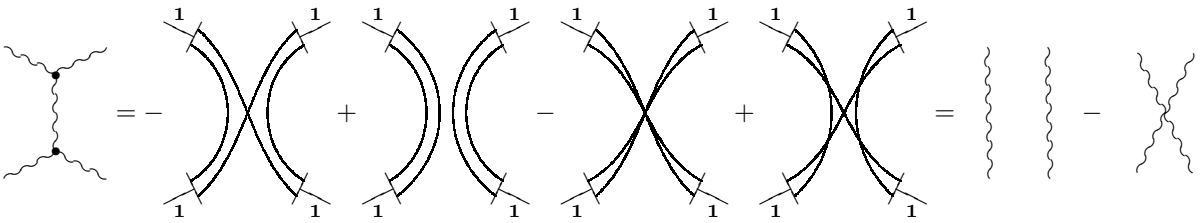
The really interesting identity, however, is this:



We call this the "spin-1 skein relation". It should remind you a bit of the spin-$\frac{1}{2}$ skein relation. It's different, but it was also discovered by Penrose. To prove it, we go back to the definitions:

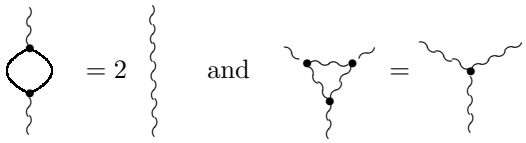


Using the spin-$\frac{1}{2}$ skein relation on the horizontal lines we get



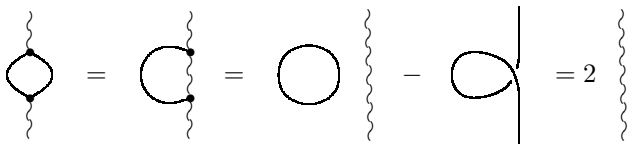
since the symmetrisers allow us to cancel the first and last terms, and the remaining two are just what we need.

The spin-1 skein relation implies the following identities:



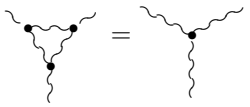
These are both great for simplifying vector networks. How do we prove them? The trick is just to find a place to apply the spin-1 skein relation, and apply it!

To prove the first identity, we just take a trace on the "left" legs of the sipn-1 skein relation. We will be bending things around a bit but the topological invariance of our rules says that's okay.

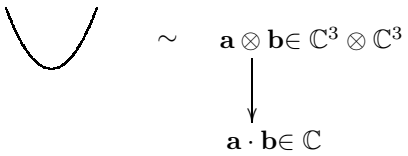


We leave the other one as an exercise.

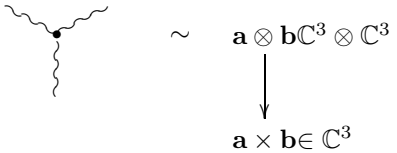
Exercise 14 *Prove the "star-triangle" relation*



We now reveal a secret that we have been hiding all along—while still dropping small clues so you could have caught on if you were really paying attention. What we are actually doing here is good old-fashioned vector algebra: the study of the dot product and cross product! We usually think of these as operations involving vectors in $\mathbb{R}^3$, but we can use the same formulas to define them as operations involving vectors in $\mathbb{C}^3$, which is the spin-1 represetation of $SU(2)$. The dot product corresponds to this vector network:

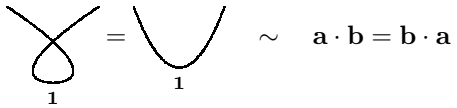


while the cross product corresponds to this one:

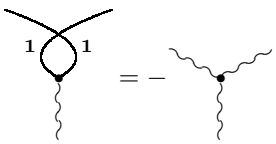


One can show this either by direct computation or a little group representation theory; we leave this as an exercise. By the way, if you were wondering *which* square root of -2 we used in our redefinition of the trivalent vertex, it doesn't really matter as long as we're consistent: the two choices correspond to using either a left-hand rule or right-hand rule for the cross product!

The classical identities of vector algebra are easily proved using diagrammatic methods. We have already seen the commutativity of the dot product and the anticommutativity of the cross product:

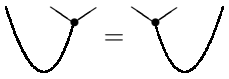


$$\sim \quad \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$$



$$\sim \quad \mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$$

If we take the output of the trivalent vertex and turn it into an input we obtain the triple scalar product, and the following identity follows from the topological invariance of the diagrammatic calculus:



$$\sim \quad \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$$

If we turn one or two of the outputs of the spin-1 skein relation we obtain the remaining two identities:

$$\sim \quad \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$$

$$\sim \quad (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c})$$

In short, spin networks know all about vector algebra!

Now, it's not hard to show that we can evalulate any closed vector network using just the identities we have derived so far. The resulting number is always an *integer*, since our identities only involve integers. This was the reason for changing our normalization of the trivalent vertex. We have also seen that this integer is zero if our vector network has an edge connecting a vertex to itself—or in the language of graph theory, a "loop".

That much is obvious. The following result is less obvious. We say a vector network is **planar** if it can be drawn in the plane without any edges crossing each other. Now, there is a remarkable fact about **1** spin networks, and that is

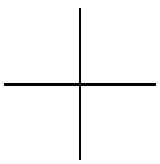
Theorem 2 *If a planar, trivalent graph without loops is interpreted as a spin-*1 *spin network, it evaluates to a nonzero integer.*

This is one of the hardest theorems in all of mathematics, and its only known proofs involve reducing it to thousends of special cases, and then checking these one by one by computer. The reason is that this result is equivalent to the four-colour theorem!

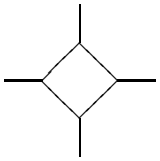
The story of this theorem began one day in October, 1852 when a fellow named Francis Guthrie was coloring a map of England. He wondered whether it was always possible to color any map with only 4 colors, in such a way that no two countries (or counties!) sharing a common stretch of boundary were given the same color. Guthrie's brother passed the question on to the famous mathematician De Morgan, who passed it on to students and other mathematicians. In 1878 Cayley publicized it in the Proceedings of the London Mathematical Society. In just one year the mathematician Kempe proved it. However, in 1890, Heawood found an error in Kempe's proof! And then the real fun began...

We shall not tell the rest of the story, leading up to how Appel and Haken finally proved it in 1976, with the help of a computer calculation involving 10^{10} operations and taking 1200 hours. Instead, let us explain how the four-color theorem is equivalent to the above result about vector networks.

First, note that to prove the 4-color theorem, it suffices to consider the case where only three countries meet at any "corner," since if more meet, say four:

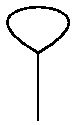


we can stick in a little country at the corner:



so that now only three meet at each corner. If we can color the resulting map, it's easy to check that the same coloring with the little corner countries deleted gives a coloring of the original map.

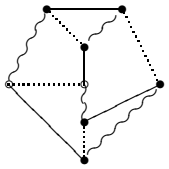
Let us use the language of graph theory, calling the map a "graph," the countries "faces," their borders "edges," and the corners "vertices." If the graph coming from our map never has a loop like this:



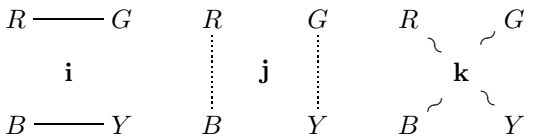
then the region enclosed by the loop and the edge connecting it to the rest of the graph can be removed without affecting whether the graph can be coloured or not.

What we've shown is that, in order to prove the four-color theorem, it suffices to consider planar trivalent graphs without loops.

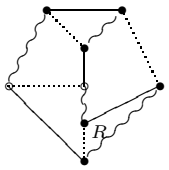
Now, to four-color the faces of a graph like this, it turns out to be enough to three-color the edges. In other words, it's enough to label its edges with the letters i, j and k in such a way that no two edges labelled by the same letter meet at a vertex. For example:



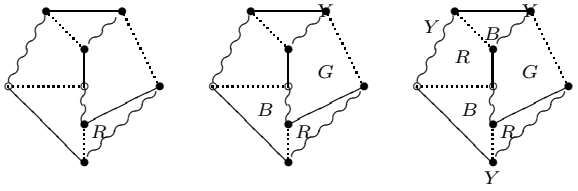
How does this give a four-coloring of the faces? The trick is to use this profound relationship between the numbers 3 and 4: there are 3 ways to divide the set of 4 colors into two pairs:



To get a four-coloring of the regions from a three-coloring of the edges, we start by coloring one face arbitrarily:



Then we color the rest using this rule: whenever we cross an edge, the face color switches to the other color in the pair corresponding to the letter on this edge. In the above example we get this four-coloring:



We leave it as an exercise to check that this rule always gives a well-defined four-coloring. In particular, you should check that as you march around a vertex following this rule, the 3 countries you go through have different colors, and by the time you get back where you started, you get back to the same color. This isn't hard.

Even better, we can run this rule backwards. If we start with a four-coloring of the faces, it uniquely determines a way to label the edges by **i**, **j**, **k** so that no two edges meeting at a vertex get the same letter.

In short, the 4-color theorem is equivalent to this result:

Theorem 3 *For any trivalent planar graph without loops, there exists a way to 3-label it so that no two edges meeting at a vertex carry the same label.*

Next, let us see why *this* theorem is equivalent to the fact that any planar vector network without loops evaluates to a nonzero integer.

Think of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ as the usual basis for $\mathbb{C}^3$, and imagine evaluating a vector network by contracting a lot of tensors, one for each vertex. Since the vertex of a vector network is just the cross product in disguise, the tensor at each vertex is just

$$\epsilon_{abc} = \begin{cases} 1 & \text{if } (a,b,c) \text{ is an even permutation of } (\mathbf{i},\mathbf{j},\mathbf{k}) \\ -1 & \text{if } (a,b,c) \text{ is an even permutation of } (\mathbf{i},\mathbf{j},\mathbf{k}) \\ 0 & \text{otherwise} \end{cases}$$

where (a, b, c) are the labels encountered on the edges ar we march around the vertex. The left-versus-right-hand rule ambiguity implies that it does not matter whether we march around each vertex clockwise or counterclockwise, as long as we do the same at every vertex. The spin network is thus given by a sum over all ways of labelling the edges by i, j, and k, where each term in this sum is a product over all vertices of numbers 1, -1, or 0, computed using the above rule.

If there is *no* way to three-color the edges of our graph, all the terms in this sum will vanish, so our network will evaluate to zero. Thus the only thing left to check is the converse: if the network evaluates to zero, all the terms in the sum must vanish—so there are no ways to three-color its edges. We leave the proof of this as an exercise for the courageous reader.

This means that the four-color theorem is really a theorem about vector networks! We can also reformulate it purely in terms of the vector cross product, as follows:

Theorem 4 *If* $\mathbf{v_1} \times \cdots \times \mathbf{v_n}$ *is parenthesised in two ways, there exists an assignment* $\mathbf{v_l} \in \{\mathbf{i},\mathbf{j},\mathbf{k}\}$ *such that both expressions are equal and nonzero.*

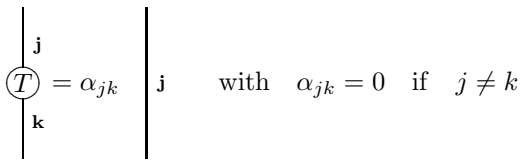
Can we use any of these reformulations to find a shorter proof of the four-color theorem? Nobody knows! People have tried and, so far, failed. We need some new ideas - perhaps some tools from physics. As we shall see, the "profound relationship between the numbers 3 and 4" that we used above is also important in quantum gravity. Maybe this is a clue.

Chapter 15

Schur's Lemma and Related Results

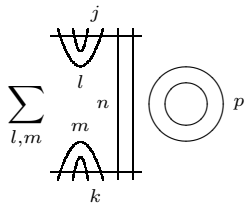
Theorem 5 (Schur's Lemma) *If T is an open spin network with two legs, then it is proportional to the identity.*

We have to show that



where $\textcircled{T}$ represents a spin network.

The proof is remarkably simple. We reexpress all vertices as triangles and expand all symmetrisers—except those at the top and bottom—as sums over permutations. The spin network falls apart into a sum of diagrams consisting of "spaghetti" and "spaghetti-ohs":

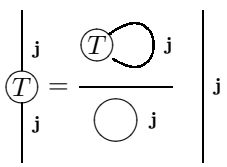


but

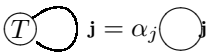
implies that $T = 0$ except possibly if $j = k$.

We will call this proof technique "rewiring" or "rerouting".

Corollary 1 *If T is an open spin network,*

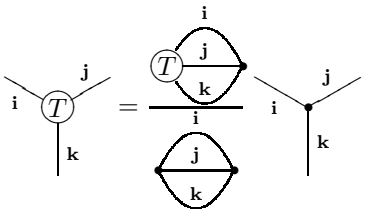


The formula follows by taking the trace of both sides of Schur's Lemma:



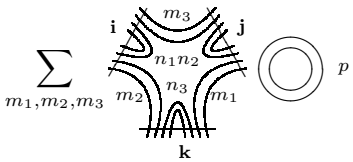
Now, we observe that a closed spin network is just a complex number, so we can divide by it if it is nonzero! Since $\bigcirc\,\mathbf{j}$ is nonzero for nonnegative j, we can solve for α_j and the result follows.

Theorem 6 (Wigner-Eckart) *If T is an open spin network with three legs, it is proportional to the corresponding trivalent vertex.*



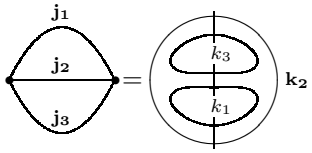
We have to prove two things, that T is proportional to the vertex and that the proportionality factor is as stated. Assuming that proportionality holds, the proportionality factor follows by joining the three free ends to a single vertex. Proportionality to the vertex can be obtained almost trivially by rewiring (in essentially the same way as the uniqueness of the vertex itself).

We expand each leg and each of the **j**-lines inside T into $2j$ spin-1/2 lines with a symmetriser across them. Then, the spin network breaks up into



but, again, $= 0$ implies that all the terms vanish except those proportional to the trivalent vertex.

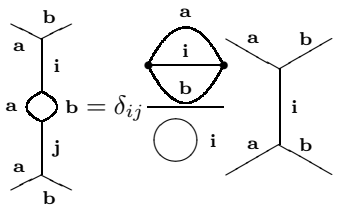
The spin network



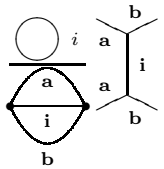
is called "θ-net" or, in physics, "$3j$-coefficient".

We do not have an analogue of Schur's Lemma for spin networks with four legs, but we can come really close.

First, we have

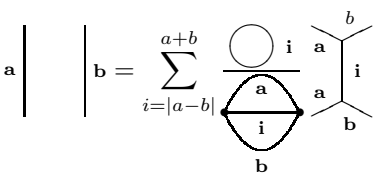


(where we apply Schur's lemma to the intermediate "bubble") which implies that



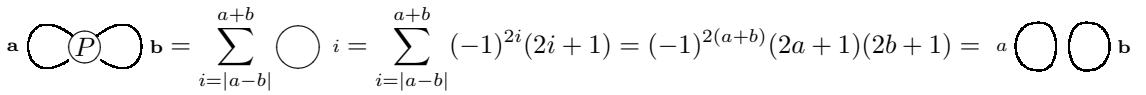
is a projector, and the product of different projectors vanishes. Now,

Theorem 7 (resolution of the identity)

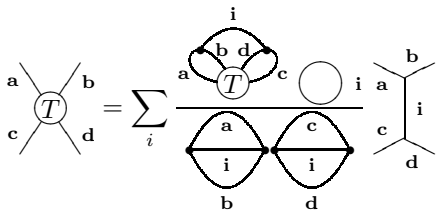


where the sum is extended to the values of i satisfying the compatibility conditions with a and b.

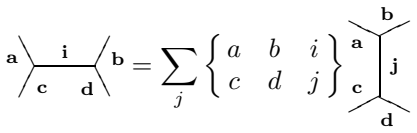
The right-hand side is a projector onto a subspace of $\mathbf{a}\otimes\mathbf{b}$. All we need to do is prove that the dimension of its image is the dimension of $\mathbf{a}\otimes\mathbf{b}$. Calling the left-hand side P, we have



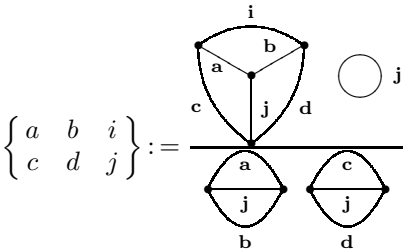
Corollary 2 (tetravalent Schur's lemma) *If T is a spin network with four legs,*



This time the proof follows by applying the resolution of the identity to the top and bottom pairs of legs. An important special case is

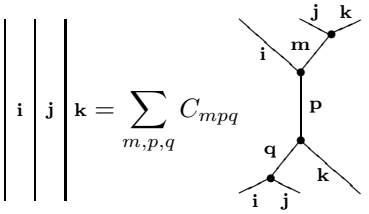


where $\left\{ \begin{smallmatrix} a & b & i \\ c & d & j \end{smallmatrix} \right\}$ are what spectroscopists call "$6j$-symbols" or "Racah $6j$ coefficients". We leave it as an exercise to prove that



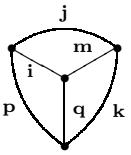
As is usual with these kinds of quantities, there are several definitions in the literature which differ by signs or multiplicative factors. In fact, if it wasn't for historical reasons and consistency with the physics literature, it would certainly make sense to call the tetrahedron diagram "$6j$-symbol".

Lurking in all this is the associative property of the tensor product, $\mathbf{i}\otimes(\mathbf{j}\otimes\mathbf{k}) \simeq (\mathbf{i}\otimes\mathbf{j})\otimes\mathbf{k}$. The linear operator which effects the isomorphism is the "associator", drawn as



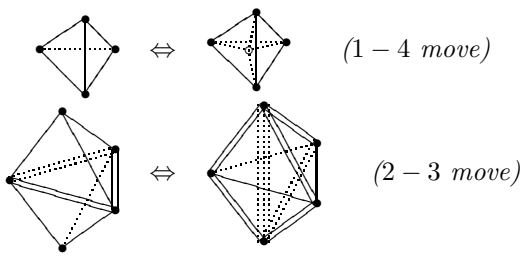
where the C_{mpq} stand for a coefficient too tedious to calculate here and to cumbersome to print, although as an exercise, the reader is encouraged to evaluate it as a function of bubbles, thetas and tetrahedrons. In fact, everything we will do next quarter on $3D$ quantum gravity will be based on the associator!

The amazing thing that happens is that each term on the right-hand-side of the definition of the associator can be closed up to give a tetrahedron diagram



Then it is possible to triangulate a three-dimensional manifold, label all the edges by half-integer numbers and multiply together the complex numbers associated to all tetrahedra. This will turn out to be an invariant of three-manifolds. This last fact follows if we can show that the resulting number is independent of the triangulation.

Theorem 8 *Any two triangulations of the same three-dimensional manifold can be obtained from one another by a sequence of "Pachner moves":*



All we need to do is evaluate the spin networks on both sides and show that they are equal[1].

[1]In fact, the $1-4$ move leads to a divergent sum, and this will be one of the major arguments for the introduction of "quantum (or cosmological)" groups.

Chapter 16

Physics from Lagrangians (V)

16.1 Lagrangians in Field Theory (V)

We've seen the following theories and equations:

- Chern Theories:

$$S = \int_{M^{2n}} \mathrm{Tr}(\wedge^n F) \longrightarrow \delta S = 0 \Leftrightarrow \begin{cases} 0 = 0 & \text{if } n = 1 \\ (F^{\wedge(n-2)}) \wedge d_A F = 0 & \text{if } n > 1, \end{cases}$$

  which hold identically by the Bianchi identity $d_A F = 0$.

- Electromagnetism:

$$S = \int_M \mathrm{Tr}(F \wedge *F) \longrightarrow \delta S = 0 \Leftrightarrow d{*}F = 0.$$

  We will see that, in the non-abelian Yang–Mills case,

$$S = \int_M \mathrm{Tr}(F \wedge *F) \longrightarrow \delta S = 0 \Leftrightarrow d_A{*}F = 0,$$

  which reduces to Maxwell's equations in the abelian case.

- EF theory:

$$S = \int_M \mathrm{Tr}(E \wedge F) \longrightarrow \delta S = 0 \Leftrightarrow \begin{cases} F = 0 \\ d_A E = 0 \end{cases}$$

To finish up this story, we will now derive the equations of motion for the rest of the Lagrangians we wrote at the beginning, and next quarter we will quantise them.

16.1.1 EF Theory with Cosmological Constant

We recall that, in dimensions less than 5, The EF Lagrangian admits additional terms not involving F at all. We will ignore the $2D$ case for the moment, since it involves infinitely many coupling constants.

In $4D$ we have

$$S = \int_M \mathrm{Tr}(E \wedge F + \lambda E \wedge E)$$

The possible Chern term $F \wedge F$ does not affect the classical equations of motion, so we ignore it. On quantisation, or on manifolds with a boundary, this term does have an effect. The variation of the action is

$$\delta S = \int_M \mathrm{Tr}(\delta E \wedge F) + \mathrm{Tr}(E \wedge \delta F) + \lambda \mathrm{Tr}\big[(\delta E \wedge E) + (E \wedge \delta E)\big] = \int_M \mathrm{Tr}(\delta E \wedge F) + \mathrm{Tr}(E \wedge d_A \delta A) + 2\lambda \mathrm{Tr}(\delta E \wedge E),$$

where E and δE are 2-forms so their wedge product is symmetric. We now integrate by parts:

$$\begin{aligned}\delta S &= \int_M \mathrm{Tr}(\delta E \wedge F) + \mathrm{Tr}(E \wedge d_A \delta A) + 2\lambda \mathrm{Tr}(\delta E \wedge E) = \\ &= \int_{\delta M} \mathrm{Tr}(E \wedge \delta A) + \int_M \mathrm{Tr}(\delta E \wedge F) - \mathrm{Tr}(d_A E \wedge F) + 2\lambda \mathrm{Tr}(\delta E \wedge E)\end{aligned}$$

and the equations of motion turn out to be

$$\delta S = 0 \Leftrightarrow \begin{cases} d_A E = 0 \\ F + 2\lambda E = 0 \end{cases}$$

If $\lambda \neq 0$, the second equation just says $E = -\frac{1}{2\lambda} F$, in which case the first equation is Bianchi's identity. The equations of motion turn out to be vacuous again, and this is just like Chern theory in that the space of solutions is the same, namely the space of all connections. The equations of motion also imply that λ is a number, not a function on the manifold.

The bottom line is that as soon as there is a nonzero cosmological constant the character of the theory changes drastically.

Let us now look at $3D$ EF theory with cosmological constant. We have

$$S = \int_M \mathrm{Tr}(E \wedge F + \lambda E \wedge E \wedge E).$$

It is interesting that in lower dimension we can do more things with the E. Each dimension has a distinct personality as far as the available Lagrangians go, and this is related to how topology is very different in different dimensions. Also, things become boring in high-enough dimensions. We have seen reflections of these facts already in track 1.

The variation of the action is

$$\delta S = \int_M \mathrm{Tr}(\delta E \wedge F) + \mathrm{Tr}(E \wedge \delta F) + 3\lambda \mathrm{Tr}(\delta E \wedge E \wedge E) =$$

using the cyclic property of the trace and the fact that cyclic permutations of three 1-forms do not involve a change of sign.

$$\begin{aligned}\delta S &= \int_M \mathrm{Tr}(\delta E \wedge F) + \mathrm{Tr}(E \wedge \delta F) + 3\lambda \mathrm{Tr}(\delta E \wedge E \wedge E) = \\ &= \int_M \mathrm{Tr}(\delta E \wedge F) + \mathrm{Tr}(E \wedge d_A \delta A) + 3\lambda \mathrm{Tr}(\delta E \wedge E \wedge E) = \\ &= -\int_{\delta M} \mathrm{Tr}(E \wedge \delta A) + \int_M \mathrm{Tr}(\delta E \wedge F) + \mathrm{Tr}(d_A E \wedge \delta A) + 3\lambda \mathrm{Tr}(\delta E \wedge E \wedge E),\end{aligned}$$

integrating by parts at the last step. The equations of motion are

$$\delta S = 0 \Leftrightarrow \begin{cases} d_A E = 0 \\ F + 3\lambda E \wedge E = 0 \end{cases}$$

If we differentiate the second equation and use Bianchi's identity,

$$0 = 3\lambda (d_A E \wedge E - E \wedge d_A E).$$

Since E is a 2-form, its wedge products are symmetric, so this is zero independently of whether $d_A E$ vanishes. Therefore, the two equations turn out to be independent.

We can now justify our previous statement that, if $G = \mathrm{SU}(2)$ or $G = \mathrm{SO}(3)$, these are the equations of motion for Riemannian gravity with cosmological constant. E is a frame field, so it gives a metric on the manifold! Then $d_A E$ ends up meaning that A is the Levi-Civita connection of that metric. Lorentzian general relativity can be obtained by making $G = \mathrm{SO}(2,1)$ or $\mathrm{SL}(2,\mathbb{R})$

16.1.2 $4D$ General Relativity

To finish, let's derive the equations of motion for $4D$ general relativity. We have a connection A and an SO(4) frame field e. The lagrangian is $\mathcal{L} = e \wedge e \wedge F$. With a cosmological constant,

$$S = \int_M \mathrm{Tr}(e \wedge e \wedge F + \lambda e \wedge e \wedge e \wedge e)$$

Here we are using the embedding of $\Lambda^2\mathbb{R}^4$ as antisymmetric matrices, which are isomorphic to $\mathfrak{so}(4)$. Therefore, we can pretend that $e \wedge e$ is an $Ad(P)$-valued 2-form. The action of a gauge transformation on $e \wedge e$ is $e' \wedge e' = ge \wedge eg^{-1}$. This is called the Palatini formulation of general relativity.

The variation of the action is

$$\delta S = \int_M 2\mathrm{Tr}(\delta e \wedge e \wedge F) + \mathrm{Tr}(e \wedge e \wedge \delta F) + 4\lambda \mathrm{Tr}(\delta e \wedge e \wedge e \wedge e)$$

because $e \wedge f = f \wedge e$, as $\wedge$ is antisymmetric both on $\mathbb{R}^4$ and on the 1-form part. Substituting $\delta F = d_A \delta A$ and integrating by parts,

$$\delta S = \int_{\partial M} \mathrm{Tr}(e \wedge e \wedge \delta A) + \int_M 2\mathrm{Tr}(\delta e \wedge e \wedge F) - \mathrm{Tr}\big(d_A(e \wedge e) \wedge \delta A\big) + 4\lambda \mathrm{Tr}(\delta e \wedge e \wedge e \wedge e)$$

so the equations of motion are

$$\delta S = 0 \Leftrightarrow \begin{cases} d_A(e \wedge e) = 0 \\ e \wedge F + 2\lambda e \wedge e \wedge e = 0 \end{cases}$$

The first equation characterises A as the Levi-Civita connection of the metric $e \wedge e$, and the second equation is Einstein's equation in the presence of a cosmological constant.

To finish, recall that the equations of motion of $4D$ EF theory are

$$\delta S = 0 \Leftrightarrow \begin{cases} d_A E = 0 \\ F + 2\lambda E = 0, \end{cases}$$

which means that solutions of general relativity are solutions of EF theory. This suggest the following way to generate solutions of GR: We "just" have to find solutions to EF theory in which $E = e \wedge e$.

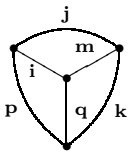
Part II

Quantum gravity/Category Theory

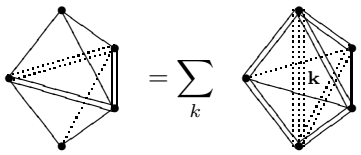
Chapter 17

Quantum Gravity (I)

Last time we did a lot of work to show how to calculate a number of linear operators based on spin networks. In particular, a closed spin network is a linear operator from $\mathbb{C}$ to $\mathbb{C}$ (i.e. a number). We concentrated on

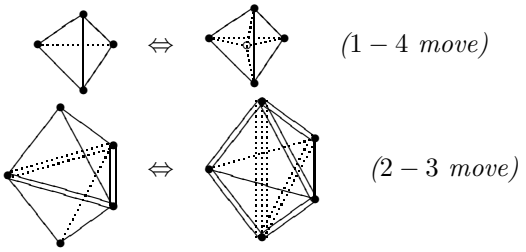


which associates a complex number to each tetrahedron. It turns out that (up to fudge factors),



Now the following theorem suggests an intimate relationship between topology and spin networks:

Theorem 9 (Pachner) *Any two triangulations of a compact* 3*-manifold can be obtained from one another by a sequence of "Pachner moves":*



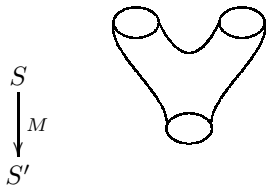
To the wise, this suggests that we should use spin networks to generate *three-dimensional topological quantum field theories*. So, our objective for the next series of lectures is: let's get wise!

17.1 Topological Quantum Field Theories (I)

In the struggle to reconcile quantum mechanics and general relativity, we are helped by an analogy between two otherwise very different subjects.

General relativity is about space and space-time, and space is very flexible object. Space will be any $(n-1)$-dimensional manifold, and space-time will be an n-dimensional manifold limited by two "choices of space". So spacetime is a manifold with boundary, technically a cobordism between two disconnected parts

of its boundary, which are labelled "input" and "output".



In Quantum mechanics, on the other hand, we describe the possible states of a system using a vector space (Hilbert space) and we describe the passage of time by linear operators. In other words, the space of states is a Hilbert space, and processes are described by linear operators.

$$\begin{array}{c} \psi \in \mathcal{H} \\ \Big\downarrow T \\ T(\psi) \in \mathcal{H}' \end{array}$$

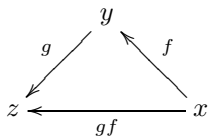
In sum, in general relativity and quantum mechanics we have notions of what things can be, and how things can change to become other things. A topological quantum field theory will be a way to go from general relativity to quantum mechanics, i.e. given a manifold called "space", it will spit out a Hilbert space, and given a spacetime it will spit out a linear operator. Therefore, we are looking for some kind of map between the world of manifolds and cobordisms and the world of Hilbert spaces and linear operators. This was the approach taken by Atiyah in his axiomatisation of topological quantum field theories.

17.1.1 Category Theory (I)

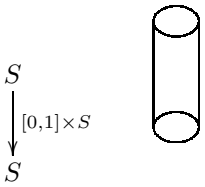
Those in the know will have realised that, in the above exposition, by "world" we mean "category", which we now define.

Definition 4 *A* **Category** $\mathcal{C}$ *consists of*

- *a collection*[1] *of* ***objects****;*
- *given two objects* x, y*, a set* $\mathrm{Hom}(y, x)$ *of* ***morphisms****. Generalizing from the categories where* $\mathrm{Hom}(y, x)$ *is a set of functions, we denote* $f \in \mathrm{Hom}(y, x)$ *by* $f : x \to y$*. Morphisms satisfy the following properties:*
  - *given morphisms* $f : x \to y$ *and* $g : y \to z$*, we can compose them and obtain*[2] $g \circ f : x \to z$*. When there is no possibility of confusion* $g \circ f$ *is abbreviated* gf*.*



  - *for any* x*, there is an* ***identity*** *morphism* $1_x : x \to x$ *such that, for any* $f : x \to y$*, we have* $f1_x = f = 1_y f$*. For example,*



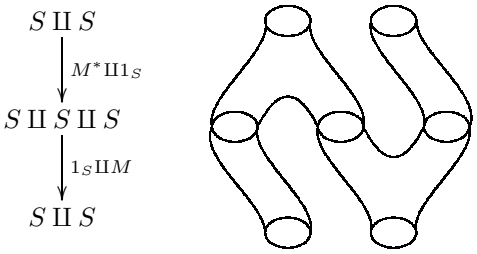
[1] This collections is not in general a set, but a proper class. Consider the category Set, in which the collection of all objects cannot be a set because of the famous Russell paradox.

[2] At this point, category theorists split into warring factions, depending on the order in which they write the composition of morphisms.

This definition captures the most primitive notions of "things" and the "processes" that things can undergo, in other words, the ways that things can "be" and the ways that things can "happen".

Examples of categories are:

- Set, where objects are sets and morphisms are functions.
- nCob,where objects are $(n-1)$-dimensional compact manifolds, and morphisms are n-dimensional cobordisms.



- Vect, where objects are (finite-dimensional, complex) vector spaces, and morphisms are linear operators.
- Hilb, where objects are (finite-dimensional, complex) Hilbert spaces, and morphisms are linear operators.

Quantum mechanics uses Hilb rather than Vect because (among other things)

- given state vectors (i.e. unit vectors) in a Hilbert space, say ϕ and ψ, then $\langle \phi \mid \psi \rangle$ is the **amplitude** and $|\langle \phi \mid \psi \rangle|^2$ is the **probability** that a system prepared in state ψ will be found in state ϕ. There is no such structure in Vect.
- given an operator $T\colon \mathcal{H} \to \mathcal{H}'$, the condition $\langle T^*\phi \mid \psi \rangle = \langle \phi \mid T\psi \rangle$ defines an **adjoint** operator $T^*\colon \mathcal{H}' \to \mathcal{H}$. In Vect, the best we can get is the dual $T^*\colon \mathcal{H}'^* \to \mathcal{H}^*$.
- **observables** in quantum mechanics are represented by self-adjoint operators $A\colon \mathcal{H} \to \mathcal{H}$, where $\mathcal{H}$ is the space of states of the system and $A = A^*$. Such an operator[3] has associated an orthonormal basis $\{\psi_i\}$ of $\mathcal{H}$ such that $A\psi_i = a_i\psi_i$ with $a_i \in \mathbb{R}$. The interpretation is that ψ_i is a state in which A will always be measured to be a_i.

The fact that in Hilb we have a canonical antiisomorphism $\mathcal{H} \to \mathcal{H}^*$ induced by $\langle \cdot \mid \cdot \rangle$ is very different from Vect or Set, but a lot like nCob, where the "dual" of a space is the same space with the opposite orientation, and the "adjoint" of an n-cobordism is its time-reversal. Time reversal is of utmost importance in physics.

[3]More generally, any **normal** operator, i.e. any operator such that $NN^* = N^*N$, has an orthonormal basis of eigenvectors with complex eigenvalues.

Chapter 18

Quantum Mechanics from a Category-Theoretic Viewpoint (I)

We have come short of defining topological quantum field theories because we still haven't explained just what is a map between categories. A topological quantum field theory is, among other things, a *functor* $Z\colon n\mathrm{Cob} \leftarrow \mathrm{Hilb}$. What this means is:

Definition 5 *If $\mathcal{C}$ and $\mathcal{D}$ are categories, a* **functor** $F\colon \mathcal{C} \to \mathcal{D}$ *consists of:*

- *an object $F(x) \in \mathcal{D}$ for each $x \in \mathcal{C}$;*
- *a morphism $F(f)\colon F(x) \to F(y)$ for each $f\colon x \to y$ such that*
  - $F(1_x) = 1_{F(x)}$ *for all $x \in \mathcal{C}$*
  - $F(gf) = F(g)F(f)$ *for all $f\colon x \to y$ and $g\colon y \to z$.*

This looks a lot like a group homomorphism, and that should be no surprise because a group is a special kind of category. In fact, for any object x in a category $\mathcal{C}$, $\mathrm{Hom}(x,x)$ is a monoid and $F\colon \mathrm{Hom}(x,x) \leftarrow \mathrm{Hom}\big(F(x),F(x)\big)$ is a monoid homomorphism.

18.1 Schrödinger's Equation

In ordinary quantum mechanics we don't talk about how the topology of space changes, and also time is a parameter (there is some kind of fixed clock which ticks to "universal time"). So we assume that there is a single Hilbert space, not a whole collection of them. Also, for each time $t \in \mathbb{R}$ we have an operator $U(t)\colon \mathcal{H} \to \mathcal{H}$ which describes time evolution in such a way that, if ψ is the state of the system at time $t = 0$, $\psi(t) = U(t)\psi$ is the state of the system at time t. Time-translation symmetry is expressed by $U(t)U(s) = U(s+t) = U(s)U(t)$, and if $U(t)$ is defined for $t < 0$ we have a group rather than a semigroup homomorphism. Moreover, we require that[1] $\langle \phi(t) \mid \phi(t)\rangle = \langle \phi \mid \phi\rangle = 1$. From this we can see that $U(t)$ must be unitary as follows.

First, observe that $U(0)^2 = U(0+0) = U(0)$ implies that $U(0)$ is a projector. If $\mathcal{H}'$ is the subspace onto which $U(0)$ projects, the image of $U(t) = U(0)U(t)U(0)$ is in $\mathcal{H}'$ for all t, so we can assume without loss of generality that $\mathcal{H} = \mathcal{H}'$ and $U(0) = 1_{\mathcal{H}}$.

Then, $\langle \psi \mid \psi\rangle = \langle U(t)\psi \mid U(t)\psi\rangle = \langle U^*(t)U(t)\psi \mid \psi\rangle$ implies $U^*(t)U(t) = 1_{\mathcal{H}}$ for all t. Since $U(t)U(-t) = U(0) = 1_{\mathcal{H}}$, we conclude that $U^*(t) = U^*(t)U(t)U(-t) = U(-t)$.

If we add the continuity assumption that $\lim_{t\to s} \|U(s)\psi - U(t)\psi\| = 0$ for all ψ, we have that $U(t)$ is a **strongly-continuous, one-parameter unitary group**. Quite a mouthful. We then have

Theorem 10 (Stone) *If $\mathcal{H}$ is a Hilbert space and $U(t)$ is a strongly-continuous, one-parameter unitary group, then $U(t) = \exp\big(-itH\big)$, where H is self-adjoint.*

[1] By the polarisation identities, knowledge of $\langle \phi \mid \phi\rangle$ for all ϕ determines $\langle \phi \mid \psi\rangle$ for all ϕ, ψ

The operator H in Stone's theorem is called the Hamiltonian operator and it corresponds to the energy observable. Stone's theorem and $\psi(t) = U(t)\psi$ imply the abstract Schrödinger equation, $i\frac{\partial}{\partial_t}\psi(t) = H\psi(t)$.

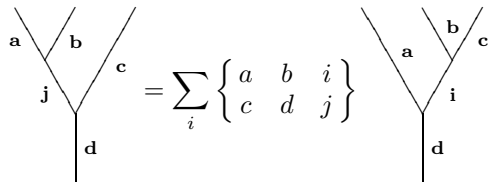
Normally, the way physicists approach a quantum-mechanical problem is, given the Hamiltonian, solve for the evolution of the system. In contrast, in quantum field theory and quantum gravity the hard part is to figure out the Hilbert space and Hamiltonian of the theory.

Chapter 19

Quantum Gravity (II)

We begin with a very important identity which will be use to obtain the $2-3$ Pachner move in $3D$ quantum gravity.

Exercise 15 (Biedenharn–Elliot Identity) *The* $6j$ *symbols are defined by*

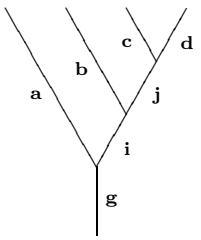


$$= \sum_i \begin{Bmatrix} a & b & i \\ c & d & j \end{Bmatrix}$$

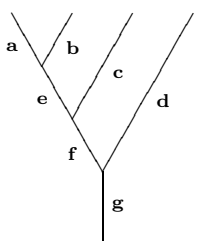
Prove that they satisfy the Biedenharn–Elliot identity:

$$\sum_h \begin{Bmatrix} a & b & h \\ c & f & e \end{Bmatrix} \begin{Bmatrix} a & h & i \\ d & g & f \end{Bmatrix} \begin{Bmatrix} b & c & j \\ d & i & h \end{Bmatrix} = \begin{Bmatrix} e & c & j \\ d & g & f \end{Bmatrix} \begin{Bmatrix} a & b & i \\ j & g & e \end{Bmatrix}$$

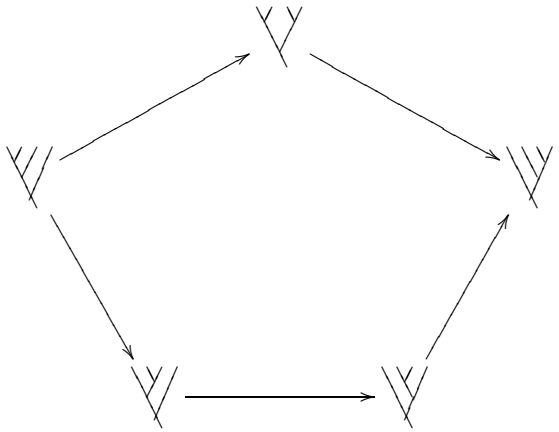
[Hint: Use the fact that



form a basis of spin networks from $\mathbf{a} \otimes \mathbf{b} \otimes \mathbf{c} \otimes \mathbf{d}$ *to* $\mathbf{g}$*, and express*



as a linear combination of the former in two different ways; then equate coefficients.] [Big hint:



Follow the steps in the above diagram.]

19.1 Topological Quantum Field Theories (II)

We have seen that nCob and Hilb have lots of things in common, and topological quantum field theories exploit that fact.

To get it over with, let us just state the definition of a topological quantum field theory and explain all the terms we need for it to make sense. First, nCob and Hilb are symmetric monoidal $*$-categories.

Definition 6 *A* **unitary topological quantum field theory** *is a symmetric, monoidal, $*$-functor*

$$Z\colon n\mathrm{Cob} \leftarrow \mathrm{Hilb}$$

19.1.1 Category Theory (II)

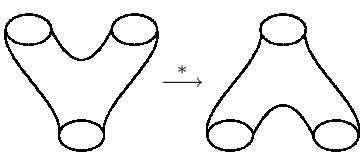
Definition 7 *We say that a category $\mathcal{C}$* **has adjoints** *or* **duals for morphisms** *or* **is a $*$-category** *if there is a contravariant functor $*\colon \mathcal{C} \to \mathcal{C}$ which takes objects to themselves and such that $*^2 = 1$ (the identity functor). For any object x or morphism f, the dual is denoted $*(x) = x^*$ or $*(f) = f^*$.*

Spelling out the definition, $*$ has to satisfy the following properties:

- $x^* = x$ for any $x \in \mathcal{C}$,
- for any $f\colon x \to y$ there is a morphism $f^*\colon y \to x$ (this is what "contravariant" means),
- for any $x \in \mathcal{C}$, $(1_x)^* = 1_{x^*} = 1_x$,
- for any morphisms $f\colon x \to y$ and $g\colon y \to z$, we have $(gf)^* = f^*g^*$, and
- $(f^*)^* = f$ for any morphism f.

Examples of $*$-categories are:

- nCob, where M^* is obtained by exchanging the roles of input and output. If the cobordism is imbedded, this can be represented as reflection along the "time" direction.

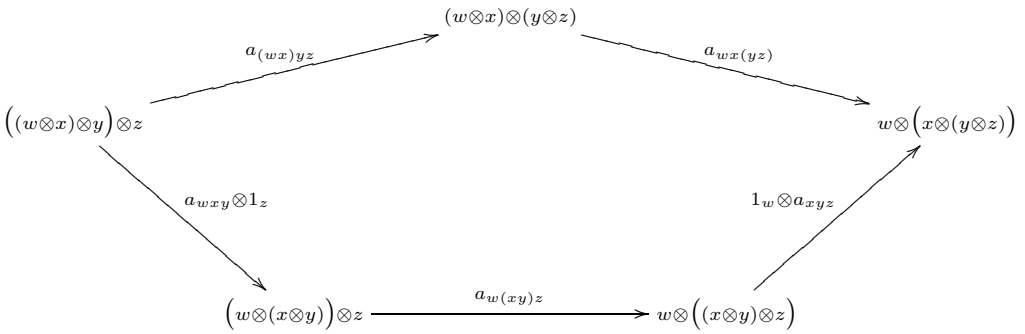


- Hilb, where the adjoint T^* of a linear operator $T\colon \mathcal{H} \to \mathcal{H}'$ is defined by $\langle T^*\phi \mid \psi\rangle_{\mathcal{H}} = \langle \phi \mid T\psi\rangle_{\mathcal{H}'}$.
- any **groupoid** (a category where every morphism is invertible), as then the inverse has the properties required of $*$.

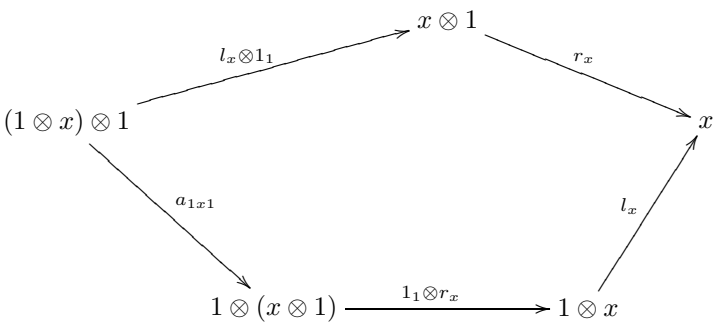
Definition 8 *We say that $F\colon \mathcal{C} \to \mathcal{D}$ is a $*$-functor if, and only if, given $f\colon x \to y$, we have $F(f^*) = F(f)^*\colon F(x) \to F(y)$.*

Definition 9 *A category $\mathcal{C}$ is* **monoidal** *if it is equipped with an operation $\otimes$ with the following properties:*

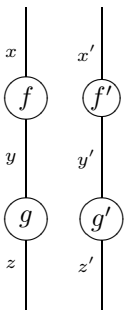
- *for any $x, y \in \mathcal{C}$, there is an object $x \otimes y \in \mathcal{C}$;*
- *for any $f\colon x \to x'$ and $g\colon y \to y'$, there is a morphism $f \otimes g\colon x \otimes y \to x' \otimes y'$.*
- *for any objects $x, y, z \in \mathcal{C}$ there is an isomorphism $a_{xyz}\colon (x \otimes y) \otimes z \to x \otimes (y \otimes z)$ called* **associator** *and satisfying the* **pentagon identity***:*



- *there is an object 1 such that, for any object $x \in C$, there are isomorphisms $l_x\colon 1 \otimes x \to x$ and $r_x\colon x \otimes 1 \to x$ called* **units** *satisfying the* **other identity***:*



- *finally, given $f\colon x \to y$, $g\colon y \to z$, $f'\colon x' \to y'$ and $g'\colon y' \to z'$, we require that $(g \otimes g')(f \otimes f') = (gf) \otimes (g'f')$, which just says that the following diagram is unambiguous:*



MacLane's theorem guarantees that if the above two diagrams commute, then any diagram that can be constructed from the associator and the units commutes.

Examples of monoidal categories are

- Grp: objects are groups, morphisms are group homomorphisms and $\otimes$ is the direct product of groups.
- nCob: the $\otimes$, both for objects and for morphisms, is the disjoint union of manifolds.
- Vect or Hilb: the $\otimes$ is the tensor product. This is how, in quantum mechanics, two things are put together.
- Elect: it has just one object, morphisms are electrical circuit elements, composition is serial combination of components, and $\otimes$ is parallel or shunted combination of components.

Notice that parenthesised expressions can be drawn as trees. This is how we start to see that category theory can be useful for physics, and also we get a hint of what the Biendenharn-Elliot identity means.

Note also that we are promoting all the theorems we proved in the first quarter to axioms. We can draw morphisms in any monoidal category as two-dimensional diagrams just like we did in Vect.

The moral of this story is that categorification—the process of replacing equalities by isomorphisms—is a way to understand processes (that is, time) at a deeper level.

Chapter 20

Quantum Mechanics form a Category-Theoretic Viewpoint (II)

Recall that in quantum mechanics, we describe the evolution of the state vector $\psi \in \mathcal{H}$ by a one-parameter, strongly continuous, unitary group U, i.e., $\psi(t) = U(t)\psi(0)$. Then Stone's theorem guarantees that $U(t) = \exp(-itH)$, where H is self-adjoint and corresponds to the energy of the system. Schrödinger's equation

$$i\frac{d}{dt}\psi(t) = -iH\psi(t),$$

which is equivalent to $\psi(t) = U(t)\psi(0)$ if $\psi(t)$ is differentiable, just states that "energy is the same as rate of change with respect to time". This is one of the most significant discoveries of the twentieth century, and although there were indications of this coming from analytical mechanics in the nineteenth century, it wasn't until the advent of quantum mechanics that the connection was established so explicitly.

Now the question is, how do we get the Hamiltonian H for a particular problem? A possible answer, and the first we shall explore here, is to "steal it" from classical mechanics.

20.1 A Point particle on a line

The classical Hamiltonian for a particle on a line is

$$H = \frac{p^2}{2m} + V(q), \qquad \text{with} \quad \{q, p\} = 1,$$

where q is an affine coordinate on the line and (q, p) are coordinates on the cotangent bundle. The Lagrangian is

$$L = p\dot{q} - H(q, p),$$

and the Hamilton equations of motion are

$$\dot{q} = \frac{\partial H}{\partial p} = p/m \qquad \text{and} \quad \dot{p} = -\frac{\partial H}{\partial q}.$$

This is translated into quantum mechanics as follows.

The Hilbert space of a point particle on a line is $\mathcal{H} = L^2(\mathbb{R})$, the space of complex functions on the line. States are unit-norm functions $\psi(x)$, and $\int_A |\psi(x)|^2 \, dx$ is the probability to find a particle in $A \subset \mathbb{R}$. The fact that ψ is normalised just means that the probability that the particle is *somewhere* is 1.

To get a Hamiltonian operator, we interpret $(q\psi)(x) = x\psi(x)$ and $(p\psi)(x) = -i\psi'(x)$. The form of these operators is restricted by the Dirac quantization condition $[q, p] = i\{q, p\}$ relating the commutator of the quantum operators and the Poisson bracket of the classical variables. Then the Hamiltonian is

$$H = \frac{p^2}{2m} + V(q) = \frac{-1}{2m}\frac{\partial^2}{\partial x^2} + V(q),$$

where $\big(V(q)\psi\big)(x) = V(x)\psi(x)$.

The Schrödinger equation is

$$i\frac{\partial}{\partial t}\psi(t,x) = \frac{-1}{2m}\frac{\partial^2}{\partial x^2}\psi(t,x) + V(x)\psi(x,t)$$

so, for a free particle,

$$\Big(i\frac{\partial}{\partial t} + \frac{1}{2m}\frac{\partial^2}{\partial x^2}\Big)\psi(t,x) = 0.$$

We can obtain solutions by guessing. Since the equation is linear we use exponential trial functions

$$\psi(t,x) = e^{-i(Et-kx)},$$

which satisfy Schröedinger's equation if $E = \frac{k^2}{2m}$. In fact, every solution is of the form

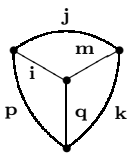
$$\psi(t,x) = \int \frac{dk}{2\pi} \exp\Big[-i\Big(\frac{k^2}{2m}t - kx\Big)\Big] f(k),$$

and if $f \in L^2(\mathbb{R})$, then $\psi(t) \in L^2(\mathbb{R})$ for all t. There are special solutions generically called "wavepackets" that preserve their form as they evolve, apart from getting gradually more and more spread out. These can be described as "Gaussian bumps with a corkscrew twist".

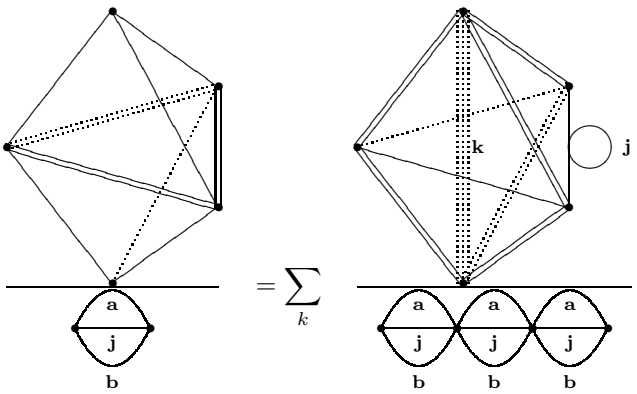
Chapter 21

Quantum Gravity (III)

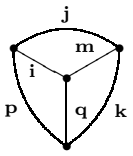
Exercise 16 ($2-3$ Pachner Move) *Using the formula relating the tetrahedron*



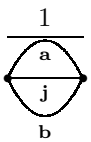
to the $6j$-symbols and the Biedenharn–Elliot identity, prove the following identity, which is the $2-3$ Pachner move up to "fudge factors" for lower-dimensional simplices (triangles and edges).



[Hint: Each side has a factor of



for each tetrahedron, a factor of



for each triangular face, and a factor of

for each edge. Factors appearing on both sides have been cancelled.]

21.1 Topological Quantum Field Theories (III)

21.1.1 Category Theory(III)

Definition 10 *Given monoidal categories $\mathcal{C}$ and $\mathcal{D}$, a functor $F\colon \mathcal{C} \to \mathcal{D}$ is* **strictly monoidal** *if, and only if,*

- $F(x \otimes y) = F(x) \otimes F(y)$ *for all objects* $x, y \in \mathcal{C}$,
- $F(f \otimes g) = F(f) \otimes F(g)$ *for all morphisms* $f, g \in \mathcal{C}$, *and*
- $F(1_{\mathcal{C}}) = 1_{\mathcal{D}}$, *where* $1_{\mathcal{C}}$ *and* $1_{\mathcal{D}}$ *are the identity objects for the* $\otimes$ *operation.*

Definition 11 *A monoidal category is* **braided** *if, and only if, for every objects $x, y \in \mathcal{C}$ there is an isomorphism $B_{x,y}\colon x \otimes y \to y \otimes x$ such that*

- *the* **hexagon identity** $BaB = aBa$ *is satisfied, and*
- *for any morphisms $f\colon x \leftarrow x'$ and $g\colon x' \to y'$, we have $B(f \otimes g) = (g \otimes 1)B(1 \otimes f) = (g \otimes f)B = (1 \otimes f)B(1 \otimes g)$.*

A braided category is **symmetric** *if $B_{x,y} = B_{y,x}^{-1}$.*

Definition 12 *A monoidal functor $F\colon \mathcal{C} \to \mathcal{D}$ between braided categories is* **braided** *if, and only if, $F(B_{x,y}) = B_{F(x),F(y)}$.*

A braided functor between symmetric categories is automatically symmetric.

Definition 13 *A monoidal category has* **duals for objects** *if, for every object x there is an object x^* and morphisms $i_x\colon 1 \to x \otimes x^*$ and $e_x\colon x^* \otimes x \to 1$ such that $(1_x \otimes e_x)(i_x \otimes 1_x) = 1_x$ and $(e_x \otimes 1_{x^*})(1_{x^*} \otimes i_x) = 1_{x^*}$.*

Any monoidal functor will preserve duals for objects.

Now that all the terms have been defined we can restate the definition of Topological Quantum Field Theory.

Definition 14 *A* **topological quantum field theory** *is a symmetric monoidal functor $Z\colon n\mathrm{Cob} \to \mathrm{Vect}$.*

Definition 15 *A* **unitary topological quantum field theory** *is a symmetric, monoidal $*$-functor $Z\colon n\mathrm{Cob} \to \mathrm{Hilb}$.*

Depending on the dimension of space-time, topological quantum field theories have varying degrees of complexity.

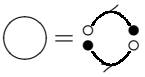
- TQFTs on 1Cob are extremely simple;
- TQFTs on 2Cob are related to commutative algebras;
- TQFTs on 3Cob involve the spin-network technology we have been developing, and are related to monoidal categories;
- nobody really understands TQFTs on 4Cob. This is unfortunate, as we have seen that General Relativity is like a TQFT on 4Cob with some extra features.

21.1.2 One-dimensional Topological Quantum Field Theories (I)

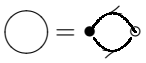
To see just how simple one-dimensional TQFTs are, let us characterise them completely.

A 1-cobordism is a one-dimensional manifold with boundary. This means it is a disjoint union of arcs and circles, and the boundary is formed by the endpoints of the arcs. If we consider oriented cobordisms, the boundary is also oriented, in such a way that, if $\gamma\colon p \to q$ is an oriented arc between p and q, then p is considered to be positively oriented and q is considered to be negatively oriented. If we represent positive points by $\circ$ and negative points by $\bullet$, then all the 1-cobordisms are disjoint unions of the following elementary ones:

Cobordisms are composed in such a way that the orientation of the arrows is preserved, so we need to identify a positive and a negative endpoints, which cancel due to the opposite orientations. For example,

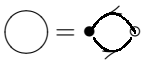


For our purposes we need identity cobordisms, so we adopt the convention of reversing the orientation of the output. Then, the elementary morphisms are

so that 

So let us characterise 1Cob as a category:

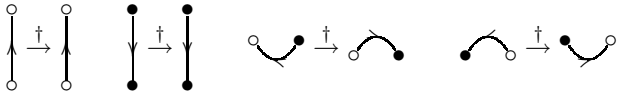
- objects are 0-dimensional compact oriented manifolds, i.e., finite disjoint unions of oriented points. $\mathbb{N}[\bullet, \circ]$.
- morphisms are 1-dimensional oriented manifolds, i.e., disjoint unions of

  and 

  Two 0-dimensional oriented manifolds $n \bullet + m\circ$ and $n' \bullet + m'\circ$ are cobordant—i.e., there is a 1-dimensional oriented cobordism between them—if, and only if, $n - m = n' - m'$.
- $\otimes$ is the disjoint union $\amalg$ of both 0-dimensional and 1-dimensional manifolds. This product is braided and symmetric because the manifolds are not assumed to be imbedded in any dimension, but are considered as abstract manifolds, so the disjoint union is diffeomorphic to the union of any permutation of the spaces.
- duals are as follows:

  $\bullet^* = \circ$ $\quad \circ^* = \bullet$ $\quad e_\bullet =$ $\quad e_\circ =$ $\quad i_\bullet =$ $\quad i_\circ =$

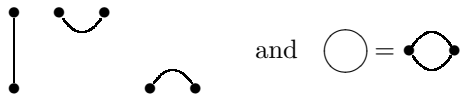
  and
- adjoints are

We now notice that as soon as, say, $\bullet$ is mapped to some object in a different category, the images of all the other objects and morphisms are fixed. Therefore, a 1-dimensional topological quantum field theory is determined by a single object $Z(\bullet)$.

If we consider unoriented cobordisms, the category simplifies even more. Indeed, an unoriented 1-cobordism is just a disjoint union of circles and segments, and the boundary is a union of (unoriented) points. As a category, unoriented 1Cob is described as follows:

- objects are 0-dimensional compact manifolds, i.e., finite disjoint unions of oriented points. $\mathbb{N}[\bullet]$.
- morphisms are 1-dimensional unoriented manifolds, i.e., disjoint unions of

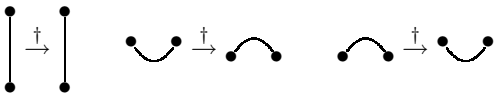


  Two 0-dimensional unoriented manifolds $n\bullet$ and $m\bullet$ are cobordant—i.e., there is a 1-dimensional unoriented cobordism between them—if, and only if, $n = m \pmod 2$.

- $\otimes$ is the disjoint union $\amalg$ of both 0-dimensional and 1-dimensional manifolds. For the same reasons as in the case of 1-dimensional oriented cobordisms, this product is braided and symmetric.
- duals are as follows:

  and

- adjoints are



  Note, however, that the adjoint of a morphism is the same as its dual, so this will be a special case in which a topological quantum field theory will automatically be unitary.

Oriented 1Cob goes to $\mathbb{C}$Hilb and $\mathbb{C}$Vect
Unoriented 1Cob goes to $\mathbb{R}$Vect = $\mathbb{R}$Hilb (see the remark about adjoints being equal to duals).
A 1-cobordism is a space (vector or Hilbert).
[does this mean that a 2-cobordism should map to a subcategory of Vect or Hilb?]

Chapter 22

Quantum Mechanics from a Category Theoretic Viewpoint (III)

22.1 Fields on $(1+1)$-dimensional spacetime (I)

Last time we quantised a free particle on a line. Now let's quantise the vacuum Maxwell equations on a $(1+1)$-dimensional spacetime with the topology of a cylinder. We will find that this is isomorphic to a particle on a line!

Our spacetime is $\mathbb{R} \times S^1(l)$ with metric $ds^2 = -dt^2 + l^2 dx^2$, where l is the "perimeter of the Universe" and $x \in \mathbb{R}/\mathbb{Z}$. The volume form is $d\mathrm{Vol} = l\, dt \wedge dx$ and the induced metric on 1-forms is $\langle dt, dt\rangle = -1$ and $\langle dx, dx\rangle = 1/l^2$. The Hodge dual, satisfying $\omega \wedge *\omega = \langle \omega, \omega\rangle d\mathrm{Vol}$, is therefore

$$\begin{cases} *1 = l\, dt \wedge dx, \\ *dt = -l\, dx, \\ *dx = -\frac{1}{l} dt, \\ *dt \wedge dx = -\frac{1}{l}. \end{cases}$$

Note that $*^2 = -1$.

Maxwell's theory is Yang–Mills theory with gauge group $\mathbb{R}^+$ or $U(1)$. These two cannot be distinguished at the level of local degrees of freedom since their Lie algebras are isomorphic. However, as we shall see they have different global properties that affect the quantum theory.

We first tackle the case of $\mathbb{R}^+$ gauge group, which is easier because the topology of the group itself is trivial. The exterior derivative is $d = d\theta \frac{\partial}{\partial x} + dt \frac{\partial}{\partial t}$, and the exterior covariant derivative is

$$\nabla = d + A \wedge \qquad \text{with} \quad A = a_t dt + a_x dx.$$

The curvature of the connection is

$$\begin{aligned} F &= \nabla^2 = dA = (\partial_t a_x - \partial_x a_t) dt \wedge dx, \qquad \text{and} \\ F &= \frac{1}{l}(\partial_x a_t - \partial_t a_x). \end{aligned}$$

We see that in $2D$ there is a one-component electric field and no magnetic field. Now, $*F$ is a *function* on space-time, and Maxwell's equation $d*F = 0$ simply implies that $\frac{1}{l}(\partial_t a_x - \partial_x a_t)$ is a constant, say e. On $\mathbb{R}^2$ this would be the end of the story, since $\partial_t a_x - \partial_x a_t = el$ uniquely determines A up to gauge transformations. On $\mathbb{R} \times S^1$, however, we have nontrivial solutions even if $e = 0$, as in the Aharonov-Bohm effect.

The "physical field" is A modulo gauge transformations, which are of the form $A \to A' = A - df$. By taking $f(t,x) = \int_0^t ds\, a_t(s,x)$ we get $A' = a(t,x)dx$. This is called "temporal gauge". Since $\partial_t a(t,x) = el$ is constant, we actually have $A' = \big(a(x) + elt\big)dx$. However there is still some gauge freedom left, because

by taking a gauge transformation generated by a suitable $f(x)$ we can cancel the variation in $a(x)$. To be precise,

$$f(x) = \int_0^x dy\, a(y) - x \oint_0^1 dy\, a(y)$$

generates the gauge transformation

$$a(x) \to a'(x) = a(x) - \frac{\partial}{\partial x} f(x) = \oint_0^1 dy\, a(y),$$

which is a gauge-invariant constant. Therefore, a solution up to gauge transformations of the vacuum Maxwell equations on $\mathbb{R} \times S^1$ is $A = (a + let)dx$. This, of course, looks like the trajectory of a free particle with initial position a and velocity le. The analogy is supported by the observation that the equation $d(*F) = 0$, which implies that e is constant, is like Newton's law stating that momentum is conserved for a particle of mass $1/l$.

The above is the way a gauge is traditional chosen in physics. However, this is just an explicit example of computing the first de Rham cohomology of a spacetime. Now, we can obtain the same parametrization of the space of connections modulo gauge transformations in a slicker way as follows. We observe that the first cohomology class of $\mathbb{R} \times S^1$ is generated by dx, therefore $A = a(t)dx$ up to gauge transformations. The curvature of the connection (which is gauge-invariant and therefore observable) is $F = \dot{a}(t)dt \wedge dx$, so $*F = -\frac{1}{l}\dot{a}(t) = -e(t)$ and the constant part of the connection can be observed by sending a test particle around a loop in the x direction and comparing it with a particle at rest. A more invariant experiment can be performed by sending two test particles at the speed of light in opposite directions and having them interact when they meet. Mathematically, these are much more complicated than the equivalent $a(t) = \oint_{t=\text{const}} A$. So we have two functions $a(t)$ and $e(t)$ related by the constraint $e(t) = \dot{a}(t)/l$. The analogy with classical mechanics is clear, as in that case $q(t)$ and $p(t)$ are related by the constraint $p = m\dot{q}$.

The Maxwell Lagrangian density for $A = a(t)dx$ is

$$\mathcal{L} = \frac{-1}{2} F \wedge *F = \frac{1}{2l}\dot{a}^2 dt \wedge dx = \frac{1}{2l^2}\dot{a}^2 d\text{Vol},$$

so integrating over dx we obtain the Lagrangian

$$L = \frac{1}{2l}\dot{a}^2 dt.$$

The momentum conjugate to $a(t)$ is $e(t) = \dot{a}/l$, so define $E = -*F$. The Hamiltonian

$$H = e\dot{a}dt - L$$

can be obtained by integrating the following Hamiltonian density over dx:

$$\mathcal{H} = E \wedge dA - \mathcal{L}.$$

Now, if we express the Hamiltonian and Hamiltonian density in terms of (a, e) and (A, E) respectively, we get

$$H = \frac{l}{2} e^2$$

and

$$\mathcal{H} = \frac{1}{2} E \wedge *E.$$

Now, in terms of independent fields (A, E), the Lagrangian density is

$$\mathcal{L} = E \wedge dA - \frac{1}{2} E \wedge *E = -\frac{1}{2} F \wedge *F$$

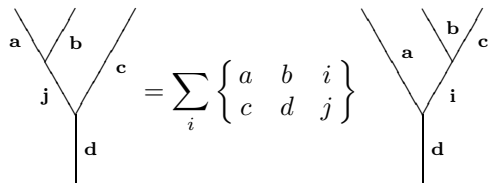
This is EF theory with an extra term involving the Hodge $*$. From this discussion it is clear that $\mathbb{R}^+$ Maxwell's theory on the cylinder of radius l is analogous to a free particle of mass $1/l$ on a line, where the 1-form $A = a(t)dx$ plays the role of position, and the 0-form (function) E plays the role of momentum conjugate to A. If we quantise the Hamiltonian $H = \frac{l}{2}e^2$ we obtain the following Schrödinger equation:

$$i\frac{\partial}{\partial_t}\psi(t, a) = -\frac{l}{2}\frac{\partial^2}{\partial a^2}\psi(t, a).$$

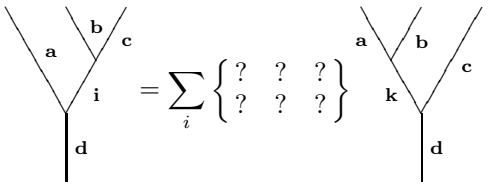
Chapter 23

Quantum Gravity (IV)

Exercise 17 (Orthogonality relations) *Recall the definition of the* $6j$ *symbols:*



Express



and derive a quadratic orthogonality relation that the $6j$ *symbols satisfy. [Hint: rotate the definition of the* $6j$ *symbols so that it goes "backwards".]*

23.1 Topological Quantum Field Theories (IV)

Recall the definitions of (U)TQFTs:

Definition 16 *A* **topological quantum field theory** *is a symmetric monoidal functor* $Z\colon n\mathrm{Cob} \to \mathrm{Vect}$.

Definition 17 *A* **unitary topological quantum field theory** *is a symmetric, monoidal* $*$*-functor* $Z\colon n\mathrm{Cob} \to \mathrm{Hilb}$.

Having analysed the categories of (un)oriented one-dimensional cobordisms, we can characterise all one-dimensional (U)TQFTs.

23.1.1 One-dimensional Topological Quantum Field Theories (II)

In one dimension we have the categories 1Cob and Un1Cob, and TQFTs will map them into Vect or Hilb.

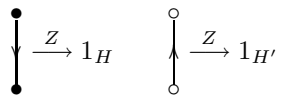
Assume $Z\colon 1\mathrm{Cob} \to \mathrm{Hilb}$ is given. Then, we have

- 0-dimensional oriented manifolds map to Hilbert spaces:

$$Z(\bullet) = H \qquad \text{and} \quad Z(\circ) = H'.$$

We will be able to find a relation between H and H' when we consider morphisms.

- The only morphisms we have are the identity morphisms:



- The union of manifolds maps to the tensor product: $Z(\amalg) = \otimes$, so

$$Z(n\bullet + m\circ) = H^{\otimes n} \otimes H'^{\otimes m}$$

and the tensor product of identity morphisms is the identity morphism. We also have $Z(\emptyset) = \mathbb{K}$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ is the base field of our vector spaces.

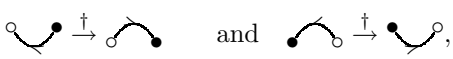
- Duality is given by $\bullet^* = \circ$, so $H' = Z(\circ) = Z(\bullet^*) = Z(\bullet)^* = H^*$. We also have

e_H: $\begin{array}{ccc} H^* \otimes H & \to & \mathbb{K} \\ \alpha \otimes v & \mapsto & \alpha(v) \end{array}$ and
i_H: $\begin{array}{ccc} \mathbb{K} & \to & H \otimes H^* \\ 1 & \mapsto & e_i \otimes e^i \end{array}$

where (e_i) is any basis of H and (e^i) is the associated dual basis of H^*. Observe that

$e^i(e_i) = \dim H = \dim H^*$.

- So far the whole construction works for $\mathbb{C}$Vect, but if we want to preserve adjoints we must map into $\mathbb{C}$Hilb. Recall that adjoints are



so

$$e_H^\dagger = i_{H^*} \qquad \text{and} \quad i_H^\dagger = e_{H^*}.$$

In conclusion, 1Cob admits topological quantum field theories into $\mathbb{C}$Vect or $\mathbb{R}$Vect which are determined by the single datum $Z(\bullet)$. Unitary TQFTs incorporate the adjoint operation and map into $\mathbb{C}$Hilb or $\mathbb{R}$Hilb. The image of a TQFT can be given a Hilbert space structure by adding the image of the 1Cob adjoint, i.e., the relations $e_H^\dagger = i_{H^*}$ and $i_H^\dagger = e_{H^*}$.

The moral of this story is that an nTQFT associates a complex number to each closed n-dimensional manifold. This is called the partition function of the manifold. If the manifold is self-adjoint, this will be a real number. In the case of 1TQFTs, a closed manifold is a disjoint union of circles, which are self-adjoint and evaluate to the dimension of $H = Z(\bullet)$, which is a real number.

The case of unoriented cobordisms is simpler, and it can be obtained from the previous one by adding relations:

- There is only one basis Hilbert space $H = Z(\bullet)$, which is *self-dual*: $\bullet^* = \bullet$. This means that H is isomorphic to H^*, which requires H to be a real Hilbert space, as in the complex case H is anti-isomorphic to H^*. Note that in this case it cannot be merely a vector space.

- The evaluation map is a bilinear inner product:

$e_H =: H \otimes H \to \mathbb{R}$

and,

- as we observed last time, duals and adjoint coincide, so $\mathbb{R}$Vect and $\mathbb{R}$Hilb coincide, where $\mathbb{R}$Hilb has the inner product given by the evaluation.

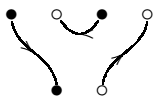
Therefore, we have $Z\colon \mathrm{Un1Cob} \to \mathbb{R}\mathrm{Hilb} = \mathbb{R}\mathrm{Vect}$.

23.1.2 Associative algebras

In any dimension there is a God-given closed manifold, and that is the n-sphere (the boundary of an $(n+1)$-disk). The 0-sphere is

$$S^0 \simeq \bullet \quad \circ \xrightarrow{Z} H \otimes H^* \simeq \mathrm{End}(H) := \mathrm{Hom}(H, H).$$

That is, S^0 maps to the associative algebra of endomorphisms of $H = Z(\bullet)$. Multiplication is given by

$e_\bullet =$ via 

and the unit is

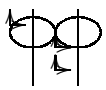
$i_\bullet =$

This defines an associative unitary algebra. We already proved this in a different context when we were developing the diagrammatic methods for linear algebra.

23.1.3 Two-dimensional Topological Quantum Field Theories (I)

Two-dimensional topological quantum field theories map (un)oriented 2-dimensional cobordisms to Vector or Hilbert spaces. Now, the objects in 2Cob are one-dimensional compact manifolds, that is, disjoint unions of circles. By analogy with the 0-spheres of the previous paragraph, we would expect these circles to be elements of an algebra. It turns out that all 2TQFTs actually map not only into Hilbert spaces, but into Hilbert algebras.

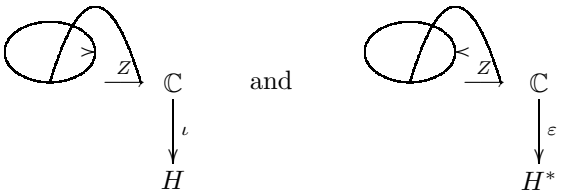
Let us now study oriented two-dimensional cobordisms and the possibilities for 2TQFTs. First, a note about orientation. As we know, an oriented 2-cobordism will be an oriented surface with boundary, which will be divided into an "input" and an "output" part. We orient the input in such a way that a positively oriented basis for the input, followed by an outward vector, is a positively oriented basis for the 2-cobordism. The output is oriented similarly, but with an inward vector completing the basis. In pictures:



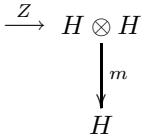
has both circles positively oriented.

Any two-dimensional cobordism can be constructed from the following elements:

- Two units



- two multiplications



...etc...

Chapter 24

Quantum Mechanics from a Category Theoretic Viewpoint (IV)

24.1 Fields on $(1+1)$-dimensional spacetime (II)

Last time we saw that (1+1) vacuum Maxwell equations on a cylinder were secretly the same thing as the point particle on a line. (This is a great thing because we know how to quantise the particle.) Now let's play around with the theory. Some things we can do are changing the gauge group or deforming the manifold.

We saw that, on $\mathbb{R} \times S^1(l)$ (a cylinder with perimeter l), the Maxwell connection takes the form $A = (a + elt)dx$, where e is a constant. The connection has curvature $F = dA = eld\text{vol}$, and $*F = e$. Apart from e there is another gauge-invariant quantity, which is the integral of A around a loop going around the cylinder once. We call this constant

$$a = \oint_{t=t_0} A.$$

We saw that, in temporal gauge, $a(t) = a + elt$, which comes from the equations

$$\dot{a} = el \qquad \text{and} \quad \dot{e} = 0.$$

This motivates the identification with a point particle of position a, momentum e and mass $1/l$, with equations

$$\dot{q} = p \qquad \text{and} \quad \dot{p} = 0.$$

The classical solution is determined by a pair (a, e) or (q, p) (initial conditions). This is also good because the Hamiltonians match up, as we saw last time. The quantised maxwell equations are therefore

$$i\frac{\partial}{\partial t}\psi(a) = -\frac{l}{2}\frac{\partial}{\partial a}\psi(a).$$

Now let's consider Maxwell's theory with gauge group $U(1)$ as opposed to $\mathbb{R}^+$. If the connection is a $U(1)$ connection, on each point of $M = \mathbb{R} \times S^1$ we have a fiber (not canonically) isomorphic to the gauge group, and the connection tells us how to parallel transport objects around a loop.

The only context in which we encountered this parallel transport was in computing

$$a = \oint_{t=t_0} A \in \mathfrak{u}(1).$$

If the group is $U(1)$, a is only determined up to $2\pi i$, because we can take $f(x) = e^{2\pi i n x}$, and $A \to A' = A - 2\pi i n$. In fact, the holonomy of the connection around the loop is technically a group element representing the effect of parallel transport on the various representations of the group, and this is written as

$$e^{\oint_{t=t_0} A} \in U(1).$$

Let's just call this a. Then, the role of $\dot{a}$ is played by $a^{-1}\dot{a}$.

A solution of Maxwell's $U(1)$ theory up to gauge transformations is, then, $(a, e) \in (\mathbb{R}/2\pi\mathbb{Z}) \times \mathbb{R}$. This is a free point particle on a circle or radius 1 with the same Hamiltonian as the free particle, but Schrödinger's equation will be able to tell the difference.

The theory is quantised by taking the Hilbert space $\mathcal{H} = L^2(S^1)$ and the Hamiltonian $H = -\frac{l}{2}\frac{d^2}{da^2}$. Schrödinger's equation is

$$\Big(i\frac{\partial}{\partial t} + \frac{l}{2}\frac{\partial^2}{\partial a^2}\Big)\psi(t,a) = 0 \qquad \text{with} \quad \psi(t,a) = \psi(t, a+2\pi).$$

So the exponentials $\psi(t,x) = \exp\big(-i(Et - kx)\big)$ are solutions if

$$E = \frac{l}{2}k^2 \qquad \text{with} \quad k \in \mathbb{Z}.$$

Note how this differs from the solution in the case of a particle in a line, with a continuous spectrum and generalised eigenvectors.

We therefore have two theories that no local experiments can distinguish, but quantum mechanics or global experiments can. Of course, the energy E in this case is the energy of the whole universe, so that is not a local observable and the fact that quantum gravity can tell the difference does not contradict the fact that $\mathbb{R}^+$ and $U(1)$ theory are locally indistinguishable.

We now want to relax the three fixed structures we have: the topology, the metric and the foliation into space-like slices. First notice that classically, $2D$ vacuum Maxwell theory has a lot of symmetry that is not entirely obvious: the whole theory is invariant under all area-preserving diffeomorphisms of space-time.

For a proof, all we need to show is that the action itself is invariant under area-preserving diffeomorphisms.

$$S = \frac{1}{2}\int F \wedge *F.$$

Under a diffeomorphism ϕ, A is pulled back to $\phi^*(A)$ and so is the curvature: $F \to \phi^*(F)$. So we need to calculate

$$S = \frac{1}{2}\int_M F \wedge *F = \frac{1}{2}\int_M \phi^*(F') \wedge *\phi^*(F')$$

but notice that any 2-form is equal to a function times the volume form, so we have $F' = E' d\mathrm{Vol}'$ and, since ϕ is area-preserving, $\phi^*(d\mathrm{Vol}') = d\mathrm{Vol}$. Therefore,

$$\begin{aligned} S &= \frac{1}{2}\int_M \phi^*(F') \wedge *\phi^*(F') = \frac{1}{2}\int_M \phi^*(E' d\mathrm{Vol}') \wedge *\phi^*(E' d\mathrm{Vol}') = \frac{1}{2}\int_M \phi^*(E')^2 d\mathrm{Vol} = \\ &= \frac{1}{2}\int_M \phi^*(E'^2) d\mathrm{Vol} = \frac{1}{2}\int_M \phi^*(E'^2 d\mathrm{Vol}'), \end{aligned}$$

since $*d\mathrm{Vol} = 1$.

We now use

$$\int_M \phi^*\omega = \int_{\phi^{-1}(M)} \omega,$$

and we obtain

$$S = \frac{1}{2}\int_{\phi(M)} E'^2 d\mathrm{Vol}' = \frac{1}{2}\int_M F' \wedge *F'.$$

Now, there is a theorem of Moser asserting that, if there are two non-intersecting circles, there is an area-preserving diffeomorphism that takes them to two "nice" ($t = \text{const}$) circles. So trying to understand time evolution between any two closed curves is not harder than on a very simple cylinder.

Therefore the time-evolution operator $U\colon L^2(\mathbb{R}) \to L^2(\mathbb{R})$ corresponding to time evolution from now to then, is just a function of the area of the spacetime between "now" and "then"!

$$U = e^{-iAH},$$

where A is the area of the spacetime enclosed by the two curves. So, in $(1+1)$-Yang-Mills theory, time *is* area!

What we will do next is do $t \to it$, which will turn our theory into statistical mechanics or stochastic processes... Then we will be able to study more general spacetimes, like the ones we were talking about on track 1:

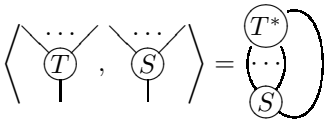
(diagram: trousers)

This will give us something not unlike the topological quantum field theories of Track 1.

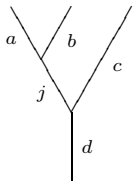
Chapter 25

Quantum Gravity

The orthogonality relation gets its name from the following inner product on the space of intertwining operators:

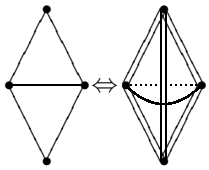


With respect to this inner product, the

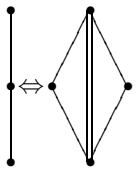


form an orthogonal basis of intertwiners. The orthogonality relation just says that the $6j$ symbols are an orthogonal matrix.

Exercise 18 *Show that the orthogonality relation for the* $6j$ *symbols corresponds to this move on* $3d$ *triangulations:*

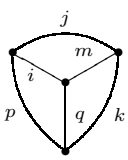


This is the $3d$ *analogue of the following move in two dimensions:*

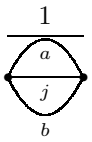


[Hint: translate the move into an equation between spin networks using these rules:

1. *one tet-net per tetrahedron*



2. the reciprocal of a theta net for each triangle



3. one loop per edge.

Using the relation between tet nets and $6j$ symbols, show that this move is equivalent to the orthogonality identity!]

This move together with the $2-3$ move are insufficient to go from any triangulation to any other, because the number of vertices of the triangulation is preserved. The $1-4$ move allows us to change the number of vertices, but it turns out that, when it is translated into spin networks it diverges because we get to sum over all representations of $SL(2,\mathbb{C})$. This problem is solved by introducing quantum groups, which have only finitely many irreducible representations.

25.1 Topological Quantum Field Theories (V)

25.1.1 Two-Dimensional Topological Quantum Field Theories

Suppose we have

$$
\begin{array}{lcl}
Z\colon & \text{2Cob} & \to \quad \text{Vect} \\
 & & \mapsto \quad \mathcal{A} \\
 & \text{(pants)} & \mapsto \quad m\colon \mathcal{A}\otimes\mathcal{A}\to\mathcal{A} \\
 & \text{cap} & \mapsto \quad \iota\colon \mathbb{C}\to\mathcal{A} \\
 & \text{cup} & \mapsto \quad \tau\colon \mathcal{A}\to\mathbb{C}
\end{array}
$$

This is an Abelian associative algebra with unit ι, and the additional structure τ gives rise to an inner product $g=\tau\circ m$:

$$\text{(diagram: U)} \overset{Z}{\mapsto} g\colon \mathcal{A}\otimes\mathcal{A}\to\mathbb{C}.$$

As usual, g determines a mapping $\sharp\colon \mathcal{A}\to\mathcal{A}^*$, and g is non-degenerate if, and only if, $\sharp$ is an isomorphism. The map $\sharp$ is obtained from g by composition with the unit endomorphism, which we represent as a bundt:

$$\text{(diagram)}$$

$$\mathcal{A}\simeq\mathcal{A}\otimes\mathbb{C}\xrightarrow{1_{\mathcal{A}}\otimes i_{\mathcal{A}}}\mathcal{A}\otimes\mathcal{A}\otimes\mathcal{A}^*\overbrace{\xrightarrow{m\otimes 1_{\mathcal{A}^*}}\mathcal{A}\otimes\mathcal{A}^*\xrightarrow{\tau\otimes 1_{\mathcal{A}^*}}}^{g\otimes 1^*_{\mathcal{A}}}\mathbb{C}\otimes\mathcal{A}^*\simeq\mathcal{A}^*$$

But now we observe that the cobordism corresponding to $\sharp$ is invertible, and its inverse (corresponding to $\flat$) is

$$\text{(diagram and proof)}$$

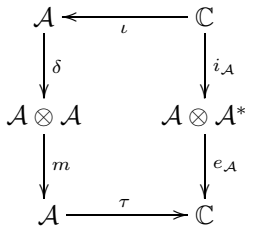
Therefore, g is a non-degenerate metric. An algebra with a non-degenerate inner product is called a Frobenius algebra.

Theorem 11 *If Z is a $2dTQFT$, then $Z($ $)$ is a* **commutative Frobenius algebra** *with the operations described above. Conversely, give any commutative Frobenius algebra $\mathcal{A}$ there is a unique $2dTQFT$ such that $Z($$)=\mathcal{A}$.*

Uniqueness is proven by showing that the multiplication on $\mathcal{A}^*$ and the comultiplications on $\mathcal{A}$ can be obtained from the product on $\mathcal{A}$ by means of $\sharp$ and $\flat$.

Existence is harder, because one has to show that if M and M' are diffeomorphic cobordisms obtained by composing elementary cobordisms in different ways, then $Z(M) = Z(M')$. This gives rise to all sorts of identities that a Frobenius algebra must satisfy, for example,

flat torus $\simeq$ vertical torus $\Leftrightarrow$



The primary tool for carrying out this proof is "serf theory".

As an example, consider the matrix algebra $\mathcal{A} = M_n(\mathbb{C})$. The most general bilinear form is of the form $g(A,B) = \mathrm{Tr}(\tau AB)$, where τ is some matrix. We require that $g(A,B) = g(B,A)$, so $\mathrm{Tr}(\tau\big[A,B\big]) = 0$ for all A, B. Let $\tau = \tau^j_i E^i_j$, where E^i_j are the elementary matrices of $\mathcal{A}$, and consider $A = E^i_j$ and $B = E^k_l$. Then, $g(A,B) = \mathrm{Tr}(\tau^h_m E^m_h E^i_j E^k_l) = \tau^i_l \delta^k_j$, so $\tau^i_l \delta^k_j = \tau^k_j \delta^i_l$. This implies that $\tau^i_j = \alpha \delta^i_j$ for some α, so we have $\tau_\alpha(\cdot) = \alpha \mathrm{Tr}(\cdot)$.

Theorem 12 *Any Frobenius algebra is isomorphic to a direct sum of algebras of this sort.*

Theorem 13 (Wedderburn) *Any simple Frobenius algebra is isomorphic to a matrix algebra.*

Corollary 3 *Any commutative Frobenius algebra is of the form $\oplus(\mathbb{C}, \alpha)$, with component-wise addition and multiplication; alternatively, any commutative Frobenius algebra is an algebra of diagonal matrices with an invertible diagonal matrix as its α.*

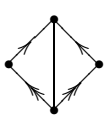
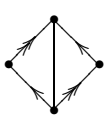
In higher dimensions, this "hard knuckles" approach of enumerating all manifolds gets pretty hard pretty quick, so in higher dimensions it is necessary to switch to "lattice TQFTs". This is a reflection of the fact that only in two dimensions is the problem of classifying all compact manifolds completely solved. In higher dimensions, one resorts to various combinatorial-topological techniques like triangulations, cellulations, CW-complexes, handle-body theory, Heegård decompositions, and so on, to obtain algorithms that allow one to determine whether two manifolds can be homeomorphic. In a similar spirit, we will specify TQFTs by giving an algorithm for extracting an algebraic structure from a manifold specified combinatorially. The trick will be to prove that the result does not depend on the combinatorial presentation, and to this effect we will have to take theorems like that of Pachner, which give conditions for two combinatorial presentations to be equivalent, and translate the theorems into relations that the algebraic objects must satisfy. This is what we have been doing so far in our exercises about $6j$ symbols and Pachner moves.

25.2 Lattice Field Theory

Until we get more sophisticated, we will just think of a **lattice field theory** ad a recipe to get an $nTQFT$ from a combinatorial presentation (usually a triangulation) of an n-dimensional cobordism. Triangulation independence has to be checked using the Pachner moves. There is much more to lattice field theory than this, but we have to start somewhere, and this seems like a good starting point.

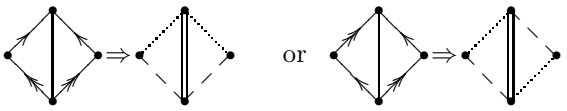
Let's illustrate the principle in the $2d$ case, where we already know the answer, by computing $Z(M)\colon \mathbb{C} \to \mathbb{C}$ for a closed 2-manifold. Let's follow the procedure step by step:

1. Triangulate M:

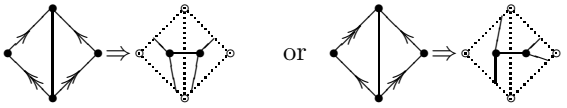
(diagram: sphere) $\Rightarrow$  or (diagram: torus) $\Rightarrow$ 

2. There are two things one can do at this stage, and they determine how the next steps are phrased, but it will be clear that they are actually the same thing:

   - "Explode the triangulation" (following Fukuma and Hosano).



   - Take the "dual cellulation" of the triangulation.

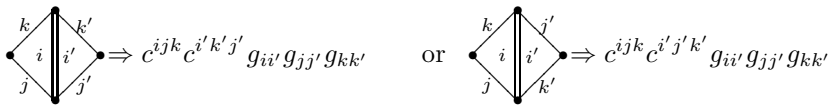


   (Here we are trying to use the distance from the intersection of primal and dual edges to the primal vertices as a way to indicate how the edges of the primal triangulation are to be identified. We see that for the triangulation of the sphere there are three dual faces, three dual edges and two dual vertices; while for the triangulation of the torus there is one dual face, three dual edges and two dual vertices. Note that the dual vertices are trivalent.)

3. Choose a vector space V, and elements $c \in V \otimes V \otimes V$ and $g \in V^* \otimes V^*$. If you are a physicist, or if you are using abstract index notation, you are allowed to write c^{ijk} and g_{ij}.

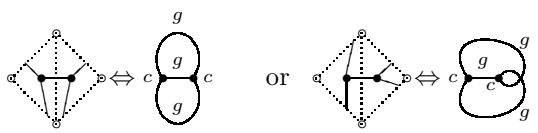
4. Now we come to defining the linear operators associated to each triangulation. Both approaches, the primal and the dual, illuminate each other.

   (a) Label the three edges around each triangle with $i_\Delta, j_\Delta, k_\Delta$, where Δ is a label indicating to which triangle the edge belongs. This is why we exploded the triangulation by duplicating the edges. Then, for each triangle Δ write $c^{i_\Delta j_\Delta k_\Delta}$, if (i, j, k) is the order in which you encounter the labels as you go around the positively oriented boundary of the triangle, and for each exploded edge shared by triangles Δ and Δ' write $g_{i_\Delta l_{\Delta'}}$:



   We now observe that there is an ambiguity: apart from the cyclic order of the edges around a triangle, how does the triangulation know the difference between (i, j, k) and (j, k, i)? We conclude that we must impose the condition $c^{ijk} = c^{jki} = c^{kij}$. Similarly, because there is no reason to enumerate the triangles in one order or another, we have to impose the condition $g_{ij} = g_{ji}$.

   (b) Label each vertex of the dual triangulation with a c and each edge of the dual triangulation with a g. Clearly c must be cyclic-symmetric and g must be symmetric. Note also that any cyclic-symmetric $c \in V \otimes V \otimes V$ is the average of a completely symmetric element and a completely antisymmetric element.



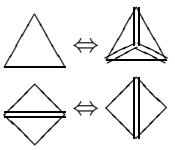
   Translating the diagrams into linear algebra, the same expressions as above are obtained. Notice the similarity with spin networks, and also the differences. The $\mathrm{SL}(2, \mathbb{C})$ spin networks have labels on the edges, which is the same as to say that we can assign different vector spaces V to each edge of the triangulation. This means that there is some extra structure associated to spin network models, over and above the topological structure that TQFTs are able to see. The key here is provided by the compatibility conditions for intertwining operators. Recall that the compatibility was somewhat mysteriously related to the triangle inequality. Notice that the

j label edges of the dual cellulation, which correspond to sides of triangular faces of the primal triangulation, and that the compatibility conditions just say that the three
j on the sides of a triangle must be valid lengths for the sides of a flat triangle. In the topological setting there is no notion of length, so we have no labels.

Now we come to the final question that we need to ask to validate the model, and that is, is this computation triangulation-independent, and if so, how do we prove it?

The answer is provided by Pachner's theorem. This theorem is valid in any dimension, and we have been exploring the three-dimensional case in the exercises. The two-dimensional case was probably proved by Alexander in the 1920's, and is as follows:

Theorem 14 (Alexander, Pachner) *Two triangulations of the same two-dimensional compact manifold can be obtained from each other by a succession of moves of the form:*



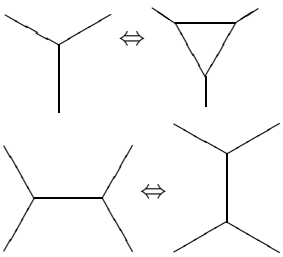
$(1-3$ move)

$(2-2$ move)

There are two things to notice in the statement of this theorem. The first is that both moves can be obtained from a tetrahedron by colouring the faces of the tetrahedron with two colours, say, black and white. Then, the triangulations on either side of $\Leftrightarrow$ in each move are simply the black and white triangulations. Similarly, the $(1-4)$ and $(2-3)$ moves in three dimensions are the two ways of bi-colouring the five tetrahedral faces of a four-simplex.

Exercise 19 *State Pachner's theorem in one dimension and use it to prove that the only one-dimensional triangulable manifolds are the circle and the segment. [Hint: a two-dimensional simplex is a triangle.]*

The second thing is that, taking duals of the Alexander-Pachner moves, we have

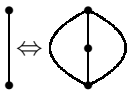


(Star-triangle relation)

(Crossing symmetry)

These diagrams crop up virtually everywhere in statistical physics and field theory.

Theorem 15 *Given the $(2-2)$ move, the $(1-3)$ move is equivalent to*

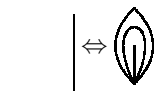


(Bubble move)

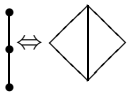
The proof is simple, but it rests on the assumption that the bubble move must be interpreted as being inside some triangulation, that is, the initial edge must be a side of some triangle in the triangulation. First, $(1-3)$ implies (bubble):

And now, (bubble) implies $(1-3)$:

Exercise 20 *Prove the following moves on two-dimensional triangulations:*



(Cone move)



(Other Bubble move)

Chapter 26

Quantum Mechanics from a Category Theoretical Viewpoint

We are going to continue sneaking up on TQFTs from the field theory side.

In $2d$ vacuum electromagnetism, we have worked out the time evolution for the classical theory, and we have seen that it is determined by the area enclosed by the initial and final circles.

(diagram: cylinder with wiggly boundaries)

Area and time are in fact so similar in this theory that we are going to call the area t. Then we have seen that (a, e) is changed to $(a + et, e)$. Similarly, the quantum version is as follows.

$$\begin{array}{c} \psi \in L^2(\mathbb{R}) \\ \Big\downarrow {\scriptstyle e^{-itH}} \\ \psi \in L^2(\mathbb{R}) \end{array}$$

where H is the Hamiltonian

$$(H\psi)(a) = \frac{-1}{2}\frac{d^2}{da^2}\psi(a)$$

More generally, if g is any Lorentzian metric on $[0, 1] \times S^1$ such that the boundary circles are space-like, then there is an area-preserving diffeomorphism taking it to the standard metric $ds^2 = -dt^2 + dx^2$ on $[0, t] \times S^1$, where t is the area of the original manifold. Since vacuum electromagnetism is invariant under area-preserving diffeomorphisms, the same time-evolution operator is valid.

But we would like to work with more interesting two-dimensional topologies like the ones we are considering in Track 1.

Now for more general spacetimes (i.e. cobordisms). The problem shows up right away when one realises that there is no way to put a Lorentzian metric on

(diagram: upside-down trousers)

that makes all the boundary circles space-like. The reason for this is that the light cones don't match up globally. If it were possible, there would be a nowhere-vanishing time-like vector field, but by the Poincaré-Hopf theorem any vector field of the trousers must vanish at a point.

So we switch to studying Riemannian metrics of the form $ds^2 = dt^2 + dx^2$.

For example, our cylinder admits the metric $ds^2 = dt^2 + dx^2$. Formally, all we have done is replace t by $-it$, and this means that time evolution of the state in quantum theory is given by

$$\begin{array}{c} \psi \in L^2(\mathbb{R}) \\ \Big\downarrow {\scriptstyle e^{-tH}} \\ \psi \in L^2(\mathbb{R}) \end{array}$$

which is a solution of the heat equation:

$$\frac{\partial}{\partial t}\psi(a) = \frac{1}{2}\frac{d^2}{da^2}\psi(a).$$

If "spacetime" is a Riemannian cylinder, we get

$$\begin{array}{c} \psi \in L^2(\mathbb{R}) \\ \Big\downarrow{\scriptstyle e^{-tH}} \\ \psi \in L^2(\mathbb{R}) \end{array}$$

But now we can think of what happens if "spacetime" has a more interesting topology. For example:

(diagram: upside-down trousers)

$$\begin{array}{c} \psi \in L^2(\mathbb{R}) \otimes L^2(R) \\ \Big\downarrow{\scriptstyle T} \\ \psi \in L^2(\mathbb{R}) \end{array}$$

What form can the time evolution operator take in this case?

(diagram: upside-down trousers with seams: two circles to a circle)

$$\begin{array}{c} L^2(\mathbb{R}) \otimes L^2(\mathbb{R}) \\ \Big\downarrow{\scriptstyle e^{-t_1 H} \otimes e^{-t_2 H}} \\ L^2(\mathbb{R}) \otimes L^2(\mathbb{R}) \\ \Big\downarrow{\scriptstyle m} \\ L^2(\mathbb{R}) \\ \Big\downarrow{\scriptstyle e^{-t_3 H}} \\ L^2(\mathbb{R}) \end{array}$$

We can assume that T is of this form, but we have to impose the consistency condition that it cannot depend on how the $t = t_1 + t_2 + t_3$ is divided.

Classically, m just corresponds to summing the holonomies: $(a_1, e_1) \otimes (a_2, e_2) \mapsto (a_1 + a_2, e + 1, e + 2)$. Now, there is a natural way to take an operation on a space and obtain an operation on functions on that space. From $f(a, a')$, we get $f^*(\psi)(a, a') = \psi\big(f(a, a')\big)$. In other words, if the classical configuration spaces are mapped as

$$\begin{array}{c} \mathbb{R} \times \mathbb{R} \\ \Big\downarrow{\scriptstyle f} \\ \mathbb{R} \end{array}$$

the quantum version turns out to be the pull-back of f:

$$\begin{array}{c} L^2(\mathbb{R}) \\ \Big\uparrow{\scriptstyle f^*} \\ L^2(\mathbb{R}) \otimes L^2(\mathbb{R}) \end{array}$$

and it represents not time evolution on upside-down trousers, but on the regular ones.

So, in fact, we can hope that the following (where $f(a,a') = a + a'$) is the right time evolution for a different spacetime:

(diagram: trousers with seams)

$$\begin{array}{c} L^2(\mathbb{R}) \\ \downarrow{\scriptstyle e^{-t_3 H}} \\ L^2(\mathbb{R}) \\ \downarrow{\scriptstyle f^*} \\ L^2(\mathbb{R}) \otimes L^2(\mathbb{R}) \\ \downarrow{\scriptstyle e^{-t_1 H} \otimes e^{-t_2 H}} \\ L^2(\mathbb{R}^2) \end{array}$$

But there is a problem: $f^*(\psi)$ is not in $L^2(\mathbb{R}^2)$, as we can readily see!

$$\|f^*(\psi)\|^2_{L^2(\mathbb{R}^2)} = \int da_1\, da_2\, \psi(a_1 + a_2) = \int d(a_1 + a_2) \frac{d(a_1 - a_2)}{2} \psi(a_1 + a_2) = \frac{1}{2} \|\psi\|_{L^2(\mathbb{R}^2)} \int_{\mathbb{R}} dx = \infty.$$

The reason, as is evident from the above calculation, is that the measure of configuration space is infinite!

There are two ways out of this quandary:

- relax, stop worrying and generalise the heck out of quantum mechanics so that it works on vector spaces rather than on Hilbert spaces.

  We can still explore the structure of the theory in this way, and it may be instructive even if we cannot calculate finite probabilities.

- switch from R^+ to $U(1)$, so that the infinity does not arise because everything is the same, except that the circle has finite measure.

  Let's do this in detail, because there are subtleties associated with the need to take sums modulo $2\pi i$. In fact, I am going to use multiplicative notation so $z_1 = e^{a_1}$ and $z_2 = e^{a_2}$ are complex numbers of norm 1, ψ is a function defined on $S^1 = \{z \in \mathbb{C} \colon |z| = 1\}$ and $f(z_1, z_2) = z_1 z_2$. Note that the uniform unit measure on the circle is $\frac{dz}{2\pi i z}$. Now we can calculate

$$\|f^*(\psi)\|^2_{L^2(S^1 \times S^1)} = \int_{S^1 \times S^1} \frac{dz_1}{2\pi i z_1} \frac{dz_2}{2\pi i z_2} \psi(z_1 z_2)$$

  Let $u = z_1 z_2$ and $v = z_1 / z_2 = z_1 z_2^*$. Then,

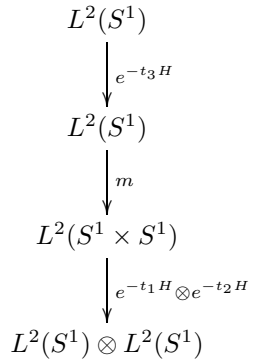
$$\begin{cases} \frac{du}{2\pi i u} = \frac{dz_1}{2\pi i z_1} + \frac{dz_2}{2\pi i z_2} \\ \frac{dv}{2\pi i v} = \frac{dz_1}{2\pi i z_1} - \frac{dz_2}{2\pi i z_2} \end{cases} \quad \text{implies} \quad \begin{cases} \frac{dz_1}{2\pi i z_1} = \frac{1}{2}\big(\frac{du}{2\pi i u} + \frac{dv}{2\pi i v}\big) \\ \frac{dz_2}{2\pi i z_2} = \frac{1}{2}\big(\frac{du}{2\pi i u} - \frac{dv}{2\pi i v}\big) \end{cases}$$

  so $\frac{dz_1}{2\pi i z_1} \wedge \frac{dz_2}{2\pi i z_2} = \frac{1}{2} \frac{dv}{2\pi i v} \wedge \frac{du}{2\pi i u}$. Therefore,

$$\|f^*(\psi)\|^2_{L^2(\S^1 \times S^1)} = \frac{1}{2} \int_{S^1 \times S^1} \frac{dv}{2\pi i v} \frac{du}{2\pi i u} \psi(u) = \frac{1}{2} \|\psi\|^2_{L^2(S^1)} \int_{S^1} \frac{dv}{2\pi i v} = \frac{1}{2} \|\psi\|^2_{L^2(S^1)}.$$

  and if we want m to be normalised we must let $m(\psi)(z_1, z_2) = \frac{1}{\sqrt{2}} \psi(z_1 z_2)$ or, equivalently, $m(\psi)(a_1, a_2) = \frac{1}{\sqrt{2}} \psi\big((a_1 + a_2) \mathrm{mod} 2\pi i\big)$.

(diagram: trousers with seams)



Moral: compactness is a good thing!

Next, we will have to check the consistency condition, and explore other space-times.

Part III

Spin foam models of 3D QG

Part IV

Miscellaneous

Chapter 27

Building Spacetime from Spin (I)

The Dirac equation

$$(\not{\partial} + im)\psi = 0$$

describes electrons and, when Dirac discovered it, predicted the existence of a particle with the same mass and opposite charge, the positron. The terms of the equation are

- ψ, called a "spinor field", is a function $\psi\colon \mathbb{R}^4 \to \mathbb{C}^4$, where $\mathbb{R}^4$ is $(3+1)$-dimensional Minkowski space and $\mathbb{C}^4$ is the space of "Dirac spinors";
- m is the mass of the spinors described by ψ (electrons/positrons); and
- $\not{\partial}$ is called the Dirac operator, and has the form $\not{\partial} = \gamma^\mu \partial_\mu$, where γ^μ are the Dirac matrices which ensure that $\not{\partial}$ is a Lorentz-covariant operator.

27.1 Spinors

The space of Dirac spinors $\mathbb{C}^4$ carries a representation of the Lorentz group on $\mathbb{R}^4$, but the relationship between $\mathbb{R}^4$ and $\mathbb{C}^4$ is more complex that just complexifying spacetime coordinates. The relationship is most easily explained in terms not of Dirac spinors but of Weyl spinors in $\mathbb{C}^2$,

$$\begin{array}{ccc} \text{Dirac spinors} & & \text{Weyl spinors} \\ \mathbb{C}^4 & \simeq & \mathbb{C}^2 \oplus (\mathbb{C}^2)^* \end{array}$$

The notation $(\mathbb{C}^2)^*$ denotes the dual of $\mathbb{C}^2$, which is isomorphic to $\mathbb{C}^2$ as a vector space but carries the dual representation of the Lorentz group. In a certain sense, $\mathbb{C}^2$ is the space of left-handed Weyl spinors, and its dual is the space of right-handed spinors.

To understand Weyl spinors we need to study four symmetry groups that appear all over the place in physics: SO(3), SU(2), $\mathrm{SO}_0(3,1)$ and $\mathrm{SL}(2,\mathbb{C})$.

27.1.1 The rotation group

The rotation group SO(3) is the space of linear transformations $R\colon \mathbb{R}^3 \to \mathbb{R}^3$ which leave an inner product and an orientation on $\mathbb{R}^3$ invariant. In matrix notation, we write the inner product as

$$g(x,x) = x \cdot x = x^T x = x_1^2 + x_2^2 + x_3^2,$$

where $x = (x_1, x_2, x_3)$ is a point of $\mathbb{R}^3$.

The condition that R preserve the inner product is $R^* R = 1$, where the adjoint of R is defined by $g(x, Ry) = g(R^* x, y)$ for all $x, y \in \mathbb{R}^3$. In matrix notation,

$$g(R^* x, y) = g(x, Ry) = x^T R y = x^T R^{TT} y = (R^T x)^T y = g(R^T x, y),$$

so $R^* = R^T$ and the inner product is preserved if $R^T R = 1$. The orientation is preserved if $\det R = 1$. Hence,

$$\mathrm{SO}(3) = \{R\colon \mathbb{R}^3 \to \mathbb{R}^3 \mid R^T R = 1 \quad \text{and} \quad \det R = 1\}.$$

27.1.2 The Lorentz group

The proper Lorentz group $\mathrm{SO}(3,1)$ is defined similarly, and consists of all linear transformations $\Lambda\colon \mathbb{R}^4 \to \mathbb{R}^4$ preserving preserving an orientation and the indefinite inner product

$$g(x,x) = x^T \eta x = x_0^2 - x_1^2 - x_2^2 - x_3^2,$$

where $\eta = \mathrm{diag}(1,-1,-1,-1)$. The invariance of the inner product is still equivalent to $\Lambda^*\Lambda = 1$ or, in matrix notation, $\Lambda^T \eta \Lambda = \eta$. Orientation-preserving transformations satisfy $\det \Lambda = 1$, so

$$\mathrm{SO}(3,1) = \{\Lambda\colon \mathbb{R}^4 \to \mathbb{R}^4 \mid \Lambda^T \eta \Lambda = \eta \quad \text{and} \quad \det \Lambda = 1\}.$$

We encounter a complication that was absent in the SO(3) case, since every element of SO(3) can be connected to the identity by a path staying in SO(3), but not so in $\mathrm{SO}(3,1)$. To see this, consider the transformation $\Lambda x = -x$ (interpreted as reversal of time and orientation simultaneosly), which cannot be obtained from the identity by a continuous transformation.

We don't want our symmetry groups to have disconnected components, so we define $\mathrm{SO}_0(3,1)$ to be the connected component of the identity in $\mathrm{SO}(3,1)$. We then have the inclusion

$$\mathrm{SO}(3) \hookrightarrow \mathrm{SO}_0(3,1)$$

$$R \longmapsto \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}$$

27.1.3 The unitary group

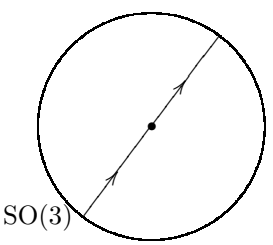
There is an amazing physical fact about electrons, and that is that they can distinguish a rotation by 360° from no rotation at all, but a rotation by 720° is indistinguishable from no rotation.

The coffee-cup trick Yadda, yadda, yadda.

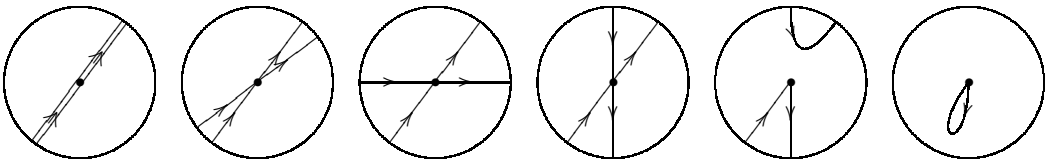
Topology of SO(3) An explanation of the coffee-cup trick or of the behaviour of electrons under 360° rotations is that there are certain special curves in SO(3) that join the identity to itself but cannot be contracted to a point.

We can find a convenient representation of the rotation group SO(3) in terms of axes and angles. It is well-known that every rotation fixes an axis in space. Every rotation can be described by a direction in space and an angle in $[0,\pi]$ of counterclockwise rotation about the given direction. All rotations of angle 0 are equivalent and represent the identity and, more importantly, rotations of π about opposite directions represent the same rotation.

Therefore, the rotation group can be represented by a solid 3-dimensional sphere of radius π, with opposite points of the boundary identified. The following curve represents all the rotations of angle 0 to 2π about a single axis.



Because of the identification of the opposite points of the boundary, it is impossible to "detach" the curve from the boundary by deformations, and therefore this curve is noncontractible. On the other hand, the sequence of rotations of angles from 0 to 4π is contractible, as shown by the following diagram.



The Pauli spin matrices When Pauli was trying to reconcile Quantum mechanics with the geometry of 3-dimensional space, he found the spin matrices associated to SU(2), the double cover of SO(3). Similarly, Dirac was led to the group $\mathrm{SL}(2,\mathbb{C})$, the double cover of $\mathrm{SO}_0(3,1)$, when discovering the Dirac equation. The relationship between the four groups is summarised in the following diagram:

$$\begin{array}{lcc} & \text{space} & \text{spacetime} \\ \text{quantum} & \mathrm{SU}(2) & \mathrm{SL}(2,\mathbb{C}) \\ & \downarrow 2{:}1 & \downarrow 2{:}1 \\ \text{classical} & \mathrm{SO}(3) & \mathrm{SO}_0(3,1) \end{array}$$

Now, SU(2) is the group of complex linear transformations $U\colon \mathbb{C}^2 \to \mathbb{C}^2$ with unit determinant and preserving the complex inner product

$$\langle z, z\rangle = z^\dagger \cdot z = \bar{z}^T z = \bar{z}_1 z_1 + \bar{z}_2 z_2 + \bar{z}_3 z_3,$$

where $\bar{z}$ denotes the complex conjugate, and $z^\dagger$ the conjugate transpose, of z. That U preserves the inner product is equivalent to $U^* U = 1$ or, in matrix notation, $U^\dagger U = \bar{U}^T U = 1$. In other words,

$$\mathrm{SU}(2) = \{U\colon \mathbb{C}^2 \to \mathbb{C}^2 \mid U^\dagger U = 1 \quad \text{and} \quad \det U = 1\}.$$

To understand how SU(2) can represent rotations, we have to use the fact that $\mathbb{R}^3$ with its usual inner product is isomorphic to 2×2 complex, hermitian, traceless matrices:

$$\begin{array}{ccc} (x_1, x_2, x_3) & \longrightarrow & \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix} \\ \downarrow \cdot & \swarrow -\det & \\ x_1^2 + x_2^2 + x_3^2 & & \end{array}$$

We denote

$$H_0 = \left\{ \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_3 \end{pmatrix} \colon (x_1, x_2, x_3) \in \mathbb{R}^3 \right\}$$

(H is for hermitian and 0 for traceless). We now need an action of SU(2) on H_0 preserving the determinant and the trace of elements of H_0. Such an action is given by

$$g\colon X \mapsto gXg^{-1}$$

where $X = \left(\begin{smallmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_3 \end{smallmatrix}\right)$ is the element of H_0 associated to $x = (x_1, x_2, x_3)$. To show that this is the appropriate action of SU(2) on H_0, we calculate the adjoint, trace and determinant of gXg^{-1}:

- $\det(gXg^{-1}) = \det g \det X \det g^{-1} = \det X$;
- $\mathrm{Tr}(gXg^{-1}) = \mathrm{Tr}(Xg^{-1}g) = \mathrm{Tr}X$; and

- $(gXg^{-1})^\dagger = (gXg^\dagger)^\dagger = g^{\dagger\dagger}X^\dagger g^\dagger = gXg^\dagger$, where we have used the fact that X is Hermitian ($X^\dagger = X$) and g is unitary ($g^\dagger = g^{-1}$).

There is just one detail left, and that is that there are two elements of SU(2) corresponding to each rotation.

Exercise 21 *Check that $gXg^\dagger = X$ has exactly two solutions in* SU(2)*, namely $g = \pm 1$, so that the map* $\mathrm{SU}(2) \to \mathrm{SO}(3)$ *is* 2:1*.*

27.1.4 The special linear group

We can illustrate the 2:1 map $\mathrm{SL}(2,\mathbb{C}) \to \mathrm{SO}_0(3,1)$ by very similar arguments to the above. In short, $\mathbb{R}^4$ with the Minkowski metric is isomorphic to the space of Hermitian 2×2 matrix. Indeed,

$$\det \begin{pmatrix} x_0 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{pmatrix} = x_0^2 - x_1^2 - x_2^2 - x_3^2.$$

Now, we can define an action of $SL(2,\mathbb{C})$ on the space H of Hermitian matrices by $X \mapsto gXg^\dagger$. The proof that this maps H to itself and preserves $\det X$ is the same as for $SU(2)$ acting on H_0. It is because there is no need to preserve the trace of X that the condition $g^\dagger = g^{-1}$ can be dropped.

Other dimensions This "trick" using complex matrices only works in 4-dimensional spacetime. It is a coincidence that Pauli and Dirac were inclined to use complex numbers because they wanted to interpret the components of spinors as complex probability amplitudes, and were thus led to these formulations in terms of small matrices. Similar constructions exist for the other real division algebras (the real numbers $\mathbb{R}$, quaternions $\mathbb{H}$ and octonions $\mathbb{O}$), according to the following table

Groups			Weyl spinors
$\mathrm{SL}(2,\mathbb{R})$	$\xrightarrow{2:1}$	$\mathrm{SO}_0(2,1)$	$\mathbb{R}^2$
$\mathrm{SL}(2,\mathbb{C})$	$\xrightarrow{2:1}$	$\mathrm{SO}_0(3,1)$	$\mathbb{C}^2$
$\mathrm{SL}(2,\mathbb{H})$	$\xrightarrow{2:1}$	$\mathrm{SO}_0(5,1)$	$\mathbb{H}^2$
$\mathrm{SL}(2,\mathbb{O})$	$\xrightarrow{2:1}$	$\mathrm{SO}_0(9,1)$	$\mathbb{O}^2$

Chapter 28

Building Spacetime from Spin(II)

We have seen that $\mathrm{SL}(2,\mathbb{C})$ acts on

- $\mathbb{C}^2$, the space of "Weyl spinors" via
$$g\colon \psi \mapsto g\psi, \quad \psi \in \mathbb{C}^2;$$
and
- $H = \{2\times 2 \quad \text{hermitian matrices}\}$ via
$$g\colon X \mapsto gXg^\dagger, \quad X \in H.$$

The space H is another way to look at Minkowski spacetime because the action of $\mathrm{SL}(2,\mathbb{C})$ on H preserves $\det X$ and, if we consider the basis

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

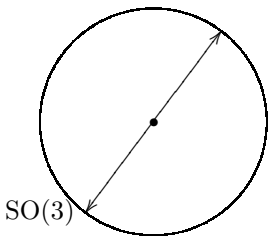
so that all $X \in H$ are of the form $X = x_0\sigma_0 + x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3$, then

$$\det X = x_0^2 - x_1^2 - x_2^2 - x_3^2$$

is the metric on Minkowski space. In this way we get a 2:1 and onto homomorphism $p\colon \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SO}_0(3,1)$.

Also, $H_0 = \{X \in H \mid \mathrm{Tr}X = 0\}$ is isomorphic to 3-dimensional Euclidean space and $\mathrm{SU}(2) \subseteq \mathrm{SL}(2,\mathbb{C})$ is the subgroup mapping H_0 to itself.

We studied the topology of $\mathrm{SO}(3)$ and saw that one good way to visualize this group is as a sphere of radius π with opposite points of the boundary identified.



In this way, we see that $\mathrm{SO}(3) \simeq \mathbb{R}P^3$, the three-dimensional real projective space. We have the following 2:1 maps

$$\begin{array}{ccc} \mathrm{SU}(2) & & S^3 \\ \downarrow{\scriptstyle 2:1} & & \downarrow{\scriptstyle 2:1} \\ \mathrm{SO}(3) & \xrightarrow{\sim} & \mathbb{R}P^3 \end{array}$$

This suggest that we should try to visualize $\mathrm{SU}(2)$ as a 3-dimensional sphere, and one way to do this is to realize that $\mathrm{SU}(2)$ is the space of unit quaternions!

Exercise 22 *Check that*

$$\mathrm{SU}(2) = \{x \in \mathrm{SL}(2,\mathbb{C}) \mid x^\dagger x = 1, \det x = 1\}$$

is equal to the space of unit quaternions

$$\{x_0\sigma_0 + i(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3) \mid x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1\}.$$

A way to visualize the 2:1 map from SU(2) to SO(3) is to visualize S^3 as $\mathbb{R}^3$ with a single "point at infinity", and consider the orbits of

$$x \mapsto -\pi^2 \frac{x}{|x|^2}.$$

It is clear that each point of the interior of the sphere of radius π maps to one and only one exterior point (the origin mapping to the "point at infinity") and that opposite points of the surface map to ech other. This is the "antipodal map" on S^3 and identifying pairs of antipodal points leads to a the representation of $\mathbb{R}P^3$ as a ball in $\mathbb{R}^3$ with antipodal boundary points identified.

28.1 Understanding spinors

So, $\mathrm{SL}(2,\mathbb{C})$ acts on $\mathbb{C}^2$ (spinors) and on H (Minkowski spacetime).

28.1.1 States of Weyl spinors

In quantum mechanics the space of states of a system is the space of unit vectors in a complex Hilbert space, "modulo phase" (i.e., up to multiplication by a complex scalar of norm 1). In our case, if we give Weyl spinors $\psi = \binom{\psi_1}{\psi_2}$ the Hilbert space norm $\|\psi\|^2 = |\psi_1|^2 + |\psi_2|^2$, unit spinors form a 3-dimensional sphere S^3. Hence,

$$\{\text{states of a Weyl spinor}\} \simeq S^3/\mathrm{U}(1).$$

28.1.2 The complex projective line

Each 1-dimensional subspace of $\mathbb{C}^2$ can be associated uniquely to a unit vector modulo phase. In other words, each state of a Weyl spinor uniquely determines the 1-dimensional subspace of all spinors proportional to it. Weyl spinors can then be interpreted as homogeneous coordinates in the projective space $\mathbb{C}P^1$. We say that two nonzero Weyl spinors $\binom{\psi_1}{\psi_2}$ and $\binom{\phi_1}{\phi_2}$ represent the same state if, and only if, $\psi_1\phi_2 = \psi_2\phi_1$. This is an equivalence relation.

28.1.3 The Riemann sphere

The complex projective line $\mathbb{C}P^1$ is well known to be isomorphic to $\mathbb{C}$ with a point at infinity, an extension of the complex numbers known as the Riemann sphere.

To see this, consider a 1-dimensional complex subspace of $\mathbb{C}^2$

$$\psi = \left\{\alpha\tbinom{\psi_1}{\phi_2} : \alpha \in \mathbb{C}\right\}.$$

If $\psi \neq 0$, we can write this as

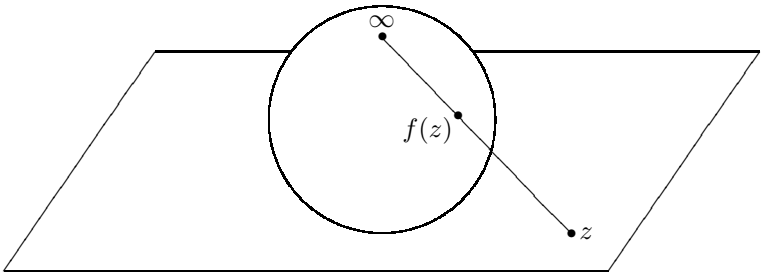
$$\psi = \left\{\alpha\tbinom{\psi_1/\psi_2}{1} : \alpha \in \mathbb{C}\right\},$$

and we can associate the complex number ψ_1/ψ_2 to ψ. Otherwise,

$$\psi = \left\{\alpha\tbinom{1}{0} : \alpha \in \mathbb{C}\right\},$$

corresponding to $\psi_1/\psi_2 = \infty$.

A plane with a single point at infinity is topologically equivalent to a sphere, as can be borne out by the following construction (the stereographic projection)



The Hopf fibration

We have seen that the states of a Weyl spinor can be interpreted as points on a 2-dimensional sphere, so we have $S^3/\mathrm{U}(1) \simeq S^2$. This is called the Hopf fibration

$$\begin{array}{ccc} S^1 & \hookrightarrow & S^3 \\ \downarrow & & \downarrow \\ \{*\} & \hookrightarrow & S^2 \end{array}$$

which shows that it is possible to fill the 3-dimensional sphere S^3 with an S^2 worth of circles.

Exercise 23 *Draw the Hopf fibration in* $\mathbb{R}^3 \cup \{\infty\}$.

28.1.4 Spin directions

Intuitively, the Riemann sphere is just the set of all directions in which a unit vector can point.

There are "angular momentum operators" acting on $\mathbb{C}^2$, qhich take the form

$$J_i = \frac{1}{2}\sigma_i \quad (i = 1, 2, 3)$$

The factor of $\frac{1}{2}$ is there so the commutation relations come out to be

$$[J_1, J_2] = J_3 \quad \text{(and cyclic permutations)}$$

Given a unit spinor ψ, we say that its expected angular momentum is the vector with components $\langle \psi \mid J_i \psi \rangle$. This is in no way special about spinors, it is just the way quantum mechanics works. Note that $\langle \psi \mid J_i \psi \rangle$ does not change if we multiply ψ by a phase, which explain why states are considered modulo phase.

Exercise 24 *Show that the length of the expected angular momentum vector is* $\sqrt{\frac{1}{2}(\frac{1}{2}+1)}$.

In this way we can identify unit spinors modulo phase with directions in space.

28.1.5 1-dimensional projectors in $\mathbb{C}^2$

Each point on the complex projective line represents a 1-dimensional subspace of $\mathbb{C}^2$. We can associate to each 1-dimensional subspace a projector, that is, a linear map $\rho\colon \mathbb{C}^2 \to \mathbb{C}^2$ such that $\rho^2 = \rho$ and $\rho = \rho^\dagger$. The fact that ρ projects onto a 1-dimensional subspace is equivalent to the condition $\mathrm{Tr}\rho = 1$.

Exercise 25 *Show that* $\rho\colon \mathbb{C}^2 \to \mathbb{C}^2$ *is a hermitian* 1*-dimensional projector if, and only if,* $\rho = |\psi\rangle \langle\psi|$ *for some state* ψ.

28.1.6 The celestial sphere

The equation $\rho^2 = \rho$ and $\rho \neq 0, 1$ implies that $\det \rho = 0$, so each 1-dimensional projector is a hermitian matrix with unit trace and vanishing determinant. In other words, it is an element of H (Minkowski space) such that $x_0^2 - x_1^2 - x_2^2 - x_3^2 = 0$, or a point of the light cone. The sphere of spinor states is a set of null vectors. The light cone is the union of the light rays through the origin, and we will see that in a precise sense the S^2 of spinor states is the sphere of all directions of light rays through the origin, also know as "the sky" or the "celestial sphere".

Chapter 29

Building Spacetime from Spin(III)

The following spaces are the same:

- states of a Weyl spinor;
- the complex projective line;
- the Riemann sphere;
- the angular momentum vectors of a spin-1/2 particle;
- the hermitian 1-dimensional projectors in $\mathbb{C}^2$;
- the celestial sphere.

Let us see in detail how the last two are isomorphic. Hermitian 1-dimensional projectors are linear maps $\rho\colon \mathbb{C}^2 \to \mathbb{C}^2$ such that

$$\rho = \rho^\dagger, \qquad \mathrm{Tr}\rho = 1, \qquad \text{and} \quad \rho^2 = \rho.$$

The equation $\mathrm{Tr}\rho = 1$ rules out the solutions $\rho = 0, 1$ of the equation $\rho^2 = \rho$. Therefore, the rank of ρ is 1 and $\det \rho = 0$. This means that hermitian projectors correspond to elements of H of the form $X = x^0\sigma_0 + x^1\sigma_1 + x^2\sigma^2 + x_3\sigma^3$, and since $1 = \mathrm{Tr}X = 2x^0$ we have $x^0 = \frac{1}{2}$. This is a circle that intersects each light ray through the origin precisely once, so we have a unique light ray associated to each projector.

Conversely, we can show that every $\rho = \rho^\dagger$ such that $\mathrm{Tr}\rho = 1$ and $\det \rho = 0$ is a 1-dimensional projector. Since $\det \rho = 0$ and $\mathrm{Tr}\rho = 1$, the rank of ρ is 1 and we can find vectors ϕ, ψ such that $\rho = |\phi\rangle \langle\psi|$ and $\langle\psi \mid \phi\rangle = \mathrm{Tr}\rho = 1$. But then $\rho^2 = \langle\psi \mid \phi\rangle \rho = \rho$ automatically, and $\rho = \rho^\dagger$ implies $\phi = e^{i\theta}\psi$.

Now, since the space of all light rays through the origin can be described in terms of H and det, the group $\mathrm{SL}(2, \mathbb{C})$ acts on it.

Exercise 26 *Show that if we represent points in the celestial sphere by the corresponding points on the Riemann sphere, then the action of* $\mathrm{SL}(2, \mathbb{C})$ *is*

$$\eta \mapsto \frac{a\eta + b}{c\eta + d} \qquad \textit{with} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C}).$$

In complex analysis one learns that these are all the conformal transformations of the Riemann sphere. So, if one accelerates to high speed, the constellations will appear distorted by an angle-preserving transformation.

Also, if we replace $\mathbb{C}$ by another normed division algebra, we get similar facts:

Groups			Weyl spinors
$\mathrm{SL}(2, \mathbb{R})$	$\xrightarrow{2:1}$	$\mathrm{SO}_0(2, 1)$	$\mathbb{R}^2$
$\mathrm{SL}(2, \mathbb{C})$	$\xrightarrow{2:1}$	$\mathrm{SO}_0(3, 1)$	$\mathbb{C}^2$
$\mathrm{SL}(2, \mathbb{H})$	$\xrightarrow{2:1}$	$\mathrm{SO}_0(5, 1)$	$\mathbb{H}^2$
$\mathrm{SL}(2, \mathbb{O})$	$\xrightarrow{2:1}$	$\mathrm{SO}_0(9, 1)$	$\mathbb{O}^2$

[Do this with a \tabular environment and the text "$\mathbb{C}$ Weyl spinors are a representation of $\mathrm{SL}(2,\mathbb{C})$, which is a double cover of $SO_0(3,1)$".]

There is a nice way to visualize the correspondence between the directions of space and the Riemann sphere. Represent each spinor $\binom{\psi_1}{\psi_2}$ by the complex number $\psi = \psi_2/\psi_1$. Then, the projection operator corresponding to $\psi \in \mathbb{C}$ is $\rho_\psi = \frac{1}{1+|\psi|^2}\begin{pmatrix} 1 & \psi^* \\ \psi & |\psi|^2 \end{pmatrix}$. This is obviously hermitian, and it is easy to see that this has unit trace and vanishing determinant. Also, in the limit $\psi \to \infty$ we get $\rho_\infty = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Now, observe that

$$\rho_\psi = \frac{1}{2}\sigma_0 + \frac{\mathrm{Re}\psi}{1+|\psi|^2}\sigma_1 + \frac{\mathrm{Im}\psi}{1+|\psi|^2}\sigma_2 + \frac{1-|\psi|^2}{2(1+|\psi|^2)}\sigma_3, \qquad \text{and} \quad \rho_\infty = \frac{1}{2}(\sigma_0 - \sigma_3),$$

so we get the correspondence

$$\begin{array}{cc} |\uparrow\rangle & 0 \\ |\downarrow\rangle & \infty \\ |\rightarrow\rangle & 1 \\ |\leftarrow\rangle & -1 \\ |\otimes\rangle & i \\ |\odot\rangle & -i \end{array}$$

[Draw a picture of the sphere.]

Similarly, if we did this for $\mathbb{H}$ we would get 5 orthogonal axes for the $\mathbb{H}\cup\{\infty\}$ 4-dimensional sphere, 9 orthogonal axes for the $\mathbb{O}\cup\{\infty\}$ 8-dimensional sphere and 2 orthogonal axes for the $\mathbb{R}\cup\{\infty\}$ 1-dimensional sphere.

Chapter 30

Feynman diagrams (I)

The platonic solids provide certain important finite subgroups of SO(3). The vertices of a platonic solid inscribed in the unit sphere are left invariant by subgroups of SO(3) isomorphic to the following groups:

tetrahedron	A_4
cube/octahedron	S_4
dodecahedron/icosahedron	A_5

Up to conjugation by an element of SO(3), these are almost all the finite subgroups of SO(3). We also have the finite cyclic groups $\mathbb{Z}_n$ and the dihedral groups D_n, of which $\mathbb{Z}_n$ is a subgroup of index 2 and which is the group of symmetries of an n-gon[1]. The Dihedral groups are subgroups of SO(3) because plane reflections can be implemented by a rotation in space.

Theorem 16 *Every finite subgroup of* SO(3) *is conjugate to one of* A_4, S_4, A_5, $\mathbb{Z}_n$ *or* D_n.

Theorem 17 *Every finite subgroup of* SO(2) *is* $\mathbb{Z}_n$ *for some* n.

Note that, since SO(2) is abelian, all conjugacy classes of subgroups consist of exactly one subgroup, so there is a notion of "the" $\mathbb{Z}_n$ in SO(2).

We could classify all subgroups of SO(4) easily. Apart from the two-dimensional groups D_n and $\mathbb{Z}_n$, we have the groups of rotations of the platonic solids, and the groups obtained by adding reflections to them (since a reflection in 3 dimensions can be implemented as a rotation in 4 dimensions). Then there are the groups of rotations of the 4-dimensional regular polytopes.

The 4-dimensional regular polytopes are the 4-simplex, the 4-cube and the 4-cross (analogous to the octahedron and dual to the 4-cube); and three "exotic" polytopes which can be obtained by realizing that the group SU(2) (the double cover of SO(3)) is isometric to the 3-sphere (on which SO(4) acts naturally) and so the groups of rotations of the platonic solids have a good chance of producing regular 4-dimensional polytopes. In fact, the tetrahedron and the dodecahedron give rise to such polytopes, and the one associated to the dodecahedron/icosahedron has a dual. Then we have the following

Theorem 18 *Every finite group is a subgroup of* SO(n) *for sufficiently large* n.

To prove this, consider the real vector space of formal linear combinations of elements of G, the group ring $\mathbb{R}[G]$. G has a natural linear action on this vector space, and the orientation-preserving subgroup of G is a subgroup of $\mathrm{SO}\big(|G|\big)$. Then, reflections can be implemented in one dimension higher, so G is a subgroup of $\mathrm{SO}\big(|G|+1\big)$.

An interesting consequence of this is that it is hopeless to try to obtain a general classification of the finite subgroups of SO(n) for all n, since a classification of all finite subgroups is all but impossible.

But, what does all this have to do with physics? There is in fact a long, illustious tradition of physical theories which state that the world is built out of the finite subgroups of SO(3).

[1]Note that we are using the convention that the subscript n indicates the order of the subgroup of rotations in D_n, rather than the more usual convention D_{2n} where the subscript denotes the order of the group.

The first such theory is due to the Pythagoreans and is known to us from Plato's Timæus. They thought that the platonic solids classified the elements!

tetrahedron	fire
hexahedron	earth
octahedron	air
dodecahedron	quintessence
icosahedron	water

The next attempt at using the platonic solids to explain the order of the world was Kepler's theory that the known planets were carried by concentric spheres around the Earth, nested in between platonic solids. In this way, the five platonic solids determined six radii. Despite the fact that there are $5! = 120$ possible arrangements, the astronomical observations at the time were accurate enough to make it impossible to fit the model to them, and Kepler was forced to discard it. He then moved on to the next simplest hypothesis, that the planets moved on ellipses, and this led to Newton's law of gravitation. At this point in history there was a shift from trying to explain the state of the universe (as Kepler was trying to do) to explaining its dynamical laws (as Newton did) and viewing the state as a historical accident.

Finally, and more to our modern taste, there's quantum mechanics where the classification of atoms involves homomorphisms not into SO(3) but out of it (or, more precisely, out of its double cover SU(2)):

$$\mathrm{SU}(2) \to \mathrm{GL}(V).$$

That is, the quantum mechanics of atoms is all about the linear representations of SU(2).

30.1 Group representations

Definition 18 *A homomorphism*

$$\rho\colon G \to \mathrm{GL}(V)$$

is called a representation of G on V

In Quantum Mechanics, states are described by unit vectors in a Hilbert space and symmetries are drscribed by unitary representations, that is, homomorphisms

$$\rho\colon G \to \mathrm{U}(H),$$

where

$$\mathrm{U}(H) = \{f\colon H \to H \mid f \text{ is linear and unitary}\}.$$

Usually, expositions of Quantum Field Theory begin with Classical Field Theory, and then proceed to quantization prescriptions, at which point it gets really messy. But in the end, if all works well, thetheory turns out to be very simple and beautiful. What people never tell you is that it is possible to get there really quickly, circumventing all the messy parts, without having to build a complicated scaffolding. Of course, when trying to make contact with experiment it is useful to have the scaffolding linking the quantum theory to a Classical Field Theory.

We are going to develop the theory of Feynman diagrams and we will see that we can get surprisingly far just by studying group representations.

Definition 19 *Given two representations (ρ, V) and (ρ', V') of G, an intertwining operator or intertwiner is a map $f\colon V \to V'$ such that*

$$\begin{array}{ccc} V & \xrightarrow{f} & V' \\ \downarrow{\scriptstyle \rho(g)} & & \downarrow{\scriptstyle \rho'(g)} \\ V & \xrightarrow{f} & V' \end{array}$$

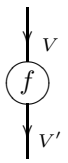
commutes for all $g \in G$.

Given this definition, it is not hard to see that we have a category where the objects are representations and the morphisms are intertwiners.

In quantum mechanics,

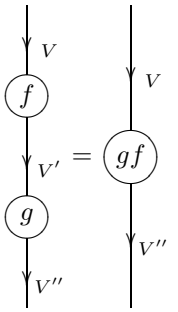
- Hilbert spaces are used to describe states,
- unitary representations describe how the symmetries of the physical system affect the states, and
- intertwiners describe processes (ways that states and systems can change) which are covariant (compatible with the symmetries).

We can draw intertwiners in this childishly simple way:



$f\colon V \to V'.$

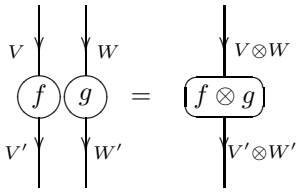
Given intertwiners $f\colon (\rho, V) \to (\rho', V')$ and $g\colon (\rho', V') \to (\rho'', V'')$, they can be composed as linear maps and a new intertwiner $gf\colon (\rho, V) \to (\rho'', V'')$ results:



$$\begin{array}{rccc} f\colon & V & \to & V' \\ g\colon & V' & \to & V'' \\ gf\colon & V & \to & V'' \end{array}$$

The key structure in a category, composition, is represented by the operation of stacking the diagrams for the intertwiners on top of each other.

We can also draw two intertwiners side by side:

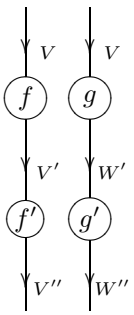


This corresponds to the operation of tensoring. Given two representations (ρ, V) and (σ, W), we get a new representation $(\rho \otimes \sigma, V \otimes W)$ defined by

$$(\rho \otimes \sigma)(g) = \rho(g) \otimes \sigma(g) \qquad \text{or, in other words,} \quad (\rho \otimes \sigma)(g)(v \otimes w) = \big(\rho(g)(v)\big) \otimes \big(\sigma(g)(w)\big).$$

Tensor representations are used in physics to describe the states of a composite system in terms of the states of its parts.

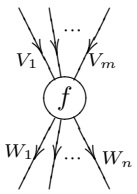
The motivation for the diagrammatic notation is that we can "let the pictures do the thinking". Take, for example, the diagram



$(f' \otimes g')(f \otimes g) = (f'f) \otimes (g'g).$

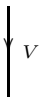
Not only is the diagram easier to understand and remember than the equation, it is also rather involved to prove that the equation is true algebraically compared to the simplicity of the corresponding diagrammatic manipulations. Of course, a number of identities like this have to be established before we can trust that the diagrams will give the right answer in any given manipulation.

We can also thing about intertwiners of the form



$$f\colon (\rho_1, V_1) \otimes \cdots \otimes (\rho_m, V_m) \to (\sigma_1, W_1) \otimes \cdots \otimes (\sigma_n, W_n),$$

We can hook these up to get more complex intertwiners. Also, we have the very simplest of all intertwiners:



$$\mathbf{1}_V\colon V \to V.$$

The Quantum field theory jargon is as follows:

- Representations are called "particles"; and
- Intertwiners are called "interactions".

The remaining ingredients that specify a "theory" are a symmetry group and the list of representations that represent physical particles.

Chapter 31

Feynman diagrams (II)

31.1 The category of representations

We have seen that for each group G there is a category of linear representations. In this category:

- Objects (ρ, V) are representations of G, i.e., group homomorphisms

  $$\rho\colon G \to \mathrm{GL}(V).$$

  For all $g \in G$, $\rho(g)\colon V \to V$ is a linear map, and $\rho(gh) = \rho(g)\rho(h)$ for all $g, h \in G$.

- Morphisms $f\colon (\rho, V) \to (\rho', V')$ are intertwiners, i.e., linear maps

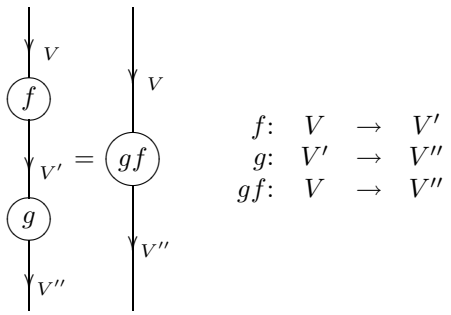
  $$f\colon V \to V'$$

  such that

  $$\begin{array}{ccc} V & \xrightarrow{f} & V' \\ \downarrow{\scriptstyle \rho(g)} & & \downarrow{\scriptstyle \rho'(g)} \\ V & \xrightarrow{f} & V' \end{array}$$

  commutes for all $g \in G$. That is, $\rho'(g)\big(f(v)\big) = f\big(\rho(g)(v)\big)$ for all $v \in V$ and $g \in G$.

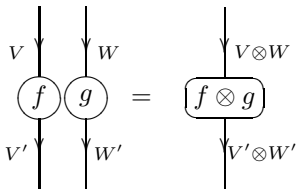
- We saw that a category has a 1-dimensional aspect, namely composition

  The properties of composition (associativity and identity morphisms) are automatic when composition is represented graphically and the indentity morphism on (ρ, V) is represented by

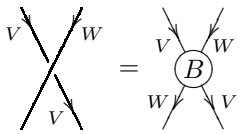
  $\mathbf{1}_V\colon V \to V.$

- In addition, the category of representations has a 2-dimensional aspect to it, given by the tensor product of representations. In technical terms, the category of representations is monoidal.



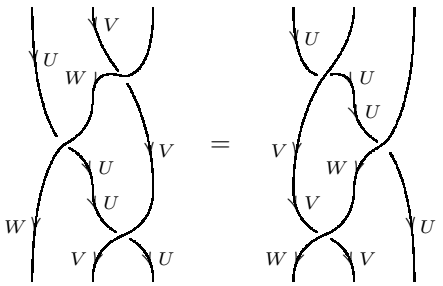
31.1.1 The braiding

The category of representations is what is called a "braided monoidal category", which gives it a three-dimensional aspect as well. Given two representations (ρ, V) and (ρ', V'), we get a special intertwiner (called the "braiding") $B_{V,W}: V \otimes W \to W \otimes V$:



$$B_{V,W}: \quad \begin{array}{ccc} V \otimes W & \to & W \otimes V \\ v \otimes w & \to & w \otimes v \end{array}$$

The braiding satisfies laws such as

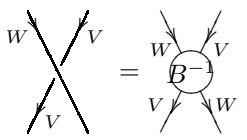


$$(\mathbf{1}_W \otimes B_{V,U})(B_{U,W} \otimes \mathbf{1}_V)(\mathbf{1}_U \otimes B_{V,W}) \quad = \quad (B_{W,V} \otimes \mathbf{1}_U)(\mathbf{1}_V \otimes B_{U,W})(B_{U,V} \otimes \mathbf{1}_W)$$

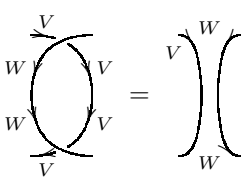
This equation is called the Yang-Baxter equation, and we write it only to illustrate the power of the diagrammatic notation.

In any case, this is yet another illustration that laws such as the Yang-Baxter equation are not only algebraic requirements, but topological as well.

The braiding is an isomorphism, but note that our diagrammatic notation makes it clear that the braiding is not required to be its own inverse. We denote

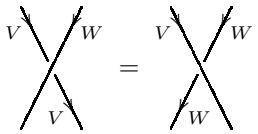


Indeed,



Symmetry

But the game of group representations has a 4-dimensional aspect, too. In 4-dimensional space, it is possible to pass crossing strings through each other!



A category in which the braiding satisfies this equation is called a symmetric monoidal category.

31.1.2 Duals and conjugates

This is great stuff, but group representations have other features.

The dual representation

Suppose we have a group representation $\rho\colon G \to \mathrm{GL}(V)$. Then we can get a representation ρ^* on the dual vector space V^*. Recall that

$$V^* := \{f\colon V \to \mathbb{C} \mid f \text{ linear}\}.$$

We need to use the representation $\rho\colon G \to \mathrm{GL}(V)$ to define the representation $\rho^*\colon G \to \mathrm{GL}(V^*)$. We try

$$\rho^*(g)(f)(v) = f\big(\rho(g)(v)\big).$$

We then have

$$\rho^*(gh)(f)(v) = f\big(\rho(gh)(v)\big) = f\Big(\rho(g)\big(\rho(h)(v)\big)\Big) = \rho^*(g)(f)\big(\rho(h)(v)\big) = \rho^*(h)\big(\rho^*(g)(f)\big)(v),$$

which means

$$\rho^*(gh) \neq \rho^*(g)\rho^*(h),$$

so this definition does not provide a representation.

To solve this problem, we need to look at the "dual pairing" or "counit"

$$\begin{array}{rccc} \varepsilon_V\colon & V^* \otimes V & \to & \mathbb{C} \\ & f \otimes v & \mapsto & f(v) \end{array}$$

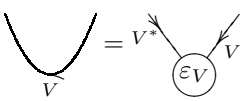
which must be an intertwiner from $\rho \otimes \rho^*$ into the trivial representation. This implies that

$$f(v) = \big(\rho^*(g) \otimes \rho(g)\big)(f \otimes v) = \rho^*(g)(f)\big(\rho(g)(v)\big), \qquad \text{or} \quad \rho^*(g)(f)(v) = f\big(\rho(g^{-1})(v)\big).$$

This is therefore the only sensible definition of ρ^*.

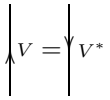
To draw the counit we need to realize that, since $\mathbb{C} \otimes \mathbb{C} = \mathbb{C}$, it makes sense to draw $\mathbb{C}$ as nothing at all rather than

. Then, we draw



$$\begin{array}{rccc} \varepsilon_V\colon & V^* \otimes V & \to & \mathbb{C} \\ & f \otimes v & \mapsto & f(v) \end{array}$$

the idea being that the "natural" picture on the left motivates the notation



Recall that in physics each arrow

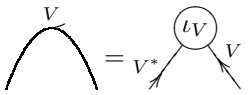
is called a "particle". Then,

corresponds to its "antiparticle" and

represents the physical process of "particle-antiparticle annihilation". We are actually neglecting conservation of energy, which explains why a particle-antiparticle pair can annihilate into nothing.

Come to think about it, when Feynman invented Feynman diagrams and proposed that antiparticles are particles moving backwards in time, he was doing nothing more than finding a graphical interpretation to the fact that, if the space of states of a particle is V, then V^* is the space of states of its antiparticle.

If V is finite-dimensional, we also get an intertwiner going the opposite way, called the "unit":

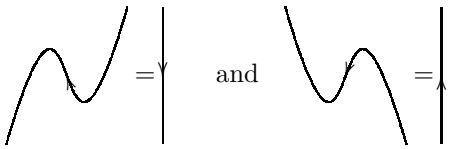


$$\begin{array}{rccc} \iota_V\colon & \mathbb{C} & \to & V \otimes V^* \\ & 1 & \mapsto & e_i \otimes e^i \end{array}$$

where $\{e_i\}$ is any basis of V and $\{e^i\}$ is its dual basis (defined by $e^i(e_j) = \delta_{ij}$). It is not entirely obvious that this definition is basis-independent, but the operator $e_i \otimes e^i$ is simply the identity in $\mathrm{End}(V)$. Indeed, if $v = v^i e_i$, then $\iota_V(v) = (e_i \otimes e^i)(v^j e_j) = e_i v^j \delta^i_j = v^i e_i = v$.

If V is infinite dimensional, we have that $V \otimes V^* \subseteq \hom(V, V)$, but $\mathbf{1}_V \notin V \otimes V^*$.

Now, combining the unit and counit, we can draw the following fun diagrams:



The conjugate representation

If we have a complex vector space V, there is a conjugate vector space $\bar{V}$, defined as follows:

- As a set, $\overline{V} = V$, but they have different vector space structures. To avoid confusion, we denote $\bar{v} \in \overline{V}$ for the vector $v \in V$, regarded as an element of $\overline{V}$ (you know you are in trouble when a mathematician says "regarded as").
- As far as addition is concerned, $\overline{V} = V$ too; that is,

$$\bar{v} + \bar{w} = \overline{v + w} \qquad \text{for all} \quad v, w \in V.$$

- Scalar multiplication involves conjugation in $\mathbb{C}$:

$$z\bar{v} = \overline{\bar{z}v} \qquad \text{for all} \quad v \in V, z \in \mathbb{C}.$$

  In particular, $i\bar{v} = \overline{-iv}$.

In this way V sprouts a whole family of related vector spaces: $V, V^*, \overline{V}, \overline{V}^*, \overline{V^*}, \ldots$. In fact, already $\overline{V^*}$ and $\overline{V}^*$ are naturally isomorphic, so there are only four different vector spaces.