Cartan Geometry and Spin Networks

submitted as a bachelor's thesis in physics by Manuel Bärenz

University of Heidelberg, Department of Physics and Astronomy

This bachelor's thesis has been carried out by Manuel Bärenz, born 2. 6. 1988 in Heidelberg, at the Ruprecht-Karls-Universität Heidelberg and the Friedrich-Alexander-Universität Erlangen-Nürnberg from 2. 5. 2011 to 13. 7. 2011
under supervision of Prof. Dr. Manfred Salmhofer (Heidelberg) with the assistance of Prof. Dr. Thomas Thiemann and Dr. Derek Wise (Erlangen).

Abstract

Spin network quantisation is a fundamental building block of loop quantum gravity, a background-independent approach to find a quantum theory of gravity. Here, we explore the possibility of applying this quantisation method to Cartan connections, which play a key role in an alternative formulation of gravity due to MacDowell and Mansouri.

In this way, we seek to relate these two approaches.

In order to make all of the mathematical tools accessible to non-experts, a self-contained introduction to principal connections, Cartan connections and their geometrical meaning is given.

Zusammenfassung

Spinnetzwerk-Quantisierung ist ein fundamentaler Baustein der Schleifen-Quantengravitation, ein hintergrundunabhängiger Ansatz um eine Quantentheorie der Gravitation zu finden. In dieser Arbeit wird die Möglichkeit untersucht, ob es möglich ist, diese Quantisierungsmethode auf Cartan-Zusammenhänge anzuwenden, die eine wichtige Rolle in einer Umformulierung der Gravitationstheorie von MacDowell und Mansouri spielen.

Auf diese Weise sollen beide Ansätze zusammengeführt werden.

Um alle mathematischen Werkzeuge für Nichtexperten nutzbar zu machen, ist eine Einführung in Zusammenhänge auf Hauptfaserbündeln und Cartan-Zusammenhänge enthalten.

Contents

Introduction

Motivation and Outline

Since Einstein's general theory of relativity, differential geometry is not only useful in physics, but essential. General relativity itself is a deep statement about the geometry of spacetime and its interplay with matter.

The Levi-Civita connection, which plays a central role in the general theory of relativity, is a prime example of a "principal connection". Surprisingly at first, gravitation is not the only area of physics where it is possible to define a theory with the language of differential geometry. For example Yang-Mills gauge theories, such as the theories of electroweak and of strong interaction, can be classically formulated with so-called principal connections, that underlie the bosons of the corresponding quantum field theories.

The introduction of Cartan connections to gravitational physics happened much more recently. They were discovered to be capable of a reformulation of MacDowell-Mansouri gravity, as discussed in [5]. They are mathematically very elegant: While the Levi-Civita-connection only contains "half" of the information of the gravitational field, the other half being the "vielbein", all of the information of the gravitational field is included in the Cartan connection: It unifies principal connection and vielbein.

Once the theory is quantised, it often loses its geometrical interpretation. Loop quantum gravity for example predicts a discrete, "combinatorial" geometry of spacetime. The reason lies in a specific quantisation process using important properties of the Lie groups of a principal connection.

Probably the most important problem of fundamental theoretical physics is finding a viable theory of quantum gravity. Loop quantum gravity is one approach to it that is built up in a mathematically rigorous way but did not yet enable anybody to calculate measurable predictions. It uses principal connections and triads (3-vielbeins) at the core of its definition. An obvious question would be if maybe the unification of those two quantities can be accomplished too, by means of Cartan connections.

In a recent article [2] about new aspects of loop quantum gravity, the notion of "projected spin networks" was introduced, where a construction that resembles certain features of Cartan geometry plays an important role.

Is it possible that Cartan connections are also a widespread concept in physics which we just were not aware of? Probably some more meaningful applications of them in physics are yet to be discovered.

In the first chapter, the two different notions of "connections" that play a role here are introduced: Principal connections and Cartan connections.

The method of how to build the Hilbert space of gauge invariant states in loop quantum gravity is explained in the second chapter. In the third chapter, the obstacles that have to be met if trying to repeat that method for Cartan connections are explained.

Who can read this paper?

There are many books and articles with lots of formulas and definitions. What is rare is one that satisfactorily explains the meaning behind them. Especially for beginners, it is extremely hard to work through the immense amount of new vocabulary used in research papers and at the same time thoroughly understanding it.

For a newcomer, it is not so hard to get a vague, intuitive understanding of the basic concepts which is sufficient to take part in small talk about the subject, but when asked to explain a concept rigorously or to calculate something explicitly, one quickly finds out not having really understood what he was talking about.

This thesis is intended to be different. I try to explain the basic definitions and *why* they were defined this way, often emphasizing the geometrical intuition.

Everything presented here should be understandable by a motivated bachelor student of theoretical physics who knows general relativity, tetrad formalism and the mathematical background thereof (smooth manifolds, vector fields, tensors, vielbeins). All other necessary concepts are explained in the appendices. Even for readers familiar with these concepts, it might be convenient to skim through them in order to be aquainted with the notations and conventions used in this thesis.

This thesis may not enable the reader to immediately start research on the mentioned fields. For this, one must to read further literature and finally papers by active researchers. But, I hope helps the reader to understand what the researchers are talking about.

Chapter 1

Connections

In what follows, we define principal connections and Cartan connections, emphasizing the geometric intuition.

A "trivial" remark

Readers that are unfamiliar with principal bundles, fibre bundles and their kind may safely skip ahead to the next section. For those who are aware of the notion of bundles I present an explanation why we do *not* need it here, and replace everything with trivial bundles such that we do not even need to consider them at all.

Since we are interested in their application to quantum field theory on a graph, with spin network quantisation as a special case, it is sufficient to define them on trivial bundles. The reason for this is that we will constantly deal with bundles over images of regular embedded curves. The image of such curves is diffeomorphic to the interval $[0, 1]$. And in fact fibre bundles over $[0, 1]$ are always trivialisable.

So if we are talking about fibre bundles over regular embedded curves, we can think of them as trivial bundles.

1.1 Principal connections

1.1.1 Geometric intuition

On an arbitrary manifold M, there is a priori no preferred way to compare tangent vectors, tensors or differential forms at *different* points.

This is demonstrated in the following example, where multiple different comparisons become possible. A connection is a way to eliminate this ambiguity.

Typically, we will have some sort of space at every point of the manifold, for example the tangent space, the space of tensors of some rank or the space of some forms, as previously mentioned. A connection enables one to compare the spaces at different points by mapping one onto the other isomorphically. The way this mapping works, contains information about the geometry of the connection.

The comparison of the two spaces may very well depend on the way we go from the first space to the second – it depends on the choice of curve that connects their base points.

The meaning of this is illustrated in figure 1.1. Suppose people on different parts on the earth construct tangent spaces and want to compare vectors in them. The method they choose is the following: The first person on the north pole picks a vector v and carries it to Congo, taking care not to turn it all the way except perpendicularly to the surface of the earth. The person in Congo receives the vector, calls it v' and carries it on to Indonesia where it is called v''.

Always referencing to their copy of the vector, all three people can compare their tangent spaces now. They have constructed a connection.

But the comparisons depend on the choice of curve the vectors were transported along: If the person on north pole wanted to send the vector to Indonesia on a great circle and have them receive the same vector as through the previous procedure, he would have to send v''', which is v rotated by 90 degrees. Equivalently, if we send v from the North pole to Congo, Indonesia and again to the North pole, it comes out rotated.

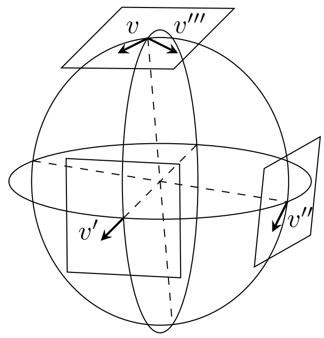

Figure 1.1: Parallel transport along a closed loop can turn around vectors.

So the tangent spaces of the earth surface get rotated when compared in this way. Generalising this to arbitrary objects (such as tensors or differential forms), we have to replace rotations by transformations acting on these objects. We will call the group of such transformations H in this section. This symmetry group, which one has to specify when talking about a specific connection, can be thought of as the generalisation of rotations to the specific space.

A connection has to provide these transformations: For every small step on the manifold, a small transformation has to be made. Infinitesimally, this means to assign an element of the Lie algebra of the transformation group to each tangent vector. Such an assignment is a Lie algebra valued 1-form, leading to...

...the definition

Definition 1.1.1. Let M be a smooth manifold and H a Lie group with Lie algebra $\mathfrak{h}$.
A **principal H-connection** is a $\mathfrak{h}$-valued 1-form A on M.

1.1.2 Properties

Lifts and holonomy

We already stated that tangent vectors and other objects get transformed when carried along a curve. The transformation that corresponds to some point of curve can be derived very simply: Intuitively, this is the transformation such that the connection is the "derivative" of it. The only thing we have to take care of is that the Lie algebra is the tangent space of the identity, not of an arbitrary Lie group element. So we have to insert the derivative of the transformation into the Maurer-Cartan-form.

The assignment of transformations, or elements of H, to every point of the curve such that their derivative as described above agrees with the connection, is called the lift of the curve.

Definition 1.1.2. Given an H-connection A on M and a curve $\gamma : [0,1] \to M$, the unique[1] map

[1] Since the equations form a sufficiently nice first-order differential equation with an initial value, the Picard-Lindelöf-Cauchy-Lipschitz-theorem asserts a unique solution.

$\hat{\gamma} : [0,1] \to G$ satisfying

$$\begin{aligned} \hat{\gamma}(0) &= \mathbb{1}_G \\ \omega_{MC}\left(\frac{\mathrm{d}\hat{\gamma}(t)}{\mathrm{d}t}\right) &= A\left(\frac{\mathrm{d}\gamma(t)}{\mathrm{d}t}\right) \end{aligned} \tag{1.1}$$

is called the **lift** of γ with respect to A.

Definition 1.1.3. Let $\hat{\gamma}$ be the lift of a curve γ with respect to a connection A. Then $\mathcal{H}_A[\gamma] := \hat{\gamma}(1)$ is called the **holonomy** of γ with respect to A.

Without a connection, one couldn't compare spaces at different points.

With a connection, one can if there is an action of the group H on these spaces: One first chooses a path from the first base point to the second. This gives the holonomy. The holonomy can now act on the objects in the space at the first base point and interpret the transformed objects as elements of the space of the other base point.

Gauging properties

One may now object that, when comparing only the vectors at the beginning and the end of the curve γ, there are lots of different connections leading to the same holonomy, and thus to the same comparison of spaces.

These connections differ only by their lifts of the curve between the endpoints, by the values $\hat{\gamma}(t)$ for $0 < t < 1$. Since the comparison at the beginning and the end is what the persons at these points are interested in, all choices of $\hat{\gamma}(t)$ for $0 < t < 1$ do not matter to them: They have to be, in physicists' terms, gauge equivalent.

This objection is in fact correct and is clarified by introducing the right gauge transformations, which act on the connections.

Two such connections A and A' that lift γ to $\hat{\gamma}$ and $\hat{\gamma}'$, respectively, are related by a multiplication with a group element for every point on the curve, that multiplies $\hat{\gamma}_1$ onto $\hat{\gamma}_2$. This is why H is normally called **gauge group**.

We want to have the definition of a gauge transformation on the whole manifold, like the connection, leading to:

Definition 1.1.4. Let M be a smooth manifold and H be a Lie group. A local H-gauge transformation is a smooth map $h : M \to H$.

Now it has to be clarified how a gauge transformation changes a principal connection. There are abstract and beautiful ways, using the language of principal bundles, explaining why this is done in one specific way, but at this point we prefer a more concrete, but equally convincing explanation.

As mentioned before, a gauge transformation h has to relate two connections such that the lift $\hat{\gamma}$ of a curve with respect to the first connection is multiplied onto the second $\hat{\gamma}'$. Let us insert this reasoning into the defining equation (1.1) of the lift:

$$\begin{aligned} A'\left(\frac{\mathrm{d}\gamma(t)}{\mathrm{d}t}\right) &= \omega_{MC}\left(\frac{\mathrm{d}\hat{\gamma}(t)'}{\mathrm{d}t}\right) = \omega_{MC}\left(\frac{\mathrm{d}\left(\hat{\gamma}(t)\cdot h(\gamma(t))\right)}{\mathrm{d}t}\right) \\ &= \mathrm{d}L_{(\hat{\gamma}(t)\cdot h(\gamma(t)))^{-1}}\left(\frac{\mathrm{d}\left(\hat{\gamma}(t)\cdot h(\gamma(t))\right)}{\mathrm{d}t}\right) \end{aligned}$$

We will introduce the short hand notation $hv := \mathrm{d}L_h(v)$ for the derivative of left multiplication, where v has to be a vector in the tangent space of the Lie group element h.

$$= (\hat{\gamma}(t) \cdot h(\gamma(t)))^{-1} \left(\frac{\mathrm{d}\,(\hat{\gamma}(t) \cdot h(\gamma(t)))}{\mathrm{d}t} \right)$$

Expanding the inversion, we get:

$$= h(\gamma(t))^{-1} \hat{\gamma}(t)^{-1} \left(\frac{\mathrm{d}\,(\hat{\gamma}(t) \cdot h(\gamma(t)))}{\mathrm{d}t} \right)$$

Let us use the product rule:

$$= h(\gamma(t))^{-1} \hat{\gamma}(t)^{-1} \left(\frac{\mathrm{d}\hat{\gamma}(t)}{\mathrm{d}t} \cdot h(\gamma(t)) \right) + h(\gamma(t))^{-1} \hat{\gamma}(t)^{-1} \left(\hat{\gamma}(t) \cdot \frac{\mathrm{d}h(\gamma(t))}{\mathrm{d}t} \right)$$

$$= h(\gamma(t))^{-1} \mathrm{d}L_{\hat{\gamma}(t)^{-1}} \left(\frac{\mathrm{d}\hat{\gamma}(t)}{\mathrm{d}t} \cdot h(\gamma(t)) \right) + h(\gamma(t))^{-1} \left(\frac{\mathrm{d}h(\gamma(t))}{\mathrm{d}t} \right)$$

Keep in mind that the derivative $\mathrm{d}L_{h^{-1}}$ of the left multiplication only agrees with the Maurer-Cartan-form ω_{MC} when it acts on vectors in the tangent space of h:

$$= h(\gamma(t))^{-1} \omega_{MC} \left(\frac{\mathrm{d}\hat{\gamma}(t)}{\mathrm{d}t} \right) \cdot h(\gamma(t)) + \omega_{MC} \left(\frac{\mathrm{d}h(\gamma(t))}{\mathrm{d}t} \right)$$

Inserting (1.1) again, we get:

$$A' \left(\frac{\mathrm{d}\gamma(t)}{\mathrm{d}t} \right) = h(\gamma(t))^{-1} A(\gamma(t)) h(\gamma(t)) + \omega_{MC} \left(\frac{\mathrm{d}h(\gamma(t))}{\mathrm{d}t} \right)$$

This has to hold for arbitrary curves. So we can rewrite it more generally and get the transformation rule of connections under gauge transformations:

$$A' = h^{-1} A h + \omega_{MC}(\mathrm{d}h)$$

Since the definition of the action of a gauge transformation on a connection was constructed like this, an almost tautological corollary follows:

Corrolary 1.1.5. If the holonomies of a curve γ with respect to two connections A and A' are the same, there is a unique gauge transformation that transforms A such that it agrees with A' on the image of the curve.

The proof is simple:
The curve is lifted twice, to $\hat{\gamma}$ and $\hat{\gamma}'$ by means of A and A', respectively. The required gauge transformation therefore is $h(t) = \hat{\gamma}(t)^{-1} \cdot \hat{\gamma}'(t)$.

Covariant derivative

For every connection, we also have what is called a **covariant derivative**.

Remember that the connection gives a comparison of spaces at different points of the manifold.

If one now applies a gauge transformation everywhere, then it will act on all the objects at either point, but it will also turn the connection such that the comparison given by the holonomy stays equivalent. So the comparison given by the connection is gauge covariant.

Looking at this infinitesimaly, one can define a differentiation that is also gauge covariant. Of course, we will only be able to differentiate something that takes values in a vector space, and be content with the derivative of a vector-valued differential form now.

The idea is going an infinitesimal step v along a curve γ starting at x, taking the value of the differential form there, transport it back to the starting point and substracting the value there.

In non-rigorous terms, the holonomy (which is a Lie group element) then is only infinitesimally different to unity, differing by a Lie algebra element to first order. Let us motivate the definition by a non-rigorous, intuitive calculation:

$$\begin{aligned}
\mathrm{d}_A\phi(v) &:= \lim_{t\to 0} \frac{\text{“}\mathcal{H}_A[\gamma]\phi(\gamma(t)) - \phi(x)\text{”}}{t} \\
&= \lim_{t\to 0} \frac{\text{“}(\mathbb{1} + tA(v))\phi(\gamma(t)) - \phi(x)\text{”}}{t} \\
&= \lim_{t\to 0} \frac{\text{“}\phi(\gamma(t)) - \phi(x)\text{”}}{t} + \text{“}A(v)\phi(\gamma(t))\text{”} \\
&= \text{“}v(\phi) + A(v)\phi\text{”} \\
&= \text{“}\mathrm{d}\phi(v) + A(v)\phi\text{”}
\end{aligned}$$

In rigorous terms, the covariant derivative is defined as:

Definition 1.1.6. Let ϕ be a V-valued differential n-form and A be a H-connection such that there is a representation R of H on V.
As explained in the appendix A.4, there is a representation of the Lie algebra $\mathfrak{h}$ on V, which we will also call R.
The **covariant derivative** of ϕ is a V-valued differential $n+1$-form defined by:

$$\mathrm{d}_A\phi := \mathrm{d}\phi + A \wedge_R \phi$$

As one can explicitly calculate, it satisfies the following formula when being gauge transformed:

$$\mathrm{d}_{A'}(h\phi) = h\,(\mathrm{d}_A\phi)$$

This is the reason why it is called the “covariant derivative”: Under a gauge transformation of the connection A and the form ϕ that is being derivated, the derivative is being gauge transformed as a whole.

Field strength is curvature

One might still doubt the usefulness of connections if nearly all of their information on a curve can be gauged away (it vanishes when applying a specific gauge transformation). But there is in fact information in the connection that does not vanish under arbitrary gauge transformations: The so-called **curvature**.

It can be defined by applying the covariant derivative two times:

$$\begin{aligned}
\mathrm{d}_A\mathrm{d}_A\phi &= \mathrm{d}_A\,(\mathrm{d}\phi + A \wedge_R \phi) \\
&= \mathrm{d}\mathrm{d}\phi + \mathrm{d}\,(A \wedge_R \phi) + A \wedge_{R,\circ} (\mathrm{d}\phi) + A \wedge_{R,\circ} (A \wedge_R \phi) \qquad (1.2)
\end{aligned}$$

We can use $\mathrm{d}\mathrm{d}\phi = 0$ for arbitrary forms, and the Leibniz rule:

$$= \mathrm{d}A \wedge_R \phi - A \wedge_R (\mathrm{d}\phi) + A \wedge_R (\mathrm{d}\phi) + A \wedge_R (A \wedge_R \phi)$$

With $\wedge_{R,\circ}$, we will now denote the wedge product of $\mathfrak{h}$-valued forms, taking as underlying multiplication the composition $\circ$ of homomorphisms in the representation R:

$$\begin{aligned}
&= (\mathrm{d}A + A \wedge_{R,\circ} A) \wedge_R \phi \\
&=: F \wedge \phi
\end{aligned}$$

As explained in appendix B.1, the 2-form F can thus be expressed as:

$$F = \mathrm{d}A + A \wedge_{R,\circ} A = \mathrm{d} + \frac{1}{2}[A \wedge A]$$

From the definition (1.2) of the curvature, we can deduce its transformation rule under gauge transformations:

$$\begin{aligned} F' \wedge_R \phi &= \mathrm{d}_{A'}\mathrm{d}_{A'}\phi \\ &= \mathrm{d}_{A'}\mathrm{d}_{A'}(hh^{-1}\phi) \\ &= \mathrm{d}_{A'}h\mathrm{d}_A(h^{-1}\phi) \\ &= h\mathrm{d}_A\mathrm{d}_A(h^{-1}\phi) \\ &= (hFh^{-1}) \wedge_R \phi \end{aligned}$$

If this is true for all ϕ and R is injective, we can conclude:

$$F' = hFh^{-1}$$

If $F = 0$, the connection is called "flat". The reason of this being a good nomenclature is given below. If $F \neq 0$, there is certainly no gauge transformations such that $F = 0$, so being flat or having nonvanishing curvature is an intrinsic information of the connection that cannot be altered by gauge transformations.

The reader will most probably already be familiar (maybe without having noticed) with two examples of curvature.

The field strength of classical electromagnetism is the first example. The potential A of electromagnetism in fact is a principal $U(1)$-connection. Since $\mathfrak{u}(1)$ is isomorphic (as a vector space) to the real numbers, having a vanishing Lie bracket, the field strength is a real-valued 2-form given by $\mathrm{d}A$.

The Riemann tensor of general relativity as well is a curvature 2-form – the curvature of the $SO(3,1)$ - Levi-Civita connection. One may at first object this and insist on the Riemann tensor being a tensor with 4 indices, but in fact the first two indices are just matrix indices that show up when looking at $\mathfrak{so}(3,1)$ in the fundamental representation and the second pair of indices just comes from the Riemann tensor being a 2-form. This also explains why the nomenclature of "flatness" was adopted for curvatures of principal connections: Vanishing curvature means a flat geometry.

1.2 Cartan geometry

The explicit use of Cartan geometry in gravity was explained by Derek Wise. His article [5] is an accessible overview of Cartan geometries and the MacDowell-Mansouri-approach reformulated in the language of Cartan geometries. Most of the calculations and notations in this section are copied from there.

1.2.1 Geometric intuition

An ancient traveller is on a journey on the earth. He doesn't know about the earth approximately being a sphere, he only knows about small lengths, such as steps, and waymarkings.

Slowly, he develops the idea of a cartesian coordinate system and finds it very helpful: To his arbitrarily chosen favorite direction, he assigns the name "x", and to the one perpendicular to x, he assigns "y". He measures distances in steps and counts the number of steps in x- and y-direction and calls this construction "coordinate system".

On his journey, he meets a lot of other travellers, who do not agree with his choice of directions. Some simply chose another pair of directions, not necessarily perpendicular, some choose what they call "latitude and longitude", which our traveller finds incomprehensible at this point, some even take completely arbitrary directions. The other travellers mainly agree that the choice of our traveller is a smart choice (for example, the length of a straight path or angles between two straight paths can be calculated easily) and – not willing to give up their own coordinate system – want to be able at least to compare small lengths: They find out, that there is always a way to translate a small step expressed in some coordinate system into a reference step in our travellers system. This translation is called a **zweibein**[2].

Making larger and larger journeys, he experiences problems with the model of a cartesian earth surface. Sometimes when he goes in large circles or other ways with a large area as interior, the starting point is at a different place than he expected, and turned by an angle (which he calls the "holonomy"). It seems as if the earth turned under his feet!

Consequently, he assigns to each step he makes a small rotation of his cartesian model earth and sees that if he includes the right rotation in his calculations, he can predict positions and orientations of all places correctly. He gives a name to this assignment, **connection**.

These two objects, the zweibein and the connection, are necessary to describe the surface of the earth sufficiently. In our very special case it turns out that the traveller can find a connection with "minimal rotation", that is, vanishing torsion, but let us focus on the general case.

After long discussions with greek philosophers, our traveller is ultimately convinced that his model earth is not very realistic on a global scale: The surface of the earth resembles the surface of a sphere, much more than the cartesian plane. Therefore, the traveller takes a ball as model earth and tries to repeat his technique of assigning small rotations and translations to each step on the earth.

Rotations around the current location still work the same on the ball. We will call them "real rotations". The analog to translations on the cartesian model are also rotations, but rotations around an axis perpendicular to the radius of the current location. We will call them "generalised translations".

So the traveller generalises: To each step on the real earth, he needs to assign a small rotation of the model earth, around a specific axis through the center. (In an even more general formulation, he could assign a so-called "homogeneous transformation", in this case an isometry, to each step.)

The traveller is amazed by the success of this technique. Unlike with the euclidean model, he nearly never needs to apply real rotations but mainly generalised translations. Most of the time, the generalised translations have the same size. Only sometimes, when he encounters high mountains or valleys, he needs to correct with real rotations or with generalised translations of other sizes. So the use of the spherical model earth enables him to "detect" mountains and valleys – the deviations of the real earth to the model earth – simply by looking where he has to assign extraordinary rotations to his steps.

[2]This is not quite correct. In fact, the zweibein is the *rule* how to find a translation in the sense described for any coordinate system, and not a specific translation.

The definition

In analogy to principal connections[3], the definition is again motivated by assigning infinitesimal transformations to infinitesimal steps on the manifold – but this time the transformations are not generalisations of rotations, but of rotations and translations!

Therefore, there has to be a larger Lie group involved, which we will call G, with Lie algebra $\mathfrak{g}$.

But the argument why we needed gauge transformations is still valid: In order to end up correctly at some special place, it doesn't matter for the traveller if he chooses a different assignment of model earth transformations to his steps, as long as the resulting model earth alignment would show him the same position. So a gauge transformation may rotate the model earth around the position of the traveller, it may however not be a generalised translation which would change the position.

We will call this smaller group, that only contains transformations leaving the position fixed, H, as before.

Definition 1.2.1. Let M be a manifold and G be a Lie group that has a subgroup H, with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, such that $\dim G - \dim H = \dim M$.
A Cartan connection is a $\mathfrak{g}$-valued 1-form A such that A projected onto the subspace $\mathfrak{g}/\mathfrak{h}$ is a vector space isomorphism at each point.

1.2.2 Properties

Metric Cartan geometries

In the definition 1.2.1, the latter condition that A needs to be an isomorphism when projected onto $\mathfrak{g}/\mathfrak{h}$ stems from the fact that this projection is normally interpreted as a vielbein field, or soldering form (which is explained in B.3). It's purpose (providing a metric on M) can be fulfilled if there is a metric on $\mathfrak{g}/\mathfrak{h}$ and if the vielbein is of full rank. If it isn't, the metric on M will become degenerate, which physicists do not believe to be true in the real world.

The curvatures of a Cartan connection

There also is something like a covariant derivative for a Cartan connection, but its geometrical meaning will not be elaborated here. Still, we can define the curvature in the same way, which does have an easily accessible geometrical meaning.

Definition 1.2.2. Let A be a Cartan connection.
Its curvature is defined to be:

$$F := \mathrm{d}A + \frac{1}{2}[A \wedge A]$$

Since $\mathfrak{g}$ is isomorphic to $\mathfrak{h} \oplus \mathfrak{g}/\mathfrak{h}$ as vector spaces, we can decompose F into its $\mathfrak{h}$-part and its $\mathfrak{g}/\mathfrak{h}$-part.

The $\mathfrak{g}/\mathfrak{h}$-part of F will be written as T. It is called the "torsion" of the Cartan connection.

The $\mathfrak{h}$-part of F will be denoted as $\widehat{F}$. It measures the deviation of the geometry determined by A from the geometry of the "model geometry", as is shown in the following paragraph.

[3]As for principal connections, we will only look at the definition of a Cartan connection on a trivial bundle. The general case is covered in detail in [5], whereas we only need to understand it in the trivial case since we want to apply it to quantum field theory on a graph.

Finally, we can also take the $\mathfrak{h}$-part ω of A and define its curvature $R := \mathrm{d}_\omega + \frac{1}{2}[\omega \wedge \omega]$ in the case that ω is a principal connection[4].

All these quantities and their geometrical meaning will be examined exemplarily in section 1.3.1 for the case of de Sitter spacetime as "model geometry".

Model geometries: Homogeneous spaces

The traveller was depicted to have a model of the earth with him that was designed to fulfil two tasks: Its geometry should be very close to that of the actual earth, but at the same time it should be very symmetric to admit a large group of isometries to map it onto itself – these were called rotations and generalised translations.

This can be generalised to arbitrary **model geometries**. A model geometry can be intuitively thought of as a (Riemannian or Pseudo-Riemannian) manifold with a large group G of isometries, like a sphere that can be rotated around any axis going through the center or Minkowski spacetime which is symmetric under Lorentz transformations (boosts and rotations) and translations.

It is possible to construct the model geometry only with the two Lie groups G and H. This construction is called a "homogeneous space" and was formulated by Felix Klein in his "Erlangen program", an attempt to unify lots of different approaches to geometry in a language based on Lie groups.

One point[5] p on the model geometry can be mapped to any other point q by a transformation in G. In fact there will be lots of transformations that map p on q. How do they differ? Intuitively, they can only differ by "rotations" around q, which are in general transformations that map q onto itself.

All transformations that map q onto itself form a group, the so-called **stabiliser group** of q. The stabiliser group of every point is isomorphic to a group H, which is also often simply called "stabiliser group".

So given only the groups G and H, every point is described by an element of G up to transformations of H: It is the equivalence class of elements of G, where equivalence means being related by an H-transformation. In rigorous terms, this means:

Definition 1.2.3. Let G be a Lie group and H be a Lie subgroup. The **homogeneous space** is the coset space

$$G/H := \{gH : g \in G\}$$

The notation gH is shorthand for the equivalence class $\{gh : h \in H\}$.

Accordingly, what we call a G/H-Cartan geometry is a $\mathfrak{g}$-valued Cartan connection with the gauge group H acting on it.

There is a prototypical example of a G/H-Cartan geometry: The homogeneous space G/H itself. Since there is a natural projection map from G to G/H (mapping every group element g onto its equivalence class gH), we can interpret the Maurer-Cartan form of G (for more details, see appendix B.3) as a $\mathfrak{g}$-valued 1-form on G/H. Taking this differential form as the Cartan

[4] This is the case if the gauge transformation acting on the Cartan connection acts on ω like the gauge transformation of principal connections. This is satisfied when elements of $\mathfrak{g}$ that lie completely in $\mathfrak{g}/\mathfrak{h}$ are transformed into $\mathfrak{g}/\mathfrak{h}$ again.

[5] This argument can be done not only for points, but also for lines, planes and other "features", as is explained in the paper [5] by Derek Wise.

connection, we notice that it is flat by construction, since the Maurer-Cartan form satisfies the following structure equation:

$$F = \mathrm{d}\omega_{MC} + \frac{1}{2}[\omega_{MC} \wedge \omega_{MC}] = 0$$

So G/H as a Cartan geometry is flat. The meaning of a general Cartan connection having vanishing curvature at some point is having locally the geometry of the model geometry.

1.3 MacDowell-Mansouri gravity

Cosmological observations suggest that the world we are living in is more similar to a de Sitter spacetime than to flat Minkowski spacetime: The cosmological constant is positive.

But describing the whole universe as a de Sitter spacetime is not detailed enough: There are galaxies and other large amounts of matter that curve spacetime slightly differently everywhere, so it will deviate from the de Sitter model from place to place. This is the perfect situation for a Cartan geometric description.

MacDowell and Mansouri did this (for the de Sitter and also anti de Sitter cases) in 1977 without knowing that they had reinvented some special cases of Cartan connections. Let us revisit their approach using the language of Cartan geometry.

1.3.1 Symmetry groups of de Sitter spacetime

A **de Sitter spacetime** is a vacuum[6] solution to the Einstein equations with a positive cosmological constant. In Riemannian geometry, a manifold with constant positive curvature is a sphere. So de Sitter spacetime may be thought of as something like the Lorentzian analog of a sphere.

As a homogeneous space, an ordinary n-sphere S^n is the quotient space $SO(n+1)/SO(n)$. This is easy to see: Every point of the sphere stays fixed by an $SO(n)$-rotation around it. If we imagine the sphere embedded in $\mathbb{R}^{n+1}$, its isometry group[7] are the rotations around the origin, $SO(n+1)$.

So we have identified $G = SO(n+1)$ and $H = SO(n)$ as relevant groups for the ordinary sphere. It is tempting to guess that for n-dimensional de Sitter spacetime we should choose $G = SO(n,1)$ and $H = SO(n-1,1)$ as Lorentzian groups with the same dimensions like the previous ones. This is in fact the right choice, but it will not be proven here.

Curvature

We will again decompose the $\mathfrak{g}$-valued connection A into the $\mathfrak{h}$-part, ω, and the $\mathfrak{g}/\mathfrak{h}$-part, $\frac{1}{l}e$. l is a characteristic length scale of the particular de Sitter geometry. Just as a sphere has a radius, de Sitter spacetime also has a length scale.

Furthermore, we will denote the curvature $\mathrm{d}_\omega\omega + \frac{1}{2}[\omega,\omega]$ as R, since it is the ordinary Riemannian curvature. It has to be emphasized that this is not the same as the $\mathfrak{h}$-part of the curvature of A, which will be called $\widehat{F}$. The $\mathfrak{g}/\mathfrak{h}$-part of F will in turn be written as T. Let us have a look how a general element of $\mathfrak{g}$ looks like in the fundamental representation of $\mathfrak{g} = \mathfrak{so}(4,1)$ as a 5×5 - matrix X. The condition that X needs to fulfil is $Xg = -gX$, where g is the "Minkowski" metric of $\mathbb{R}^{4,1}$.

[6]This means setting the energy-momentum-tensor to zero.
[7]We will only take into account the part connected to the identity.

The components of X can be written as:

$$X = \begin{pmatrix} 0 & b^1 & b^2 & b^3 & p^0/l \\ b^1 & 0 & j^3 & -j^2 & p^1/l \\ b^2 & -j^3 & 0 & j^1 & p^2/l \\ b^3 & j^2 & -j^1 & 0 & p^3/l \\ p^0/l & -p^1/l & -p^2/l & -p^3/l & 0 \end{pmatrix}$$

The upper left 4×4 - part is contained in the subalgebra $\mathfrak{h} = \mathfrak{so}(3,1)$. The b^i generate boosts and the j^i generate rotations.

Now using index notation with $a, b, c \in \{0 \dots 3\}$ and $I, J, K \in \{0 \dots 4\}$, we can decompose the Cartan connection explicitly into its $\mathfrak{h}$- and $\mathfrak{g}/\mathfrak{h}$-parts:

$$\begin{aligned} A^a{}_b &= \omega^a{}_b \\ A^a{}_4 &= \frac{1}{l} e^a \end{aligned}$$

Lowering a with the Minkowski metric of $\mathbb{R}^{3,1}$ on the right side and comparing with the general form of an element of $\mathfrak{g}$, we can deduce:

$$A^4{}_b = -\frac{1}{l} e_b$$

Let us calculate the curvature F of the connection and decompose it into the $\mathfrak{h}$-part $\widehat{F}^a{}_b = F^a{}_b$ and the $\mathfrak{g}/\mathfrak{h}$-part $T^a = F^a{}_4$:

$$\begin{aligned} F^a{}_b &= \mathrm{d}A^a{}_b + A^a{}_I \wedge A^I{}_b \\ &= \mathrm{d}A^a{}_b + A^a{}_c \wedge A^c{}_b + A^a{}_4 \wedge A^4{}_b \\ &= \mathrm{d}A^a{}_b + \omega^a{}_c \wedge \omega^c{}_b - \frac{1}{l^2} e^a \wedge e_b \\ &= R^a{}_b - \frac{1}{l^2} e^a \wedge e_b \\ F^a{}_4 &= \mathrm{d}A^a{}_4 + A^a{}_I \wedge A^I{}_4 \\ &= \mathrm{d}A^a{}_4 + A^a{}_b \wedge A^b{}_4 \\ &= \frac{1}{l} \left(\mathrm{d}e^a + \omega^a{}_b \wedge e^b \right) \\ &= \frac{1}{l} \left(\mathrm{d}_\omega e \right)^a \end{aligned}$$

To be able to formulate this in index-free notation, we need a bit more insight on the $e \wedge e$-term. First, we observe:

$$\begin{aligned} e^a \wedge e_b &= \frac{1}{2} \left(e^a \wedge e_b - e_b \wedge e^a \right) \\ &= \frac{1}{2} \left(e^a \wedge e^c - e^c \wedge e^a \right) \eta_{bc} \\ &= \left(e \wedge_\wedge e \right)^{ac} \eta_{bc} \\ &:= \left(M \left(e \wedge_\wedge e \right) \right)^a{}_b \\ &:= \left(e \wedge_{\wedge\eta} e \right)^a{}_b \end{aligned}$$

We used the antisymmetrisation defined in (4.2) and subsequently lowered an index with the ordinary Minkowski metric η. For details of the notation convention of the wedge product, see appendix B.1.

Note that $e \wedge_\wedge e$ is a $\Lambda^2\left(\mathbb{R}^{3,1}\right)$-valued form and $e \wedge_{\wedge\eta} e$ is an $\mathfrak{so}(3,1)$-valued form, so lowering an index with the Minkowski metric is an isomorphism M from $\Lambda^2\left(\mathbb{R}^{3,1}\right)$ to $\mathfrak{h}$, which will become important later on.

We can summarize all the calculation as:

$$F = \widehat{F} + T = \left(R - \frac{1}{l^2} e \wedge_{\wedge\eta} e\right) + \frac{1}{l} \mathrm{d}_\omega e$$

As promised, we can now explain what the geometrical meaning of these quantities is. For T, it is obvious: It is the torsion form known from ordinary general relativity in tetrad formulation.

But what is the meaning of $\widehat{F}$? This can be understood by asking a more special question: What does it mean to have a flat Cartan connection, that is, $F = 0$? Since $\widehat{F}$ and T reside in different subspaces, this leads to $T = 0$ and $\widehat{F} = 0$. The first equation says that the connection is torsion-free. The second equation is equivalent to $R = \frac{1}{l^2} e \wedge_{\wedge\eta} e$.

One can in fact check that the de Sitter solution satisfies this equation if we set the cosmological constant to $\Lambda = \frac{3}{l^2}$.

So a Cartan geometry with de Sitter spacetime as model space has vanishing curvature if its geometry is locally isometric to de Sitter spacetime. More generally, the quantity $\widehat{F}$ measures the deviation of the Riemann tensor R from the model de Sitter curvature.

Cartan geometry exactly accomplishes the task it was built for: If we had an overview of the Cartan curvature of the whole universe, it would be zero everywhere in the vacuum and nonzero where there are galaxies and other matter.

1.3.2 The MacDowell-Mansouri action

The interesting part about the MacDowell-Mansouri approach to gravity certainly is the action that leads to the Einstein equations. Up to now, the only thing we did was to reformulate the basic variables of general relativity with tetrads neatly with a Cartan connection. But MacDowell and Mansouri were able to rewrite also the action of general relativity in a way resembling Yang-Mills-theory strongly:

$$S_{MM} = \frac{-3}{2G\Lambda} \int \widehat{F} \wedge_K \star \widehat{F}$$

The notation of course has to be explained:

- Λ: We choose the natural $\Lambda = \frac{3}{l^2}$ found in the discussion of de Sitter spacetime.

- $\star$: Remember that there is an isomorphism $M : \Lambda^2\left(\mathbb{R}^{3,1}\right) \xrightarrow{\cong} \mathfrak{so}(3,1)$. Also be aware that there is the Hodge star isomorphism $* : \Lambda^p\left(\mathbb{R}^{3,1}\right) \to \Lambda^{4-p}\left(\mathbb{R}^{3,1}\right)$ for every $p \in \{0 \dots 4\}$ since $\mathbb{R}^{3,1}$ has a metric (the Minkowski metric) and it is oriented. In our case, this means that it is an automorphism of $\Lambda^2\left(\mathbb{R}^{3,1}\right)$. We can use this to define $\star : \mathfrak{so}(3,1) \xrightarrow{\cong} \mathfrak{so}(3,1)$: First, we use the inverse of M to get to $\Lambda^2\left(\mathbb{R}^{3,1}\right)$, use the Hodge star there and transport back to $\mathfrak{so}(3,1)$ with M. In other words:

  $$\star := M \circ * \circ M^{-1}$$

  It is often referred to as “internal hodge star”.

- $\wedge_K$: The multiplication this wedge product is based on is the Killing form of $\mathfrak{so}(3,1)$, as explained in appendix A.4. The Killing form maps two elements of the Lie algebra onto a real number, thus the integrand is a real valued differential form and qualifies as a Lagrangian density.

Let us find out why this is equivalent to general relativity.

$$
\begin{aligned}
S_{MM} &= \frac{-3}{2G\Lambda}\int \widehat{F}\wedge_K \star\widehat{F} \\
&= \frac{-l^2}{2G}\int\left(R-\frac{1}{l^2}e\wedge_{\wedge\eta}e\right)\wedge_K\star\left(R-\frac{1}{l^2}e\wedge_{\wedge\eta}e\right) \\
&= \frac{-l^2}{2G}\int\left(R-\frac{1}{l^2}e\wedge_{\wedge\eta}e\right)\wedge_K\star\left(R-\frac{1}{l^2}e\wedge_{\wedge\eta}e\right) \\
&= \frac{-l^2}{2G}\int R\wedge_K\star R-\frac{1}{l^2}(e\wedge_{\wedge\eta}e)\wedge_K\star R-R\wedge_K\star\frac{1}{l^2}(e\wedge_{\wedge\eta}e)+\frac{1}{l^2}(e\wedge_{\wedge\eta}e)\wedge_K\star\frac{1}{l^2}(e\wedge_{\wedge\eta}e)
\end{aligned}
$$

The term proportional to $R\wedge_K\star R$ can be neglected, since its variation is zero, which is a consequence of the second Bianchi identity $\mathrm{d}_\omega R=0$. Furthermore using the value $\frac{3}{l^2}$ of the cosmological constant, this simplifies to:

$$
= \frac{1}{2G}\int(e\wedge_{\wedge\eta}e)\wedge_K\star R+R\wedge_K\star(e\wedge_{\wedge\eta}e)-\frac{\Lambda}{3}(e\wedge_{\wedge\eta}e)\wedge_K\star(e\wedge_{\wedge\eta}e)
$$

Writing out everything in index notation of the fundamental representation and summing over equivalent terms finally yields a familiar result, the action of ordinary general relativity in tetrad formalism:

$$
= \frac{1}{2G}\int\varepsilon_{abcd}\left(e^a\wedge e^b\wedge R^{cd}-\frac{\Lambda}{6}e^a\wedge e^b\wedge e^c\wedge e^d\right)
$$

Thus general relativity and MacDowell-Mansouri gravity are equivalent as classical field theories. If there was a quantum theory of gravity or we would couple matter to it, the two theories maybe could be distinguished because of the topological term $R\wedge_K\star R$.

The MacDowell-Mansouri action is mathematically quite beautiful in its simplicity and its similarity to the Yang-Mills action, but there is one notable difference: The action is not built out of the full Cartan curvature, but only on its $\mathfrak{h}$-part $\widehat{F}$.

There are modifications of the theory where the full curvature is implemented and additional terms in the Lagrangian density constitute symmetry breaking from the large group G to the smaller group H. Manifestly and spontaneously symmetry breaking theories have been constructed.

Chapter 2

Spin network quantisation

2.1 Background independence of quantum field theories

"If quantum gravity is a quantum field theory, it is most probably background independent."

What is the meaning of a quantum field theory being background independent? There are different ways one could understand this term.

- The most common connotation of background independence is something like: "*A field theory is background independent if all relevant*[1] *fields are dynamical variables.*"

  For extra clarity, we will call this interpretation "background field independence". Prominent examples of background field *dependent* theories are Yang-Mills theories: Their Lagrangian density is given by $F \wedge_K *F$ with the curvature F and the Hodge star $*$. In the Hodge star, all the information about the metric (or tetrad) is contained. But the metric is the dynamical variable of general relativity, not of Yang-Mills theories, hence Yang-Mills theories are background field dependend. The metric is in fact a notorious background field, entering for example all field theories of the standard model.

- Another connotation is: "*A field theory is background independent if a solution (for the equation of motion of the dynamical fields) can be transformed by an active diffeomorphism and again be a solution.*"

  This property of a field theory is better called "diffeomorphism covariance". However, it almost implies background field independence. Imagine a theory with a background field coupled to the dynamical field and taking completely different values at for example two points p and q. Any solution configuration of the dynamical field will most probably be different at p and q. There certainly is a diffeomorphism that maps p onto q, which violates diffeomorphism covariance because solutions will differ in these points.

  Two ways of relaxing the conditions could be not coupling the background field or requiring it to be constant on all of space(time). The former would render it irrelevant, with the latter condition the field would be effectively the same thing as a constant of nature that doesn't need to be expressed as a field. So indeed all reasonable diffeomorphism covariant field theories are also background field independent.

[1] Of course we could just invent more fields that have nothing to do with the dynamical fields. Only those fields count as relevant that have some measurable impact on the dynamical variables.

Active and passive diffeomorphisms

We introduced the notion of an active diffeomorphism without defining it. An excellent discussion of that topic (and many other interesting philosophical discussions), is given in the book [3] by Carlo Rovelli.

"Passive diffeomorphisms" are better known under the name of "coordinate transformations". Basically all field theories we know admit passive diffeomorphisms: All quantities are simply expressed in the new coordinates. (The field equations do not necessarily keep their form under this transformation.)

"Active diffeomorphisms" however, leave the coordinate system fixed while the dynamical fields are pulled around by the diffeomorphisms. Theories that are covariant under this sort of transformations are very rare, general relativity is the most prominent example.

General relativity, as already mentioned, is background independent in either of the connotations. Underestimating the revolutionary aspect of this has lead to failed attempts to quantise it like an ordinary perturbative quantum field theory. If we take the core statement of general relativity serious, we should quantise it background independently.

There is no general recipe to do this – there are no numerous examples of background independent quantum field theories in text books. One specific approach to a background independent quantisation that in fact works is loop quantum gravity, and will be excerpted here.

2.1.1 Inapplicability of lattice gauge theory

In lattice gauge theory of Yang-Mills theories on Minkowski spacetime, a cubic lattice is choosen with a fixed distance of lattice points. The variables to quantise are the holonomies of the connection along the edges connecting neighbouring points.

However, this lattice approach will not work in quantum gravity for two reasons, which are precisely background field independence and diffeomorphism covariance.

First, in lattice theories, there is a specific distance of lattice points, which is crucial to the construction. Of course one needs to make use of a background metric to define it. Therefore, these theories are not background field independent.

Second, a diffeomorphism does not in general map lattice points on lattice points. Demanding this would be a severe restriction. Therefore, they're not diffeomorphism covariant either.

Additionally, generalising to arbitrary topologies of spacetime, it might not even be possible to construct a lattice.

2.2 Gauge theory on a graph

All these problems of lattice gauge theory can be solved by generalizing lattices to graphs.

The following discussion of gauge theory on a graph and spin network quantisation is based on John Baez' excellent article [1].

2.2.1 Embedded graphs

Definition 2.2.1. A finite graph (V, E) consists of a finite set of "vertices" V and a set of edges $E \subseteq V \times V$.

- We say that there is an edge from a vertex v_1 to v_2 iff $(v_1, v_2) \in E$.

- For an edge $e = (v_1, v_2) \in E$, we call the projection onto the first and the second item respectively the source $s(e) = v_1$ and the target $t(e) = v_2$ of the edge.

This definition did not make use of a manifold where the graph lives on, it is the definition of an "abstract graph". There is also the notion of an embedded graph.

Definition 2.2.2. A graph embedding of a graph (V, E) into a smooth manifold M is a map $F : E \to \{\gamma : [0,1] \to M\}$ with the following properties:

- F maps every edge onto a curve that is a smooth embedding of the interval $[0,1]$ into M.
- If, and only if two curves $F(e_1)$, $F(e_2)$ intersect somewhere, they do so at their starting points or end points while the source or target of e_1 and e_2 are the same.

Two graph embeddings F and F' are considered equivalent if $F'(e)$ is a reparametrisation of $F(e)$ for every edge e.

We will call the image of F a graph embedding. Clearly, it contains all the information about the underlying abstract graph.

The basic idea now is that we will try to approximate a gauge theory on M by describing it on embedded graphs.

2.2.2 Holonomies and their gauging properties

The dynamic field of a Yang-Mills theory is a principal connection, which has a local gauge group[2] G. This means that two connections that are related by a local gauge transformation should be regarded as physically equal and should be described by the same[3] state in the "gauge invariant Hilbert space" at the end.

We could in principle define it by defining the hilbert space over all principal connections and then somehow dividing by all the gauge transformations. But this construction might be hard to access because there is a vast amount of gauge transformations.

If we could get rid of the most of them before constructing a Hilbert space and only having to mod out a few, we might end up at a better construction[4] of the gauge invariant Hilbert space.

We may not forget that we only seek a description of the connection restricted to the embedded graph (the points corresponding to vertices and the images of the curves corresponding to the edges). It will turn out now, that it's convenient to first deal with gauge transformations on the (images of the curves of the) edges that leave the endpoints fixed and at the end take into account gauge transformations on the vertices. Together, they make up all gauge transformations.

Remember from 1.1.2, that the only information about a principal connection on a curve up to gauge transformations leaving the endpoints fixed is the holonomy. So the relevant information is completely given by the holonomy, an element of the gauge group. On each edge of the graph, we have a holonomy, so we can view the connection now as function $A : E \to G$, which is equivalent to writing:

$$A \in \mathcal{A} := G^E := \underbrace{G \times G \times \cdots \times G}_{\text{one for each element in } E} \tag{2.1}$$

We will still have to single out from the Hilbert space over $\mathcal{A}$ those few elements that are gauge invariant under the action of the group elements at the vertices.

[2]For the rest of this chapter, we will switch notation and call the gauge group G, and not H as in the previous chapter. This notation follows most literature.

[3]or at least proportional

[4]The space itself would of course be the same, but this way it will be understandable more easily, in this case by constructing an orthonormal basis.

2.2.3 Measure on the space of connections

Note that the space of connections now became significantly smaller than before. We started with the space of connections over a graph, an infinite dimensional space, and have now $\mathcal{A}$, a finite dimensional Lie group. It is straightforward to construct a Hilbert space over this configuration space since there already is a measure on Lie groups, the Haar measure.

The Haar measure

A complete and rigorous introduction into measure theory would be way beyond the scope of this thesis. Therefore only an intuitive overview of the vocabulary needed here will be given.

A measure μ on a space is a way to assign to a subset V of the space a volume $\mu(V) \in \mathbb{R}$. On the familiar space $\mathbb{R}^n$, one can for example define $\mu(V) := \int_V \mathrm{d}x^1 \mathrm{d}x^2 \cdots \mathrm{d}x^n$. In this case, one would also abbreviate $\mathrm{d}x^1 \mathrm{d}x^2 \cdots \mathrm{d}x^n$ simply by $\mathrm{d}\mu$, so $\mu(V) := \int_V \mathrm{d}\mu$.
On other spaces, in general one has to define the measure. But sometimes, there is a preferred measure, that for example respects symmetries of the space, in the sense of the following theorem:

Theorem 2.2.3. Let G be a Lie group. Then, up to a factor, there is a unique measure μ on G satisfying $\mu(gV) = \mu(V) \quad \forall g \in G$, where $gV := \{gg' : g' \in V\}$.
This measure is called the Haar measure.

Indeed, the measure in the example before was a Haar measure of $\mathbb{R}^n$ as a Lie group with the addition as group operation. The defining property can then be understood as translation invariance of the measure. For general Haar measures, the property means gauge invariance of the measure.

One might ask why the defining condition $\mu(gV) = \mu(V)$ was not written as $\mu(Vg) = \mu(V)$. This is also possible, yielding an equivalent definition of what is called a "right-invariant Haar measure" or just "right Haar measure". The right Haar measure does not necessarily have to be the same as the left Haar measure. But for semisimple and for compact Lie groups, they are the same. It turns out that these classes cover all the cases we are interested in.

Measure theory comes with a natural way to define a measure $\mu_{X\times Y}$ on a product space $X \times Y$, given measures μ_X and μ_Y on the two factors. Moreover, as suggested by the example, not only can a measure be defined from an integral, but it is also possible to define integrals from measures. This can be easily seen for example for positive real functions: Given a measure on X and a sufficiently nice function $f : X \to \mathbb{R}^+$, one defines the integral of f over X with respect to a measure μ_X as:

$$\int_X f \mathrm{d}\mu_X := \mu_{X\times\mathbb{R}}(\{(x,t) : x \in X, 0 \leq t \leq f(x)\})$$

This can be intuitively thought of as the volume under the graph of f.

For detailed and rigorous definitions, the reader is forwarded to literature on measure theory and Lebesgue integration.

The Peter-Weyl decomposition

With the Haar measure, we can define the space $L^2(G)$ of square integrable functions[5] over a Lie Group G, or in physical terms, the wave functions for a system that has G as configuration space.

[5]These are functions $f : G \to \mathbb{C}$ such that the integral $\int f(x) f(x)^* \mathrm{d}x$ makes sense.

At first, this space looks a bit unusable. A useful thing to have would be an orthonormal basis. In ordinary quantum mechanics, the knowledge of an orthonormal basis of the Hilbert space enables one to decompose vectors into basis elements and perform all calculations with their help. If an operator even is diagonalised on all the basis vectors, it is very easy to calculate its action on arbitrary vectors or expectation values of states.

In the following, we will derive an orthonormal basis of the space of connections on a graph discussed before which indeed diagonalises a lot of operators used in loop quantum gravity. However, the definition and discussion of these operators lies beyond the scope of this thesis.

Fortunately, there is a theorem that allows us decompose the Hilbert space of many Lie groups into pieces that make finding an orthonormal basis very easy.

Theorem 2.2.4 (Peter and Weyl[6], 1927)**.** Let G be a compact Lie group and Λ the set[7] of all irreducible unitary representations of G. Then:

$$L^2(G) \cong \bigoplus_{\rho\in\Lambda} X_\rho \otimes X_\rho^*$$

X_ρ denotes the vector space[8] that G acts on in the representation ρ.

Some more details about Lie groups and their representations are covered in appendix A.3.

Example 2.2.5. Consider the special case $U(1) = \left\{\mathrm{e}^{\mathrm{i}\phi} : \phi \in \mathbb{R}\right\}$.
The irreducible, unitary representations are $\rho_n, n \in \mathbb{N}_+$ where $\rho_n : U(1) \to \mathrm{Hom}(\mathbb{C})$ such that

$$\rho_n\left(\mathrm{e}^{\mathrm{i}\phi}\right)(x) = \mathrm{e}^{\mathrm{i}n\phi}(x)$$

So the spaces X_{ρ_n} are all $\mathbb{C}\otimes\mathbb{C}^* \cong \mathbb{C}$. This looks very similar to fourier decomposition. In fact, the Peter-Weyl-theorem generalises fourier decomposition (which is a feature of $U(1)$ only) to arbitrary compact Lie groups.

Labelling by representations

So the vector space $L^2(\mathcal{A})$ now can be given a simpler structure. Remember the definition of $\mathcal{A}$ in (2.1):

$$\begin{aligned}
L^2(\mathcal{A}) &= L^2(G \times G \times \cdots \times G) \\
&= L^2(G) \otimes L^2(G) \otimes \cdots \otimes L^2(G) \\
&= \bigotimes_{e\in E} L^2(G) \\
&= \bigotimes_{e\in E} \left(\bigoplus_{\rho\in\Lambda} X_\rho \otimes X_\rho^* \right) \\
&= \bigoplus_{\rho_E\in\Lambda^E} \bigotimes_{e\in E} X_{\rho_e} \otimes X_{\rho_e}^*
\end{aligned}$$

Every element of Λ^E is an assignment of an irreducible representation of G to each edge. Commonly, the elements ρ_E of Λ^E are called “labellings” of the edges by irreducible representations.

[6] The original theorem was much larger and more general. Nevertheless, let us focus on the special case relevant to us.

[7] We actually only include one representative for every equivalence class of representations.

[8] It doesn't matter which representative from the equivalence class we pick since they are all isomorphic.

2.2.4 Construction of an orthonormal base

Labelling by representations

An orthonormal basis to this space is quite easy to construct, given that the X_{ρ_e} have orthonormal bases:

Every space $H_{\rho_E} := \bigotimes_{e \in E} X_{\rho_e} \otimes X^*_{\rho_e}$ immediately gets an orthonormal base by definition of the tensor product.

The direct sum can be understood like this: We have $L^2(\mathcal{A}) = H_{\rho_E} \oplus H_{\rho'_E} \oplus H_{\rho''_E} \oplus \ldots$ which contains vectors of the form $(x, x', x'', \ldots)$. A natural orthonormal basis would be $\{(x, 0, 0, \ldots), (0, x', 0, \ldots), \ldots\}$, where $x, x', \ldots$ are members of orthonormal bases of the H_{ρ_E}.

So every basis vector of $L^2(\mathcal{A})$ can be expressed by a choice ρ of a labelling of the edges by irreducible representations of G and then choosing a basis vector for H_{ρ_E}.

Gauge symmetry leads to intertwiners

The spaces H_{ρ_E} are still very large. We can enhance our understanding of them by rewriting them in a more meaningful way and then using the remaining gauge freedom:

Lemma 2.2.6. There is a canonical isomorphism from $V \otimes W^*$ to the space $\mathrm{Hom}(W, V)$ of linear maps from W to V.

Proof for finite dimensional vector spaces:

W^* is the space of linear maps from W to the underlying field F (normally the real or the complex numbers). For a basis $\{w_1, w_2, \ldots, w_m\}$ of W there is a unique dual basis $\{w^1, w^2, \ldots, w^m\}$ of W^* with the property that $w^i(w_j) = \delta^i_j$.

With the help of a basis $\{v_1, v_2, \ldots, v_n\}$ of V, decompose a vector $u \in V \otimes W^*$ into basis elements with coefficients: $u = \sum_{i,j} A_{ij} v_i \otimes w^j$. With this choice of bases, every homomorphism can be expressed equivalently as an $n \times m$-matrix. Now assign to u the matrix A with components A_{ij}. This is clearly bijective and linear.

We can use this lemma in a meaningful way if we group the factors of the tensor product by vertices rather than by edges. For this, we introduce the notation $T(v)$ and $S(v)$ denoting all edges that have the vertex v as target or source, respectively.

Let us group the dual representations at the targets of the edges, and the other representations at the sources:

$$
\begin{aligned}
L^2(\mathcal{A}) &= \bigoplus_{\rho \in \Lambda^E} \bigotimes_{v \in V} \left(\bigotimes_{e \in S(v)} X_{\rho_e} \otimes \bigotimes_{e \in T(v)} X^*_{\rho_e} \right) \\
&= \bigoplus_{\rho \in \Lambda^E} \bigotimes_{v \in V} \mathrm{Hom} \left(\bigotimes_{e \in T(v)} X_{\rho_e}, \bigotimes_{e \in S(v)} X_{\rho_e} \right) \\
&=: \bigoplus_{\rho \in \Lambda^E} \bigotimes_{v \in V} \mathrm{Hom} \left(\mathcal{T}_v, \mathcal{S}_v \right)
\end{aligned}
$$

Note that T and S changed places in the notation in the first two lines.

Still, we did not make use of the gauge transformations at the vertices. At each vertex, we can gauge transform with an element of G, so a gauge transformation is a map from V to G, or equivalently an element of $\mathcal{G} := G^V$.

Its action on $\mathcal{A}$ preserves the previously defined measure: If we transform by g_1 at the source of an edge and by g_2 at the target, the holonomy at the edge transforms from g to $g_2 g g_1^{-1}$. Since we are considering only compact Lie groups, both the left and the right action preserve the measure on G. This is valid for all edges, so the representation of G^V on $L^2(\mathcal{A})$ is unitary. So, we are in fact looking for the space $L^2(\mathcal{A}/\mathcal{G})$. Physically, this means that a gauge transformation on the vertices does not have a measurable effect, as one would naturally require.

What condition does this requirement impose on the spaces $\mathrm{Hom}(\mathcal{T}_v, \mathcal{S}_v)$? A gauge transformation $g \in G$ at the vertex v acts on $\bigotimes_{e\in S(v)} X_{\rho_e} \otimes \bigotimes_{e\in T(v)} X^*_{\rho_e}$ with the automorphism $\bigotimes_{e\in S(v)} \rho_e(g) \otimes \bigotimes_{e\in T(v)} \rho^*_e(g)$.

So an element $f \in \mathrm{Hom}(\mathcal{T}_v, \mathcal{S}_v)$ needs to fulfill:

$$\begin{aligned} \bigotimes_{e\in S(v)} \rho_e(g)(f(x)) &= f\left(\bigotimes_{e\in T(v)} \rho_e(g)(x)\right) \\ \iff \bigotimes_{e\in S(v)} \rho_e(g) \circ f &= f \circ \bigotimes_{e\in T(v)} \rho_e(g) \end{aligned}$$

This condition is also known as **equivariance**. A linear map from representation spaces to representation spaces that is equivariant is also called **intertwiner**. The space of intertwiners at the vertex v with edges labelled by ρ will be written as $\mathrm{Inv}_v(\rho)$.

In conclusion, we can now say:

$$L^2(\mathcal{A}/\mathcal{G}) = \bigoplus_{\rho_E\in\Lambda^E} \mathrm{Inv}_v(\rho_E)$$

2.2.5 The result: "Spin network states"

Given an orthonormal basis $\{\iota_{v,1}(\rho_E), \iota_{v,2}(\rho_E), \ldots\}$ for every space $\mathrm{Inv}_v(\rho)$, then $L^2(\mathcal{A}/\mathcal{G})$ is orthonormally spanned by states

$$\Psi_{\rho_E,\iota_V} := \bigotimes_{v\in V} \iota_v(\rho_E)$$

where ρ_E is a labelling of the edges by irreducible unitary representations of G and ι_V is a choice of basis elements of the $\mathrm{Inv}_v(\rho)$.

These states are called **spin network states**.

2.3 Application to loop quantum gravity

It is time to justify spin network states by looking at their physical meaning in loop quantum gravity. There, they are supposed to be "quantum-geometry-states" of threedimensional space, as we will try to understand in a short overview.

First, general relativity in tetrad formalism is reformulated as a Hamiltonian theory. For this, a time variable is needed, so spacetime is split[9] up into a direct product of space and time. The

[9]This normally doesn't work for arbitrary topologies of spacetimes, but for a given spacelike hypersurface (a choice of 3-space at some specific time) there always is a neighbourhood around that hypersurface of the form of a direct product of space and a short time interval. So the Hamiltonian approach will work at least for small parts of 3-space and short time intervals.

conjugate variables are an $SU(2)$-connection A and the so-called “densitised triad” E.

First, why an $SU(2)$-connection? Naively, we would expect simply an $SO(3)$-connection as we have in threedimensional Riemannian differential geometry. But $\mathrm{Spin}(3) = SU(2)$ is the double cover of $SO(3)$, thus $\mathfrak{so}(3) \cong \mathfrak{su}(2)$ as Lie algebras, so there is no real change in the connection itself, just in the gauge group. Whether this preference of $SU(2)$ over $SO(3)$ does make sense, could in principle be decided by experiment. But already the physical existence of spinors that transform under $SU(2)$ and not $SO(3)$ forces one to interpret it as an $SU(2)$-connection.

Second, the densitised triad is in principle $E := e \wedge_\times e$ with the vector cross product $\times$ of $\mathbb{R}^3$. e is a $\mathbb{R}^3$-valued 1-form, hence the name “triad”. There can be some intricate issues about signs of E which will not be elaborated here. The important thing to note is that it contains all the metric information.

The classical quantities A and E are replaced by operators on the formal Hilbert space of $SU(2)$-connections[10] on 3-space. This Hilbert space is of course too bulky to be of any use, mainly because most of the states are gauge equivalent to some other states.

The trick where spin networks enter is the following: Instead of working on this large Hilbert space, one constructs a separable, accessible Hilbert space which is dense in the large one. This smaller Hilbert space is constructed as follows: The manifold is being triangulated (the details of the triangulation are topic of constant arguments) and a graph is being assigned to the triangulation by putting a vertex on each tetrahedron and an edge between two vertices if the corresponding tetrahedra are neighbours and share a face. So the edge “pierces” that face. The Hilbert space over $SU(2)$-connections on such a graph can then decomposed according to the method explained in section 2.2.3.

An interesting part is still how $L^2(SU(2))$ looks like, since it is one of the elementary building blocks of the spin network states. We already know that is decomposes into the target spaces of irreducible representations of $SU(2)$, but what are the irreducible representations of $SU(2)$?

The answer is known since a long time and is given by the quantum mechanics of spin angular momentum! Every irreducible representations corresponds to a way that an elementary particle can transform under transformations of $SU(2)$, according to what spin it has. So there is an irreducible representation for every spin quantum number $j \in \{0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots\}$. This finally solves the question why those states are called “spin network states” at all – they were originally considered in the special case of $SU(2)$ where the edges of the graphs are labelled by spins j, as we now know.

The analogy of A being the “coordinates” and E being the “momenta” goes further: Since the spin network states are the analog of “fourier modes”, one can ask if the “momenta” E are in some sense diagonalised by spin network states. Indeed, one can construct “area operators” from E, bound to every edge, which measure the area of the two-dimensional face that is pierced by the edge. It can be shown that these area operators are indeed diagonalised by the spin network states and have eigenvalues proportional to $\sqrt{j(j+1)}$, given that the spin j labels the edge. Other operators like volume operators can also be defined. Many other operators are built upon these area and volume operators.

[10] This is justified because A and E are conjugate variables. The space of A’s is interpreted to be analog to the coordinate space and the space of E’s to the momentum space. So the Hilbert space is constructed over “coordinate space”.

Chapter 3

Cartan networks

Up to this point, nothing of the material presented here is new.

Some concepts, like Cartan geometry, might be unknown to most readers, but still they are well covered in books and articles.

During May and June of 2011, my supervisor Derek Wise and I wanted to find out if there is any Cartan geometry "hidden" in loop quantum gravity. One motivation could be the following: There is a way to get the $SU(2)$-connection of loop quantum gravity by restricting the Lorentz connection of general relativity first to 3-space and then to values in $\mathfrak{so}(3) \cong \mathfrak{spin}(3) \cong \mathfrak{su}(2)$, letting the triad take the remaining degrees of freedom. Since the Levi-Civita connection takes values in the larger $\mathfrak{so}(3,1) \cong \mathfrak{spin}(3,1) \cong \mathfrak{sl}(2,\mathbb{C})$, this looks very similar to a $SO(3,1)/SO(3)$-Cartan connection.

One obvious thing to try is to find a way to repeat the spin network quantisation process for Cartan connections.

A G/H-Cartan connection can be viewed as a principal G-connection that has the additional constraint of being an isomorphism when projected to $\mathfrak{g}/\mathfrak{h}$, so in a sense a Cartan connection is a special case of a principal connection. However, in this view, the Cartan connection has a *smaller* gauge group H than the principal G-connection. This means that Cartan connections modulo gauge transformations are not at all a special case of principal connections modulo gauge transformations. So repeating spin network quantisation for Cartan connections cannot be accomplished by simply specialising the procedure.

Gauge invariant content of Cartan connections

In spin network quantisation, the holonomy of an edge of a graph is the only information about the connection that is gauge invariant under a gauge transformation leaving the endpoints fixed. This was essential when we reduced the large Hilbert space of connections to the better behaved and more physically relevant gauge invariant space.

In Cartan geometry, the situation is different: The connection is $\mathfrak{g}$-valued, but the allowed gauge transformations are only those in H, with H a subgroup of G. If we want to lift a curve with respect to the connection, we get an element of G, not of H. The space of connections is now significantly larger, and we cannot expect to have enough H-gauge transformations to transform a connection onto any other connection with the same holonomy.

One way to deal with this is trying to allow more gauge transformations. But there only is one other class of gauge transformations at hand: Diffeomorphisms. So, one could ask if, for two

Cartan connections on a curve with the same G-holonomy, there is a diffeomorphism combined with an H-transformation (or, more generally, a general diffeomorphism of the so-called "principal bundle") that maps one connection onto the other. The answer to this question has to be given in later researches.

The other approach is accepting that Cartan connections in our case contain more information than principal connections. Cartan connections combine a principal connection with a vielbein (e.g. a tetrad or a triad), and this vielbein seems just to be the information that is "too much", it cannot be gauged away on a curve. How did we actually deal with it in the loop quantum gravity approach? We declared the two quantities as...

... Conjugate variables

The densitised triad is the conjugate "momentum" of the principal $SU(2)$-connection of loop quantum gravity. Therefore there was no need to have a Hilbert space over the space of triads, but only over the space of connections.

So maybe trying to find a Hilbert space for all Cartan connections is the wrong question, and one should first try to identify canonically conjugate variables "within" the Cartan connection? This might be, but an answer to this question depends crucially on the Hamiltonian, which in turn could be built upon an action, but which one?

One way to shed light on this might be trying different actions or hamiltonians (which might be inspired by loop quantum gravity) and investigating the conjugate variables.

- **The MacDowell-Mansouri action:**
  From the viewpoint of an $SO(3,1)/SO(3)$-Cartan connection, it is tempting to repeat the MacDowell-Mansouri action: Since the field strength F is a $\mathfrak{so}(3,1)$-valued 2-form, we can even postulate a Lagrangian like $F \wedge_K \star F$ without breaking F down into $\widehat{F}$ and T. But varying this with respect to the Cartan connection A gives the field equation $\mathrm{d}_A F = 0$, which is automatically satisfied since it is the Bianchi identity. So it is a theory where every field configuration is a solution of the field equations – a physically uninteresting, trivial theory.

  In MacDowell-Mansouri gravity, the action did not lead to a trivial theory precisely because of the symmetry breaking of F to $\widehat{F}$. Varying the broken $\widehat{F} \wedge_K \widehat{F}$ with respect to the broken ω is not the same as this approach since $\widehat{F}$ is not the curvature of ω.

- **The loop quantum gravity Hamiltonian:**
  The most promising approach is probably to take the Hamiltonian that is employed in loop quantum gravity and try to reformulate it in terms of Cartan geometry. However, this has not been attempted yet.

Conclusion

The idea of "Cartan network states", an easy to handle Hilbert space of Cartan connections on a graph, is not a trivial thing to define and deserves more study. Cartan connections behave fundamentally differently from principal connections, despite their similarities. A quantum theory of Cartan connections inspired by spin network quantisation could however provide ways to reinterpret aspects of loop quantum gravity in new, surprising ways, maybe enhancing our understanding of them.

In this short time, we were unable to develop these ideas as far as we would have likes, but we understand the basic questions better and are in a good position for further investigation. Definitely, this area of research is very interesting and I will keep on following it beyond this thesis.

Acknowledgements

I want to thank Derek Wise for his neverending patience and uncountable discussions about the connections between connections and everything else.

Thanks to Thomas Thiemann and the Theorie-3-group for letting me join the fun,
to John Baez and Carlo Rovelli for enlightening discussions and having written great papers,
to Martin Bärenz and Raffaela Milner for helpful comments,
to Manfred Salmhofer and Jan Pawlowski for supporting me on this advanced topic where others wouldn't.

Chapter 4

Appendices

A Lie groups, Lie algebras and their representations

A.1 Lie groups

A Lie group unifies the properties of manifolds and groups and therefore is a very special object with rich properties.

Lie groups originated from the study of continuous symmetry transformations.

Definition A.1. Let $G = (M, *, \mathcal{A})$ be a tuple of a set M, a multiplication $*$ and a smooth atlas $\mathcal{A}$ such that:

- $(M, *)$ is a group.
- $(M, \mathcal{A})$ is a smooth manifold.
- $*$ and $\mathcal{A}$ are compatible. With this, we mean that $* : M \times M \to M$ is a smooth map regarding $\mathcal{A}$ and the product atlas of $M \times M$.

Then G is called a **Lie group**

Colloquially, the group and the underlying set often have the same letter.

Example A.2. $\mathbb{R}^n$ with addition operation $+$ can be given the structure of a Lie group:
Let $G = (\mathbb{R}^n, +, \mathrm{id})$. By the definition of vector spaces, $(\mathbb{R}^n, +)$ is a group. Trivially, $(\mathbb{R}^n, \mathrm{id})$ is a manifold. Finally, the map $+ : (x, y) \mapsto x + y$ is smooth since it is linear.

A.2 Lie algebras

Lie algebras were originally studied in order to simplify the classification of Lie groups. However, they have become a topic on their own and can be defined independently of Lie groups.

Definition A.3. Let $(\mathfrak{a}, [\cdot,\cdot] : \mathfrak{a} \times \mathfrak{a} \to \mathfrak{a})$ be an algebra. It is a **Lie algebra**, if the following two properties are satisfied for all elements $A, B, C \in \mathfrak{a}$:

$$\begin{aligned} [A, B] + [B, A] &= 0 \\ [[A, B], C] + [[B, C], A] + [[C, A], B] &= 0 \end{aligned}$$

The latter equation is called "Jacobi equation".

Intuitively, a Lie algebra element often can be thought of as a Lie group element very close to $\mathbb{1} \in G$. More precisely, it is a tangent vector to $\mathbb{1}$.

Definition A.4. It is possible to assign canonically to a Lie group G "its" Lie algebra $\mathfrak{g}$:
The underlying vector space is the tangent space $\mathrm{T}_{\mathbb{1}}G$ of the neutral element $\mathbb{1}$. Now let $\gamma_1(t)$ and $\gamma_2(t)$ be two smooth curves on G with $\gamma_1(0) = \gamma_2(0) = \mathbb{1}$ and $\dot{\gamma}_1(0) = A$, $\dot{\gamma}_2(0) = B$. Now define

$$[A, B] := \frac{\mathrm{d}}{\mathrm{d}\,t_1}\frac{\mathrm{d}}{\mathrm{d}\,t_2}\left(a(t_1)b(t_2)a(t_1)^{-1}b(t_2)^{-1}\right) \tag{4.1}$$

, which gives the structure of a Lie algebra to the tangent space: $\mathfrak{g} := (\mathrm{T}_{\mathbb{1}}G, [\cdot,\cdot])$.

A.3 Representations of Lie groups

Lie groups were discovered originally as matrix groups, that is, groups where each group element is a matrix. The motivation of this was (and still is) in a large part to have the group elements act on something, for example a rotation acting on a vector. If they should act on a vector space like $\mathbb{R}^n$ or $\mathbb{C}^n$, it is convenient to have them in matrix form.

Later, the notion of a Lie group was formalised and it turned out that is actually possible to write the group elements as matrices in different ways, depending on what sort of vectors they should act on.

Formalising the above idea yields the following definition:

Definition A.5. Let G be a group and V be a vector space.
A **group representation**[1] $R : G \to \mathrm{Aut}(V)$ of G on V is a group homomorphism from G to the automorphism group of V.

- The automorphism group $\mathrm{Aut}(V)$ of a vector space V is the group of all isomorphisms from V to itself. The group multiplication is the composition.
  Often, it is also called "general linear group".
- If V is supplied with a basis $\{v_1, v_2, \ldots v_n\}$, then $\mathrm{Aut}(V)$ can be thought of the group of invertible $n \times n$-matrices.
- R being a group homomorphism especially means for $g_1, g_2, \mathbb{1} \in G$, $f_1, f_2, \mathrm{id} \in \mathrm{Aut}(V)$:
  - $R(g_1 * g_2) = R(g_1) \circ R(g_2)$
  - $R(\mathbb{1}_G) = \mathrm{id}_V$
- Often, a group representation is also called a group action[2]. The group G acts on the vector space V through the representation R.
- The representation that maps every group element to the identity on the vector space is called "trivial representation".

Definition A.6. Let V be a Hilbert space with inner product $< .|. >$. A representation R of G on V is called **unitary** if the group action preserves the inner product:

$$\forall v_1, v_2 \in V, g \in G :$$
$$\langle v_1 |\, v_2\rangle = \langle R(g)(v_1)|\, R(g)(v_2)\rangle$$

[1] The German translation of "representation" is "Darstellung" and *not* "Repräsentation".
[2] Group actions are more general, they don't need V to be a vector space.

Definition A.7. Let R be a representation of G on V.
The **dual representation** R^* is a representation of G on the dual[3] vector space defined by:

$$R^*(g) := R(g^{-1})^{\mathrm{T}}$$

Note, that the dual representation acts by application from the right, and not from the left!

Definition A.8. Let R_1 and R_2 be two representations of G on V_1, V_2, respectively.
Then the tensor product $R_1 \otimes R_2$ of R_1 and R_2 is a representation of G on the tensor product[4] $V_1 \otimes V_2$ defined by:

$$(R_1 \otimes R_2)(g)(v_1 \otimes v_2) := R_1(g)(v_1) \otimes R_2(g)(v_2)$$

Irreducible representations

Definition A.9. Let R_1 and R_2 be two nontrivial representations of G on V_1, V_2, respectively. Then the direct sum $R_1 \oplus R_2$ of R_1 and R_2 is a representation of G on the direct sum[5] $V_1 \oplus V_2$. It is defined by:

$$(R_1 \oplus R_2)(g)(v_1 \oplus v_2) := R_1(g)(v_1) \oplus R_2(g)(v_2)$$

This can be thought of the two representations acting component-wise.

In matrix form, this means:

$$R(g) = \begin{pmatrix} R_1(g) & 0 \\ 0 & R_2(g) \end{pmatrix}$$

$R_1 \oplus R_2$ is then called **reducible**.

Definition A.10. An **irreducible representation** of G is a nontrivial representation that is not reducible i.e. it *cannot* be expressed as $R_1 \oplus R_2$, except for trivial R_1 or R_2.

Typically, we will denote irreducible representations by ρ.
Note the similarity of this definition to the definition of prime numbers. In fact, irreducible representations can be seen as the "prime numbers of representation theory", the fundamental building blocks.

A.4 Representations of Lie algebras

For a representation R of a Lie group G, there is also a representation of its lie algebra $\mathfrak{g}$ which we denote by R as well. It is defined in the most obvious way:

Definition A.11. Let R be a representation of the Lie group G. Furthermore, let $\gamma : [-\epsilon, \epsilon]$ be a curve with $\gamma(0) = \mathbb{1}_G$ and $\dot{\gamma}(0) = A \in \mathfrak{g}$.
The representation of A then is defined as:

$$R(A) := \frac{\mathrm{d}\, R(\gamma(t))}{\mathrm{d}\, t}$$

If we compare with (4.1), we note that the Lie bracket is mapped onto the commutator:

$$R(A) \circ R(B) - R(B) \circ R(A) =: [R(A), R(B)] = R([A, B])$$

[3] The dual V^* to a vector space is the space of linear maps from V to the underlying field, e.g. $\mathbb{R}$ or $\mathbb{C}$.

[4] For two finite dimensional vector spaces V and W with bases $\{v_1, v_2, \dots v_n\}$ and $\{w_1, w_2, \dots w_n\}$, the tensor product is defined as the vector space spanned by the formal basis elements $\{(v_1, w_1), (v_1, w_2), \dots (v_1, w_m), (v_2, w_1), (v_2, w_2), \dots (v_n, w_m)\}$.

[5] This is the space $V_1 \times V_2$ with addition $(v_1, v_2) + (v_1', v_2') := (v_1 + v_1', v_2 + v_2')$.

Using the property of the Lie bracket being mapped onto the commutator, we can also define Lie algebra representations abstractly:

Definition A.12. Let $\mathfrak{g}$ be a Lie algebra and V be a vector space. A representation of $\mathfrak{g}$ on V is a homomorphism $R : \mathfrak{g} \to \mathrm{Hom}(V)$ such that the Lie bracket is mapped onto the commutator of linear maps in $\mathrm{Hom}(V)$:

$$R([A, B]) = [R(A), R(B)]$$

Adjoint representation

There is one canonical representation of a Lie algebra on itself, the **adjoint representation** ad. It maps an element A onto a linear map ad_A on $\mathfrak{g}$

Definition A.13. The adjoint representation is defined as the action of the Lie bracket:

$$\mathrm{ad}_A(B) := [A, B]$$

This way, it is automatically linear. We only have to check that ad maps the Lie bracket onto the commutator.

$$\begin{aligned}
[\mathrm{ad}_A, \mathrm{ad}_B](C) &= \mathrm{ad}_A \, \mathrm{ad}_B \, C - \mathrm{ad}_B \, \mathrm{ad}_A \, C \\
&= [A, [B, C]] - [B, [A, C] \\
&= [A, [B, C]] + [B, [C, A]
\end{aligned}$$

Remember that the Jacobi identity holds:

$$\begin{aligned}
&= -[C, [A, B]] \\
&= +[[A, B], C] \\
&= \mathrm{ad}_{[A,B]}(C) \\
\implies [\mathrm{ad}_A, \mathrm{ad}_B] &= \mathrm{ad}_{[A,B]}
\end{aligned}$$

Killing form

Using the adjoint representation, we can define a bilinear form $\langle , \rangle$ on a Lie algebra:

Definition A.14. The **Killing form** $\langle , \rangle$ is defined as:

$$\langle A, B \rangle := \mathrm{tr}\,(\mathrm{ad}_A \circ \mathrm{ad}_B)$$

With $\circ$, we mean composition of linear maps on $\mathfrak{g}$. With tr, we mean the trace on the space $\mathrm{Hom}(\mathfrak{g})$ of linear maps on $\mathfrak{g}$.

The Killing form is an important tool in the classification of Lie algebras, but it also has applications in the physics of principal and Cartan connections, especially when it is non-degenerate.

A.5 Frequent examples

Definition A.15. With $GL(N)$, we denote the Lie group of invertible, linear transformations of $\mathbb{R}^n$.
With $GL(N, \mathbb{C})$, we denote the Lie group of invertible, linear transformations of $\mathbb{C}^n$.
$SL(N)$ and $SL(N, \mathbb{C})$ denote the subgroups of $SL(N)$ and $SL(N, \mathbb{C})$ where every element has determinant 1.

Definition A.16. With $U(N)$, we denote the group of **unitary**, linear maps on a complex N-dimensional Hilbert space $(H, \langle\cdot,\cdot\rangle)$.
A map is called unitary if it preserves the scalar product on H:

$$\langle v, w\rangle = \langle Uv, Uw\rangle \, \forall v, w \in H; U \in U(N)$$

With $SU(N)$, we mean the group of all elements of $U(N)$ with determinant 1.
$U(N)$ is also called **unitary group**. $SU(N)$ is also called **special unitary group**.

Definition A.17. With $O(N, M)$, we denote the group of all **orthogonal**, linear maps on a **real** $N + M$-dimensional vector space $(H, \langle\cdot,\cdot\rangle)$ with a scalar product of signature (N, M).
A map is called "orthogonal" if it preserves this scalar product:

$$\langle v, w\rangle = \langle Ov, Ow\rangle \, \forall v, w \in H; O \in O(N, M)$$

With $SO(N, M)$, we denote the group of all elements of $O(N, M)$ with determinant 1.

$O(N, 0)$ is abbreviated as $O(N)$. It is commonly called and $SO(N) := SO(N, 0)$ **"special orthogonal group"**.

Example A.18. $U(N), O(N, M), SU(N), SO(N, M)$ are real Lie groups if N and M are large enough.
They are all Lie subgroups of $GL(N, \mathbb{C}), GL(N + M), SL(N, \mathbb{C}), SL(N + M)$, respectively. Since all these groups were defined as acting as linear maps on a vector space, their definition is a group representation at the same time, the so-called **fundamental representation**.

The groups $SO(N, M)$ can have a complicated topology. Often, they are not simply connected.
There is a construction, called "double cover", giving a Lie group $\mathrm{Spin}(N, M)$ with the same Lie algebra like $SO(N, M)$. For $N > 2$ and $M = 0$, the double cover is simply connected.

It turns out that the frequently used groups $\mathrm{Spin}(3, 1)$ and $\mathrm{Spin}(3)$ are isomorphic to $SL(2, \mathbb{C})$ and $SU(2)$.

B Vector-valued differential forms[4]

To make sure to agree on notation, the essential definitions of vector-valued differential forms are given here.

B.1 The exterior algebra over a vector space

Definition B.1. Let W and V be vector spaces.
A ***W*-valued *n*-Form** ω on V is defined to be a linear function

$$\omega : \underbrace{V \times V \times \ldots \times V}_{n} \to W.$$

that is antisymmetric under permutation of two arguments:

$$\omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -\omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n)$$

The vector space $\Lambda^n(V, W)$ of W-valued n-forms on V is called the ***W*-valued exterior algebra** of V.

The simple case of the $\mathbb{R}$-valued exterior algebra $\Lambda^n(V,\mathbb{R})$ is normally denoted by $\Lambda^n(V)$ and its elements are just called "forms".
A W-valued 0-form is simply an element of W.

Remark B.2. Remember that the definition of a tensor T of rank n over V is to be a multilinear map from $\underbrace{V \times V \times \cdots \times V}_{n\text{-times}}$ to $\mathbb{R}$. Therefore, $\Lambda^n(V)$ is a subspace of the space $\mathcal{T}^n(V)$ of n-ranked tensors over V.

Furthermore, there is a canonical projection from $\mathcal{T}^n(V)$ to $\Lambda^n(V)$ which we will call Anti, or antisymmetrisation:

$$(\mathrm{Anti}(T))(v_1, v_2, \dots v_n) := \frac{1}{n!} \sum_{\sigma \text{ permutation}} \mathrm{sgn}(\sigma) \cdot T\left(v_{\sigma(1)}, v_{\sigma(2)}, \dots v_{\sigma(n)}\right) \tag{4.2}$$

Remark B.3. If we supply W with a basis, we can decompose every element w of W into components w_I. This can be done with a form evaluated at some vectors: $\omega(v_1, v_2, \dots, v_n)_I$. Since taking the I-th component is a linear operation, the result can be interpreted as a real-valued form ω_I. All ω_I together contain the same information as ω. Every linear operation like index contraction, index lowering/raising, wedge products and so on can follow.

Of course, this would have worked for matrix indices or general tensor indices as well.

Physicists often prefer this notation.

The wedge product

Definition B.4. Let $\circ : W \times U \to X$ a multilinear[6] map of vector spaces which we think of as a multiplication. Then, there naturally corresponds to it a multiplication of W-valued and U-valued forms, which will be called the **wedge product** $\wedge_\circ$.

More specifically, the wedge product of a W-valued n-form ω with a U-valued m-form η is an X-valued $n+m$-form:

$$\wedge_\circ : \Lambda^n(V,W) \times \Lambda^m(V,U) \to \Lambda^{n+m}(V,X)$$

It is defined by its action on its arguments as:

$$(\omega \wedge \eta)(v_1, \dots v_{n+m}) := \sum_{\substack{\pi \text{ is a permutation with:} \\ \pi(1)<\ldots<\pi(n) \\ \pi(n+1)<\ldots<\pi(n+m)}} \mathrm{sgn}(\pi) \cdot \omega(v_{\pi(1)}, \dots v_{\pi(n)}) \circ \eta(v_{\pi(n)}, \dots v_{\pi(n+m)})$$

Some definitions also sum over all possible permutations and therefore have to use the prefactor $\frac{1}{n!m!}$.

When multiplying two 0-forms, one can also leave out the $\wedge_\circ$ and means simply using $\circ$.

Example B.5. The multiplication $\circ$ determines most of the properties of $\wedge_\circ$. If the multiplication $\circ$ to use is unambigous, one also writes simply $\wedge$.

- One can always take the tensor product $\otimes$ as multiplication.
- If $W = U$, it is also possible to follow the tensor product by the antisymmetrisation Anti. According to remark B.2, Anti maps onto forms over W. So this wedge product produces an element of $\Lambda^n(\Lambda^2(W))$. The corresponding wedge product is denoted as $\wedge_\wedge$.

[6]Linear in both arguments independently

- If $W = U$ and it is an associative algebra, one usually takes the algebra product. The result again is a W-valued form. Associativity of $\circ$ then ensures the associativity of $\wedge_\circ$.

  If the algebra is even commutative, the exterior algebra over it is graded commutative.

- If W is an algebra and one can specify a representation R of W on U, one can define $\circ : W \times U \to U$ as the operator application:

$$\begin{aligned} w \in W, \quad u \in U \\ w \circ u := R(w)(u) \end{aligned}$$

  The corresponding wedge product is denoted as $\wedge_R$.

A comment on Lie algebra valued forms and Lie algebra representations

Let W a Lie algebra. If one wants to define the $\wedge$-product on W-valued forms, one has to give up associativity since Lie algebras are not associative in general.

It is natural to use the Lie bracket as multiplication: $a_1 \circ a_2 := [a_1, a_2]$ The wedge product of two W-valued forms ω and η is then commonly written as $[\omega \wedge \eta]$ or $[\omega, \eta]$.

However, if there is a representation R of the Lie algebra on U, one can define the $\wedge_{R,\circ}$-product which multiplies $\mathrm{Hom}(U)$-valued forms, based on the composition of homomorphisms. In some special cases, for example when multiplying forms with odd degrees with themselves, the resulting form takes again values in $R(W)$ and can then reinterpreted as W-valued forms. For 1-forms, this is easy to see:

$$\begin{aligned} R((\omega \wedge_{R,\circ} \omega)(v_1, v_2)) &= R(\omega(v_1)) \circ R(\omega(v_2)) - R(\omega(v_2)) \circ R(\omega(v_1)) \\ &= [R(\omega(v_1)), R(\omega(v_2))] \\ &= R([\omega(v_1), \omega(v_2)]) \end{aligned}$$

For this product, we have $\omega \wedge_{R,\circ} \omega = \frac{1}{2}[\omega \wedge \omega]$.

B.2 Differential forms

Intuitively, a differential form is a form on every tangent space of a manifold.

Most people simply say “form” instead of “differential form” for brevity.

Definition B.6. A W-valued differential n-form ω on a manifold M is a smooth map $\omega : TM \to W$, such that ω restricted to any tangent space is a W-valued n-form over that tangent space.

Exterior derivative

Definition B.7. Let f be a differential 0-form.
The exterior derivative $\mathrm{d}f$ of f is defined as the unique differential 1-form satisfying $(\mathrm{d}f)(v_p) = v(f)_p$ for all vector fields v and points $p \in M$.

If f is vector-valued, then this definition can be made for each component, in an arbitrary basis.

Corrolary B.8. For a coordinate chart $x : U \to V \subset \mathbb{R}^n$ ($U \subset M$ open), the derivatives ∂_i form bases of each tangent space. So taking the components x_i of x, which are just real-valued functions on U, the differential forms $\mathrm{d}x_i$ form bases of the 1-forms over U.

Therefore, the forms $\mathrm{d}x_{i_1} \wedge \mathrm{d}x_{i_2} \wedge \ldots \wedge \mathrm{d}x_{i_n}$ with $i_1 < i_2 < \ldots < i_n$ are a basis of n-forms on U.

Definition B.9. So every n-form ω can be written as:

$$\omega = \sum_{i_1 < i_2 < \ldots < i_n} f_{i_1,i_2,\ldots,i_n} \mathrm{d}x_{i_1} \wedge \mathrm{d}x_{i_2} \wedge \ldots \wedge \mathrm{d}x_{i_n}$$

The $f_{i_1,i_2,\ldots,i_n}$ are functions on U. We define the exterior derivative of ω as:

$$\mathrm{d}\omega = \sum_{i_1 < i_2 < \ldots < i_n} \mathrm{d}f_{i_1,i_2,\ldots,i_n} \wedge \mathrm{d}x_{i_1} \wedge \mathrm{d}x_{i_2} \wedge \ldots \wedge \mathrm{d}x_{i_n}$$

Corrolary B.10. Immediately, one can show that $\mathrm{d}\mathrm{d}\omega = 0$ for every form. This is often written as $\mathrm{d}^2 = 0$.

B.3 Vielbeins, or Soldering Forms

There is a special case of vector-valued differential forms, called "vielbein", or equivalently "coframe field" or "soldering form". It can replace a (pseudo-)Riemannian metric on a manifold. The idea is having a constant metric on a vector space, say Euclidean or Minkowski space, and pulling back this metric to the manifold.

Definition B.11. A **vielbein** on an n-dimensional manifold is a V-valued differential 1-form e satisfying:

- V has dimension n and is equipped with a non-degenerate metric η with signature (n_1, n_2).
- e is an isomorphism[7] of vector spaces at every point of M.

We then have a metric g on M with the same signature (n_1, n_2) like η defined by:

$$g(v, w) := \eta(e(v), e(w))$$

There is no preferred way to choose a specific e to reproduce a given metric. Indeed, a local $O(n_1, n_2)$-transformation maps e onto another vielbein that defines the same metric.

For 4 dimensions, a vielbein is also often called a **tetrad**, and in 3 dimensions **triad**.

Maurer-Cartan form

The Maurer-Cartan form is a special case of a vielbein that is given naturally on Lie groups. The intuition of this is that the tangent space at every group element looks like the Lie algebra $\mathfrak{g}$ of the Lie group G, it just "sits at the wrong place": The Lie algebra is defined only over the tangent space of the neutral element $\mathbb{1}$.

But since we are talking about Lie *groups*, there is a canonical way to map a group element g to $\mathbb{1}$: Multiplication from the left[8] with g^{-1}, which can be expressed by the operator $L_{g^{-1}}$:

$$L_{g_1}(g_2) := g_1 g_2$$

The derivative $\mathrm{d}L_{g^{-1}}$ of this map transports the tangent space $T_g G$ of g isomorphically to $T_{\mathbb{1}} G \cong \mathfrak{g}$, so it is a $\mathfrak{g}$-valued 1-form on G that serves as a vielbein.

[7] Note that this implies that e is a trivialisation of the tangent bundle of M. For general topologies, there also is a generalised notion of the vielbein formalism, involving a "fake tangent bundle", as introduced for example in [5].

[8] Note that we could use multiplication from the right as well, yielding a mathematically equivalent definition. The most frequent definition however is multiplication from the left.

Definition B.12. The **Maurer-Cartan form** on a Lie group G is the $\mathfrak{g}$-valued differential 1-form ω_{MC} defined by $\mathrm{d}L_{g^{-1}}$ acting on the tangent space $T_g G$ at every element g.

Note that the Killing form on $\mathfrak{g}$ can sometimes be degenerate. In that case, one has to choose another metric on $\mathfrak{g}$ if one wants the metric on G to be non-degenerate.

Bibliography

[1] Baez, John C.: *Spin network states in gauge theory.* Adv. Math. 117 (1996): 253-272, 1994. http://www.arXiv.org/abs/gr-qc/9411007.

[2] Livine, Etera R: *Projected spin networks for lorentz connection: Linking spin foams and loop gravity.* Class.Quant.Grav. 19 (2002) 5525-5542, 2002. http://www.arXiv.org/abs/gr-qc/0207084.

[3] Rovelli, Carlo: *Quantum Gravity.* Cambridge University Press, 2007.

[4] Wikipedia: *Vector-valued differential form.* http://en.wikipedia.org/wiki/Vector-valued_differential_form.

[5] Wise, Derek K.: *MacDowell-Mansouri gravity and Cartan geometry.* Class. Quantum Grav.27:155010, 2010. http://arxiv.org/abs/gr-qc/0611154.

Index

Erklärung

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 13. 7. 2011, Manuel Bärenz