Gravity and Gauge

Nicholas J. Teh

June 29, 2011

Abstract

Philosophers of physics and physicists have long been intrigued by the analogies and disanalogies between gravitational theories and (Yang-Mills-type) gauge theories. Indeed, repeated attempts to collapse these disanalogies have made us acutely aware that there are fairly general obstacles to doing so. Nonetheless, there is a special case (viz. that of (2+1) spacetime dimensions) in which gravity is often claimed to be identical to a gauge theory. We subject this claim to philosophical scrutiny in this paper: in particular, we (i) analyze how the standard disanalogies can be overcome in (2+1) dimensions, and (ii) consider whether (i) really licenses the interpretation of (2+1) gravity as a gauge theory. Our conceptual analysis reveals more subtle disanalogies between gravity and gauge, and connects these to interpretive issues in classical and quantum gravity.

Contents

1 Introduction

> 'The proper method of philosophy consists in clearly conceiving the insoluble problems in all their insolubility and then in simply contemplating them, fixedly and tirelessly, year after year, without any hope, patiently waiting.'
>
> Simone Weil

The above quote from Weil is certainly not to be held up as a model of philosophical methodology, but it may nonetheless serve as a description of the repeated attempts in physics to reconcile gravity (which we shall take to mean General Relativity) with gauge theory (which we shall take to mean a *Yang-Mills type* gauge theory[1]).[2] In making these stipulative definitions, we do not mean to suggest that the only reasonable thing to mean by a (classical) gravitational theory is a theory given by Einstein's field equations, or one in which gravity appears as an aspect of spacetime structure.[3] Rather, we are restricting the scope of this paper to two specific theories that everyone agrees are a gravitational theory and a gauge theory respectively, and asking when and whether they coincide.

These attempts have relied on various strategies, the most straightforward of which is to assimilate gauge theory to the model of a gravity theory, or vice versa. For instance, the Kaluza-Klein strategy tries to

[1] By gauge theory of Yang-Mills type, we mean a theory whose dynamical variable is a connection on spacetime that takes values in the gauge group, and which has a Yang-Mills type action. This does not yet commit one to modeling the gauge group G as a fiber of a principal G-bundle, as demonstrated by the Cartan gauge theory of §3.2.

[2] There is another oblique connection between Simone Weil and gauge theory: the Weil-Shimura-Taniyama conjecture, due in part to her brother Andre Weil, is related to the Geometric Langlands program, which has recently been pursued by means of gauge-theoretic methods [33].

[3] For instance, consider Nordstrom's scalar field theory of gravitation (see [22] for an introduction and references), in which gravity is not an aspect of spacetime structure. We thank an anonymous referee for the example and for helping us clarify this point.

build gauge symmetry (in a lower-dimensional space) into diffeomorphism symmetry (in a higher-dimensional space) [11], whereas the McDowell-Mansouri strategy tries to rewrite the diffeomorphism symmetry of gravity as the gauge symmetry of a Yang-Mills type theory [17];(here and in the rest of this paper, we shall take *gauge symmetry* to mean 'a local symmetry of a *gauge* theory under which its observables are invariant'). A rather more subtle strategy comes from the application of the holographic principle to identify quantum gravity on some bulk space with a quantum conformal gauge theory on the boundary of that space [15]: this introduces various complications such as the right choice of bulk geometry, the appearance of a (roughly) classical-quantum correspondence, taking the large N limit, accounting for the energy scale of the gauge theory, and (often) the use of supersymmetry on the boundary to control the interactions. In this essay, we shall however be simple-minded and concern ourselves solely with the McDowell-Mansouri strategy of recasting general relativity as a gauge theory; indeed we shall concern ourselves with a case in which it is often claimed that gravity can be *identified* with gauge theory.

Philosophers of physics, too, have been intrigued by both analogies and disanalogies between gravity and gauge theory. Some have emphasized the analogies, e.g. [5], [3], and [4]. Others (e.g. Weinstein [28]) have criticized the former group for ignoring the disanalogies between gravity and gauge, and thus failing to appreciate that gravity is not a bona fide gauge theory. Yet others warn that debates along these lines have a tendency to fall into mere semantic disagreement: witness Wallace [27], who writes that '...there are genuine dynamical differences between general relativity and more "conventional" gauge theories such as electromagnetism, but these differences are best appreciated on their own merits rather than being annexed to the essentially sterile debate as to whether or not general relativity is "really" a gauge theory.' (p. 164) We hope that the arguments of this paper will make it clear that there is at least one context in which the question 'Is general relativity a gauge theory?' is not only interesting but subtle, and connected to various issues in the interpretation of these respective theories.

The context that we have in mind is that of three-dimensional (3D) gravity, whose study was initiated by [25] in 1963, and which then experienced a revival (mostly in the 1980s) through the work of [10], [30], [1] etc. The main, and immediately apparent, feature of 3-dimensional pure gravity is that it has no local dynamical degrees of freedom: this can be seen in the $\Lambda = 0$ case by conjoining the fact that the Riemann curvature tensor is entirely determined by the Ricci tensor in three dimensions, with the fact that the vacuum Einstein equations imply that the Ricci tensor vanishes. Below, we canvas various motivations for studying three-dimensional gravity before we explain why it is relevant for our purposes here, viz. investigating the analogies and disanalogies between gravity and gauge.

1.1 Motivation

Before we list 'physical' motivations for studying 3D gravity, we should mention that 3D gravity is enormously interesting even from just the mathematical point of view: it links together subjects as various as mathematical relativity, Cartan geometry, the monster group, secondary characteristic classes, quantum groups, non-commutative geometry, knot theory, higher categories, and the geometry of 3-manifolds.[4]

Turning now to a more physical view of the subject, 3D gravity has often been advocated as

- An exploration of the (spacetime) dimension-dependent features of gravity.
- A useful toy model for exploring different approaches to quantizing gravity.
- A tractable model in which most of the 'conceptual problems' of quantum gravity (e.g., the nonlocality of observables, acausality, the problem of time, the role of topology change, and the problem of normalizing probabilities) remain, and can thus be addressed independently of the geometric and other technical complications introduced by higher dimensions.
- A setting for holography, where the boundary theory is given by a 2-dimensional conformal field theory.[5] Indeed, it frames the puzzle of black hole entropy in a particularly acute manner.
- As a model of cosmic strings.
- A setting for understanding the relationship between the invertibility of coframe fields (dreibein, in the 3D case) and the renormalizability of gravity.

These are all deeply interesting directions of research for the philosophy/foundations of physics, especially the connection with holography, but in this paper we wish to emphasize the relationship between pure 3D gravity and classical gauge theory. The importance of this relationship is not to be underestimated: for one, it forms the basis for a number of the points above, e.g. it leads to the Chern-Simons approach to quantizing gravity, and also renders apparent the sense in which pure 3D gravity is 'trivial', and thus a more tractable model for addressing conceptual questions than higher-dimensional gravity. Furthermore, if 3D gravity turns out to be a gauge theory, then (at least in three dimensions) we can reject Weinstein's charge that some theorists are making an illicit inference from 'gauge-invariant quantities are the observables of a quantum gauge theory' to 'diffeomorphism-invariant quantities are the observables of quantum gravity'.[6]

[4] [8] describes several of these links; for other references, see the bibliography of [20].

[5] The seminal insights that led to this point of view were laid out in [7].

[6] Against Weinstein, theorists who wish to take the observables of quantum gravity to be diffeomorphism-invariant quantities may nonetheless be justified in saying that they are merely pursuing an *analogy* with gauge theory, and a fairly natural one at that.

Our interest in the relationship between pure 3D gravity and gauge theory stems from the fact that, in this setting, two familiar disanalogies between gravity and gauge theory collapse completely (see §2 below). Nonetheless, even in this best-case scenario for the putative identification of gravity and gauge, there is (perhaps surprisingly) still room to ask: are gravity and gauge theory really equivalent? We shall see that the 3D setting allows us to give an incisive analysis of the differences between gravity and gauge, which in turn leads us to a more acute appreciation of the differences between gravity and gauge in *arbitrary* dimensions (although it may be impossible to articulate the differences so clearly in arbitrary dimensions; thus the utility of the 3D setting). Furthermore, the 3D setting also raises important questions about how our interpretation of GR interacts with the differences between gravity and gauge.

Before we proceed to lay out the plan of this paper, it will be important for us to forestall a particular objection to our motivation for considering the 3D case. As we mentioned earlier, pure 3D gravity has no local dynamical degrees of freedom; for this reason, it fails to have a 'good' Newtonian limit. More precisely, when we linearize gravity (i.e. make a linear perturbation about Minkowski space) and take the Newtonian limit, we obtain the equations of motion of a test particle (up to rescaling of the coupling constant)

$$\partial_t^2 x^i + \frac{2(n-3)}{n-2}\partial_i \Phi = 0, \tag{1}$$

where Φ is the gravitational potential and n is the number of spacetime dimensions. Evidently, there is no gravitational force acting on test particles in the Newtonian limit of 3D gravity!

An objector could then reasonably ask, 'What makes you think this is gravity in the sense that we know and love from four dimensions? And if it isn't, why should we be interested in this model?' While we acknowledge that 3D gravity is *dynamically* very different from gravity in higher dimensions, we also believe that this objection does not vitiate our motivations, for two reasons. First, as we mentioned above, many of the 'conceptual problems' of quantum gravity do not turn on the presence of local dynamics (i.e. local curvature). Nor are the more subtle obstructions to identifying gauge and gravity changed by the presence of local dynamics – the only difference in higher dimensions is that one is liable to focus on other, glaringly obvious, obstructions to making this identification. Second, despite the dynamical differences between pure 3D gravity and higher-dimensional gravity, there is an important similarity that is in fact dynamical, viz. the *form* of the Lagrangian remains the same in both cases (viz., it is the time-honored Einstein-Hilbert Lagrangian).[7]

[7]Contrast this with the higher-dimensional McDowell-Mansouri case, where the Lagrangian (whether it is expressed in the standard formulation or in the BF (for 'background field' formulation) is not of Einstein-Hilbert form, although varying it does yield the Einstein equations of motion with non-zero cosmological constant and vanishing torsion.

1.2 Prospectus

In §2, we review two obvious disanalogies between gravity and gauge, the second of which appears to have been neglected in the philosophical literature. §3 shows how these disanalogies can be overcome, first from the perspective of gravity, and then once more from the perspective of Chern-Simons gauge theory. §3.1 introduces the first-order (Palatini) formalism of 3D gravity and uses physical reasoning to motivate the steps through which first-order gravity can be put into the form of a gauge theory. §3.2 introduces Chern-Simons gauge theory and adopts a more geometrical and systematic approach to overcoming the disanalogies of §2. §4 then argues that, even in three dimensions – the best-case scenario where all the obvious dianalogies collapse – two subtle disanalogies still remain. We conclude by discussing how these disanalogies relate to the interpretation of GR, both from the perspective of a stand-alone classical theory, and also when viewed in light of its quantization.

2 Disanalogies

Our first disanalogy between gravity and gauge theory will strike the reader as glaringly obvious: it turns on the fact that the gauge transformations of a (Yang-Mills type) gauge theory act on 'internal space' and not on spacetime, whereas the diffeomorphism symmetry of gravity acts on spacetime itself (and also on geometric objects defined on spacetime). Let us however try to express this idea more precisely.

First, recall that a gauge theory with gauge group G on a spacetime manifold M is standardly described by a principal G-bundle P:

$$\pi : P \to M,$$

where the fiber over each point of M is a copy of G. A choice of gauge is a section of this bundle, and a gauge transformation is a vertical automorphism of the bundle. A gauge field (as physicists usually define it) is a $\mathfrak{g}$-valued one-form on M.[8]

It immediately follows that there are two related obstructions to identifying gravity with a gauge theory, one on the side of gravity and the other on the side of a gauge theory. On the one hand,

> (Obst-Grav) The (small) diffeomorphism symmetry of gravity cannot be described as a gauge group (i.e. as a fiber of a principal bundle over spacetime, at least in the standard formalism of gauge theory), on pain of its not being able to act on spacetime points at all.

[8]This is often also called a (local) connection on M, but note that it is really the pullback (via a section) of a connection form on P.

By 'diffeomorphism symmetry' here, we are only referring to the *small* diffeomorphisms, i.e. those diffeomorphisms that are smoothly contractible to the identity. The small diffeomorphisms are what most authors on GR have in mind when they discuss diffeomorphism symmetry, but we have an additional reason for making this restriction explicit here, viz. that the large (i.e. the non-small) diffeomorphisms will play an important role in our story in §4, where we elaborate upon this notion.

On the other hand,

> (Obst-Gauge) Since a gauge transformation is described by a vertical automorphism, it follows that gauge transformations cannot act on spacetime points. For this reason, authors such as Healey ([13], p. 98) claim that while 'Leibniz equivalence' is a viable strategy to defuse the hole argument in the case of gravity, this strategy is not available in the case of gauge theory.[9]

These are of course just two ways of making the same elementary point, but we are setting up the problematic in this way in anticipation of §3, where we consider how to overcome these obstructions from the perspective of gravity and gauge respectively.

(Obst-Grav/Gauge) has often been remarked on by philosophers (see e.g. [28] and [13] p. 98). A second potential obstruction to identifying a gauge theory with gravity has been far less remarked upon, and indeed only comes into view once (Obst-Grav/Gauge) has been overcome: suppose for the sake of argument that the diffeomorphisms and the metric can somehow be repackaged as gauge transformations and gauge fields – do we then have a gauge theory? Clearly not, as a gauge theory also requires an action,[10] and the definition of this action requires the existence of an appropriate (i.e. nondegenerate, invariant) trace function on the field values, i.e. the Lie algebra $\mathfrak{g}$. Furthermore, in the case of the *particular* 'gravitational' gauge theory of interest to us, this action must coincide with the Einstein-Hilbert action.[11] Let us formalize this requirement as

> (Trace) A nondegenerate bilinear form must exist on the gauge field values (i.e. the Lie algebra $\mathfrak{g}$ of the gauge group G) such that the gauge field action takes the form of an Einstein-Hilbert

[9] Note that Healey's claim can reasonably be challenged: some, e.g. Graeme Segal (private communication and unpublished remarks on gauge fields), defend an 'internal space' analog of Liebniz equivalence.

[10] A qualification: in this case, we expect to obtain the equations of motion from the Lagrangian, and so 'a gauge theory requires an action'. Nonetheless, there exist reasonable non-Lagrangian gauge theories in other contexts: see [16] for how to construct and quantize them, and [9] for an example of S-duality between Lagrangian and non-Lagrangian $\mathcal{N} = 2$ superconformal gauge theories (Gaiotto duality).

[11] Here we again remind the reader that a general gravitational theory need not have an Einstein-Hilbert action; for instance, the Brans-Dicke theory [6] only approximates the Einstein-Hilbert action under certain conditions (we thank an anonymous referee for the example). Such theories are outside the scope of our investigation.

action.

As we shall see in §3, (Trace) cannot be satisfied in arbitrary dimensions, but due to a 'minor miracle', such a trace function does exist in three dimensions! (The existence of this miracle was first observed by Achucarro and Townsend [1], and rediscovered by Witten in [30].)

We have just highlighted two sorts of disanalogies between gravity and gauge theory. In the next section, we shall see that these disanalogies collapse completely in three dimensions.

3 3D Gravity and Gauge

The goal of this section will be to introduce Einstein gravity and Chern-Simons theory on a 3-manifold, and to show that when suitably interpreted, both (Obst-Grav/Gauge) and (Trace) can be overcome.

However, before we begin, let us note that we will be making two simplifications in what follows. First, we will only discuss models with zero cosmological constant ($\Lambda = 0$), except for a brief digression in §3.2, where we indicate how to generalize our arguments to $\Lambda \neq 0$ models. This restriction is merely technical, as the addition of a cosmological constant does not change the fact that our models have no local degrees of freedom – in other words, there are no gravitational waves propagating in the classical theory, and there are no gravitons propagating in the quantum theory.

Second, we will only discuss 'pure' (or 'vacuum') gravity models, viz. those that do not include matter degrees of freedom. Adding point particles into the theory (which are modeled as 'punctures' in spacetime) will introduce additional (global) degrees of freedom (by means of conical singularities), but this is true regardless of whether we are working with a Chern-Simons gauge theory or (2+1) gravity. For the canonical technique of including matter in a Chern-Simons theory, see Witten's work relating Chern-Simons theory to knot theory [31]. This was applied to particles in 3D gravity (and also to include a boundary at infinity) in [19].

Two of the most popular methods of investigating 3D gravity are the first-order formalism ([8] p. 25) and the ADM formalism ([8] p. 12). While the ADM formalism is important for various applications (and indeed, the constraints take a simple and aesthetically pleasing form in (2+1) dimensions), we shall here restrict ourselves to the first-order formalism, in which the connection between gravity and Chern-Simons gauge theory is especially transparent. Furthermore, the solutions to the first-order equations of motion will give us the phase space of the theory directly, since (classical) phase space is nothing other than the space of solutions to the equations of motion. Treating the phase space at this level will be sufficient for the arguments of §4.

3.1 (2+1) Gravity

Let us begin by recalling the standard formulation of 3D gravity: upon varying the vacuum, $\Lambda = 0$ Einstein-Hilbert action

$$S[g] = \int_M d^3x\sqrt{-g}R \tag{2}$$

we obtain the vacuum, $\Lambda = 0$ Einstein equations

$$R_{ij} - \frac{1}{2}g_{ij}R = 0. \tag{3}$$

A model of GR is a solution (M, g) to these equations, where M is a manifold and g is a Lorentzian metric.

As we argued earlier, the solutions to pure, $\Lambda = 0$ gravity must be flat. We will not need explicit solutions in what follows, but note that there are examples of flat geometries other than Minkowski space, e.g. the torus universe $[0, 1] \times T^2$, where $[0, 1]$ is time.

Now instead of viewing the action S as a functional of the metric, à la Einstein-Hilbert, we could view it as a functional of two independent variables, viz. g and the connection ∇; this is called 'the Palatini action'. Varying the Palatini action yields not only the Einstein equations, but also the metric compatibility (or torsion-freeness) condition. In what follows, we shall pursue a variant of the Palatini strategy (which is often called the 'first-order' or 'Cartan' formalism) and take as our independent dynamical variables the dreibein e (which intutively corresponds to the 'square root' of the metric) and the spin connection ω (which intuitively corresponds to a 'gauge field' for the Lorentz group $SO(2, 1)$). Let us now work locally and explain this in greater detail. The global geometry will be explained from the perspective of Cartan gauge theory in the next section.

Consider a spacetime metric

$$ds^2 = g_{ij}(x)dx^i dx^j \qquad i, j = 0, 1, 2. \tag{4}$$

We would like to exploit the idea that the metric can be written in *any* basis, and not just a coordinate basis. So, for instance, if we let $\{E^a = e^a_i dx^i\}$, $a = 0, 1, 2$, be an arbitrary set of basis covectors, we will write the metric as:

$$\begin{aligned} ds^2 &= \bar{g}_{ab}E^a E^b \\ &= \bar{g}_{ab}e^a_i e^b_j dx^i dx^j \\ &= g_{ij}dx^i dx^j, \end{aligned}$$

so that, clearly, $g_{ij} = \bar{g}_{ab}e^a_i e^b_j$.

Now recall that at a point x of M, g can always be expressed as a Minkowski metric η (i.e. with vanishing Christoffel symbols), and in an arbitrarily small neighborhood of p, it will pick up second-order corrections due to the Riemann tensor.[12] We can thus make the judicious choice of letting $\bar{g}$ be the metric η, which will result in the curvature corrections being 'transferred' into the coefficients e_i^a. The basis covectors $\{E^a\}$ now form a pseudo-orthonormal basis at a point.

Let us pause to introduce some terminology that will make the meaning of the e_i^a's clearer: a *dreibein* or *coframe field* e is an $\mathbb{R}^{2,1}$-valued 1-form

$$e = e_i^a dx^i T^a, \tag{5}$$

where T^a is a basis element of the vector space $\mathbb{R}^{2,1}$. Derivatively, its coordinate coefficients e_i^a are often also referred to as dreibein. We use the the letters $a, b, c, \ldots$ to denote the 'fake tangent space' or $\mathbb{R}^{1,2}$ indices of the dreibein (i.e. indices that are raised/lowered by the flat metric η_{ab}) and the letters $i, j, k, \ldots$ to denote the 'spacetime' indices of the dreibein (i.e. indices that are raised/lowered by the spacetime metric g_{ij}). The dreibein e_i^a itself is used to turn spacetime indices into fake tangent space indices and vice versa (if it is invertible). Note that the invertibility of g forces the compatibility conditions $e_i^b e_b^k = \delta_i^k$ and $e_a^j e_j^b = \delta_a^b$; taken jointly, these amount to the statement that the dreibein e_i^a is invertible. (This issue about the invertibility of g will arise again in §4.2, where we also consider the case of non-invertible dreibein.)

The fact that our metric now takes the form $g_{ij} = \eta_{ab} e_i^a e_j^b$ discloses to us a pointwise symmetry that would otherwise have gone unnoticed! More explicitly, it is immediately apparent that the metric g is invariant under the pointwise Lorentz transformation

$$e_i^a \mapsto \Lambda_b^a e_i^b, \tag{6}$$

where $\Lambda \in SO(2,1)$.[13] What this invariance suggests is that there is a Lorentz group 'gauge freedom' at every point of M. By uncovering this 'hidden' gauge symmetry, we have taken a step towards overcoming (Obst-Grav) in §2, i.e. we have succeeded in identifying some of the symmetries of gravity with the 'vertical automorphisms' over a point of M. Nonetheless, the translation symmetries still need to be incorporated via our definition of the gauge field A below.[14]

Since we have (at least) an $SO(2,1)$ gauge symmetry in our theory, it is natural to introduce a gauge field (or connection) corresponding to that symmetry. This gauge field, called a *spin connection*, is an

[12] Contrast this with the case of a symplectic manifold, which is genuinely *locally* flat, and not merely pointwise flat.

[13] Here and in the rest of this paper, we shall take $SO(2,1)$ to be the connected, orthochronous component of the Lorentz group, unless we explicitly state otherwise. Similarly with $ISO(2,1)$.

[14] In other words, we do not yet have the full gauge group of the theory, but only its $SO(2,1)$ subgroup.

$\mathfrak{so}(2,1)$-valued 1-form

$$\omega = \omega^b_{ia} dx^i, \tag{7}$$

where here a and b denote *matrix* indices in the fundamental representation of $\mathfrak{so}(2,1)$ (we could equally well express ω in terms of the usual Lie algebra indices $\omega = \omega^I T^I$). A spin connection is necessary precisely in order to take the gauge covariant derivative of an object with fake tangent space indices; this adds a 'spin connection term' to the usual covariant derivative ∇ given by the Levi-Civita connection. In other words, we now have a general covariant derivative D_i of the form:

$$D_i e^a_j := \nabla_i e^a_j + \omega^a_{ib} e^b_j, \tag{8}$$

where

$$\nabla_i e^a_j := \partial_i e^a_j - \Gamma^l_{ij} e^a_l$$

is the usual (Levi-Civita) covariant derivative.[15]

This makes good physical sense, since $D_i e^a_j = 0$ (because rotation by e^a_j should not change the general covariant derivative) and so a vanishing ω would imply that e^a_j was flat according to ∇. But we know from our earlier discussion that e^a_j cannot be flat in general (since it contains the curvature corrections in the neighborhood of a point), and so we see that ω precisely measures the curvature of e^a_j. It is easy to see that the Riemann curvature tensor is entirely determined by the spin connection.[16]

In this approach to gravity, two equations are particularly important. The first, called Cartan's first structural equation, is

$$d_\omega e := de + \frac{1}{2}[\omega, e] = T, \tag{9}$$

where, as usual, d is the exterior derivative and T is the usual torsion 2-form. The second, called Cartan's second structural equation, is

$$d_\omega \omega := d\omega + \frac{1}{2}[\omega, \omega] = R, \tag{10}$$

where R is the curvature 2-form.[17]

We are now ready to try to treat e and ω as our independent variables, and to write the Einstein-Hilbert action in terms of them. Let us first proceed naively, by trying to integrate the 3-form $e \wedge d_\omega w$ over M:

[15] By considering the way a 'derivative' $D_i e^a_j$ transforms, it is easy to see that this can only be a covariant derivative if we add to the Levi-Civita connection an extra term ω that cancels out the term that has a partial derivative acting on Lorentz transformations: this extra term is precisely the spin connection.

[16] See e.g. p. 274 of [12] or p. 596 of [26].

[17] Note that the bracket in Cartan's structural equations is not that of a Lie algebra, but instead a Lie *super*-algebra, i.e. $[\mu, \nu] = (-1)^{pq+1}[\nu, \mu]$, where $\mu \in \Omega^p(M, \mathfrak{g})$ and $\nu \in \Omega^q(M, \mathfrak{g})$.

$$S(\omega, e) = \int_M \beta(e \wedge R), \tag{11}$$

where β is some nondegenerate invariant symmetric bilinear form acting on the $\mathfrak{iso}(2,1)$-values of e and R. Suppose we follow in Witten's [30] footsteps and take

$$\beta(X, Y) = \mathrm{tr}^\star(XY) = \mathrm{tr}(X \star Y), \tag{12}$$

where tr is the Killing form.[18] More explicitly, we normalize this inner product as:

$$\beta(J_a, P_b) = \eta_{ab}, \qquad \beta(J_a, J_b) = \beta(P_a, P_b) = 0, \tag{13}$$

where $P_a \in \mathbb{R}^{2,1}$ is a momentum generator, and $J_a \in \mathfrak{so}(2,1)$ is a Lorentz generator. It then follows that the integrand of the action is precisely the Ricci scalar multiplied by the volume form that appears in the Einstein-Hilbert action.

We are now in a position to try to overcome (Obst-Grav) and to satisfy (Trace) respectively. First, (Obst-Grav): earlier, we noticed that we had succeeded in reformulating some of the diffeomorphism symmetry of gravity as a pointwise $SO(2,1)$ gauge symmetry. The key to overcoming (Obst-Grav), then, is to show that all the other (small) diffeomorphisms can be captured by a larger gauge symmetry that contains $SO(2,1)$ as a subgroup. Since $SO(2,1)$ is the Lorentzian analog of rotations in Euclidean geometry, and diffeomorphisms include not just the operations of rotating objects but also moving them about, it is reasonable to conjecture that the missing diffeomorphisms come from a *translational* gauge symmetry, and thus to posit the Poincare group

$$ISO(2,1) := SO(2,1) \ltimes \mathbb{R}^{2,1}$$

as the larger gauge group. As a gauge field A for $ISO(2,1)$ we can try writing

$$A \equiv \omega + e, \tag{14}$$

since, as we discussed earlier, ω takes values in $\mathfrak{so}(2,1)$ and e takes values in $\mathbb{R}^{2,1}$. As it turns out, and as we shall see explicitly in §4, this is precisely the right choice of gauge group to capture the (small) diffeomorphism symmetry of gravity! Thus we have succeeded in overcoming (Obst-Grav).

Next, we need to show that given our choice of gauge group $ISO(2,1)$ and gauge field $A = \omega + e$, the (Trace) requirement can be satisfied. The most common Yang-Mills type action in three dimensions is the

[18] The appearance of the Hodge star in this equation may cause some confusion to newcomers, since the Hodge star is usually defined on differential forms. Its appearance here can be understood as follows: the six-dimensional Lie algebra $\mathfrak{iso}(2,1)$ is actually isomorphic to $\bigwedge^2 \mathbb{R}^4$, on which the Hodge star sends 2-forms to 2-forms.

Chern-Simon action:

$$CS(A) = \int_M \beta(A \wedge dA + \frac{1}{3} A \wedge [A, A]). \tag{15}$$

Plugging in our choice of A and β, we immediately see that we obtain the first-order action $S(\omega, e)$. (The reader should check that we would not have obtained this answer had we naively used the Killing form as an inner product, as one usually does in gauge theory!) By the same judicious choice of nondegenerate bilinear form β on $\mathfrak{so}(2,1)$ as before, it is clear that we have satisfied (Trace).

By exercising our powers of physical intuition, we have just succeeded in overcoming the disanalogies between gravity and gauge that were raised in §2. A more systematic approach to our choice of gauge group and gauge field will be taken shortly in §3.2. But before we do so, let us pause to discuss an issue that has already emerged concerning the interpretation of GR. We began with the ordinary Einstein-Hilbert action $S[g]$ and showed that it was formally equivalent to the first-order action $S[e, \omega]$, which could in turn be viewed as the action for a gauge theory. The question then arises as to whether the Einstein-Hilbert theory or the first-order theory are really the same interpretation of GR (modulo a clever relabeling of variables) or whether they should count as genuinely different interpretations.

One way to approach this question would be to say that the content of a classical theory is given by its equations of motion, and so 'having the same equations of motion' is the criterion of identity for interpretations of a physical theory. One might then try to argue that since the Einstein-Hilbert theory and the first-order theory lead to the 'same' equations of motion, these two theories are manifestly identical in physical content.

However, there are good reasons to challenge this approach. For instance, consider that (i) different actions (or Lagrangians) can give rise to the same equations of motion, and (ii) if we mean to interpret the classical theory in light of its quantization,[19] then we should take into account the fact that different actions will in general give rise to different quantizations. It would thus seem that the content of a theory is in this sense underdetermined by its equations of motion. We shall return to this point in §5; for now let us note that there is also a purely classical line of thought that counts in favor of treating the Einstein-Hibert theory and the first-order theory as different interpretations of GR.[20]

This line of thought begins from the observation that one's choice of dynamical variables for gravity constrains the sorts of matter fields that can couple to one's gravitational fields. In particular, the fact that ω is a dynamical variable in the first-order theory allows us to couple *spinors* to the theory, a move that is unavailable in the case of the Einstein-Hilbert theory. Plausibly, the content of a physical theory

[19] See the final section of [2] for a similar discussion.

[20] By 'classical' here, we mean 'prior to quantization'.

(of some field ϕ, say) does not just specify how ϕ is constrained by the theory's equations of motion, but also provides *modal* information about what other fields it is possible for ϕ to interact with. If this is right, then either either the two theories should not count as having the same equations of motion (because they have different dynamical variables), or 'having the same equations of motion' is insufficient as a criterion of identity for interpretations of a theory. Either way, first-order gravity and Einstein-Hilbert gravity turn out to be genuinely different interpretations of GR.

3.2 (2+1) Chern-Simons

We have just seen how 3D gravity can be reformulated as a 3D gauge theory, at least in the sense that (Obst-Grav) can be overcome, and (Trace) satisfied. We will now start from the opposite end, viz. that of a 3D gauge theory and show how we can convert it into 3D gravity.

It will be clearest to begin by revisiting our favorite gauge theory on a closed 3-manifold, viz. *abstract* Chern-Simons theory with gauge group G and gauge field A. We already saw its action $CS(A)$ in (15) above (typically G is taken to be a simple or semisimple Lie group, and the nondegenerate bilinear form β is taken to be the Killing form of $\mathfrak{g}$). Varying the action by using the small perturbation $A + Bt$ (where B is a $\mathfrak{g}$-valued one-form and t is a small parameter) shows us that the Chern-Simons equation of motion is precisely the condition that A is flat, i.e.

$$F := dA + A \wedge A = 0. \tag{16}$$

Those who are used to thinking of Chern-Simons theory in the abstract form above may be somewhat perplexed by the claim that anything like this could turn out to be a gravitational theory. After all, abstract Chern-Simons theory is a topological field theory of Schwarz-type, i.e. its action and equations of motion do not require the existence of a metric, whereas gravity is manifestly a theory that is in some sense about the spacetime metric. Furthermore, since Chern-Simons theory is purely topological, how could there be a hole argument for gravity described by Chern-Simons theory? These confusions are of course contrived, but they provide a good opportunity for clarifying that it is *not* abstract Chern-Simons theory that is supposed to be identified with gravity, but rather a variant of Chern-Simons that is equipped with a Cartan connection!

Before we go on to explain exactly what a Cartan connection is and how it relates to gravity, let us digress briefly to give a sketch of the intellectual canvas into which Cartan connections fit, viz. Cartan geometry.[21] This is justified not merely by the inherent beauty of the subject, but also for the reasons that (i) it is a generalization of Klein's Erlangen program, which has long captured the philosophical imagination, and (ii)

[21] Our sketch of Cartan geometry in this section follows [24] and [29] closely.

as promised at the beginning of §3, it will give us a chance to gesture at how our story goes for theories with non-zero cosmological constant Λ.

3.2.1 Cartan geometry

First, Klein geometry. In his Erlangen program, Klein studied what we would now call homogeneous spaces, viz. a space X together with a a group of transformations G acting transitively on X. His goal was to show that the study of these geometries is equivalent to the study of certain transformation groups. This is easy enough to see if we take the viewpoint that a feature of a geometry is to be identified with the subgroup H of G that stabilizes that feature. Supposing a geometry to be entirely determined by features of this sort, the study of the geometry can be identified with the study of the space of cosets $G/H \equiv \{gH : g \in G\}$, which will often have the structure of a smooth manifold. In particular, if H stabilizes points of X, then we would expect that $X \cong G/H$. (In general, H will be indexed by the particular feature that it stabilizes, e.g. a specific point of X, but this is a superficial dependence, since all stabilizers of structurally similar features (e.g. other points of X) will be related to H by conjugation.)

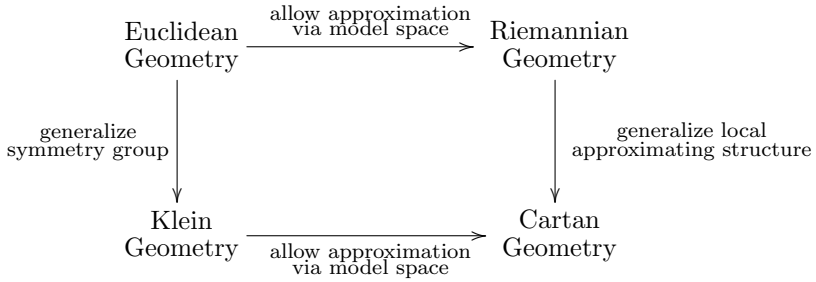
To consider an example that we have implicitly been using in §3.1, let X be the Minkowski space $\mathbb{R}^{2,1}$ and G be the connected Poincare group $ISO_0(2,1)$ that acts transitively on Minkowski space. The stabilizer of a point or 'event' of Minkowski space is the Lorentz group $SO(2,1)$, and indeed it turns out that Minkowski space is just isomorphic to the 'space of events' $ISO_0(2,1)/SO(2,1)$. (Of course, Minkowski space has other interesting geometric features, and these can be studied by forming the coset spaces of other subgroups of $ISO_0(2,1)$.)

Klein geometries are just special homogeneous spaces G/H, where we take G to be a Lie group and H to be a closed subgroup such that G/H is connected. Examples of Klein geometries (indeed metric Klein geometries) that abound in spacetime physics are the constant curvature solutions, e.g. anti de Sitter space, Minkowski space, and de Sitter space; as well as their Wick-rotated versions, i.e. hyperbolic space, Euclidean space, and spherical space.

Klein's work had a significant impact on the philosophy of geometry of his day: it led Poincare [21] to proclaim – with typical structuralist flair – that geometry is nothing but the study of a group. And Russell [23] (p. 114) was quick to appropriate the result for neo-Kantianism, by claiming that the 'form of externality' is given to us a priori by (roughly speaking) Klein geometry, whereas knowledge of the metric is determined a posteriori. But this was quickly given the lie by two advances in geometry: first, the discovery of a natural metric on many Klein geometries (which is induced from the Lie algebra); and second, the development of Lorentzian and Riemannian geometry, which are clearly not homogeneous geometries (perhaps one reason

for the relative neglect of Klein's program in comparison to Lorentzian geometry is that the latter soon emerged as the most perspicuous framework for a theory of gravity).

It is at this juncture that Cartan geometry enters the picture. Consider that Klein geometry and Riemannian geometry both generalize Euclidean geometry, but in two very different ways: Klein geometry latches onto the idea that Euclidean geometry is itself a coset manifold and generalizes the symmetry groups that define this coset manifold in order to obtain more general geometries. Riemannian geometry, on the other hand, latches onto the seemingly trivial idea that the tangent space at every point of Euclidean space is a Euclidean space, and then holds this fixed while introducing curvature into the picture, so that the flat tangent space at a point only approximates the local geometry around it.



Cartan geometry seeks to complete this 'commutative square' of approaches to geometry, as it were, by adding curvature to the Klein geometry model: in other words, just as Riemmanian geometry uses Euclidean spaces at a point to approximate a general Riemannian manifold, Cartan geometry uses the 'rolling (without slipping or twisting)' of a 'model Klein geometry' to approximate a general curved space.[22] Thus, Cartan geometry is actually a generalization of Riemannian geometry in the sense that it allows a choice of the 'model space' (beyond flat tangent spaces) that one uses to approximate a general curved manifold. (One wonders what the neo-Kantian Russell might have said in light of these developments – perhaps that the form of externality is given us by Cartan geometry? As for Poincare, he would surely have conceded that geometry is not merely the study of a group, but instead the study of a family of groups, parameterized by some space and glued together in the mathematical structure which we now call a principal bundle.)

The mathematical gadget that allows us to roll a Klein geometry along a curved manifold is called a Cartan connection, and in a moment we will explain how it works. However, let us first conclude this digression with some brief remarks about how GR models with non-vanishing cosmological constant $\Lambda \neq 0$, and the general McDowell-Mansouri construction, fit into the Cartan geometric picture. As we explained earlier, solutions of pure 3D gravity correspond to constant curvature spacetimes whose Ricci curvature is

[22] We borrow this square from [29], who in turn borrowed it from [24].

given by

$$R_{\mu\nu} = \Lambda g_{\mu\nu}.$$

Thus for $\Lambda = 0$, the model Klein geometry will be Minkowski space, whereas for Λ positive or negative, the model Klein geometries will be de Sitter and anti-de Sitter space respectively. These constant curvature spacetimes are called 'Cartan flat', meaning that the 'rolling' mentioned above is almost trivial, because a small *patch* of spacetime looks exactly like the Klein geometry that is approximating it.[23] Consequently, pure 3D gravity solutions are completely described by 'gluing together' patches of the relevant model Klein geometry, i.e. by specifying the transition functions on the overlaps of these patches.[24]

The general McDowell-Mansouri construction is somewhat more complicated than in the case of 3D gravity, precisely because the rolling can be rather non-trivial! But the overarching idea is the same: one chooses a model Klein geometry whose properties resemble the mean geometric properties of spacetime (i.e. they share the same cosmological constant), and uses this model geometry to probe spacetime.

3.2.2 Overcoming (Obst-Gauge) via Cartan connections

We now return to the main thread of our discussion, viz. how to move from abstract Chern-Simons theory to gravity. As we shall see, the main ingredient in this maneuver is to use a Cartan connection to relate the structure of spacetime to the gauge group.

Recall that abstract Chern-Simons theory with gauge group G on a spacetime M has as its basic dynamical variable a gauge field or (local) 'connection 1-form' A, i.e.

$$A : TM \to \mathfrak{g}.$$

What we need in order to approximate M with a Klein geometry of the form G/H is a *Cartan connection*, which we will now define in a restricted form before going on to generalize it to the form that we will use in the rest of the paper.

First, let us define the global version of a (restricted) Cartan connection on the principal H bundle $\pi : P \to M$. Notice that our choice of principal bundle already signals a departure from an ordinary gauge theory: each fiber of the bundle is a copy of H rather than G; it thus follows that if we choose G as our gauge group, the full gauge group cannot be modeled as a fiber of our principal bundle. A (global) Cartan connection is a $\mathfrak{g}$-valued 1-form $A : TP \to \mathfrak{g}$ such that:

[23] For a completely trivial case of 'rolling', consider rolling a Klein geometry on itself, e.g. rolling the tangent spheres of S^n along itself – clearly the only thing that changes is the point of tangency of the sphere. For a more non-trivial case of 'rolling', consider rolling a 2-sphere along a 2-manifold.

[24] This is esssentially the 'theory of geometric structures'. See Chapter 4 of [8] and the references therein.

- $A : T_pP \to \mathfrak{g}$ is a linear isomorphism for all $p \in P$
- A is H-equivariant, i.e. $R_h^* A = Ad(h^{-1})A$ for all $h \in H$.
- A takes values in $\mathfrak{h}$ on vertical vectors, i.e. vectors along the fibers of P.[25]

Locally, the global connection A (on P) is manifested as a gauge field, also called A (on M) for which the composite

$$TM \xrightarrow{A} \mathfrak{g} \xrightarrow{\pi} \mathfrak{g}/\mathfrak{h} \tag{17}$$

(where π is the canonical projection to the quotient $\mathfrak{g}/\mathfrak{h}$) restricts to a linear isomorphism at each point of M.

Let us use e to denote the composite $\pi \circ A$ – evidently this is none other than the generalization of the dreibein 1-form e that we discussed in the previous section, where we set $\mathfrak{g}/\mathfrak{h} \equiv \mathfrak{iso}(2,1)/\mathfrak{so}(2,1) \cong \mathbb{R}^{2,1}$ as the 'fake tangent bundle'! Intuitively, e should be thought of as a gadget that 'solders' the tangent space of spacetime M to the tangent space of the Klein geometry G/H.

We thus see how the extra data provided by a Cartan connection allows us to overcome (Obst-Gauge): the gauge symmetry now comprises Lorentz transformations and translations, and we know that these symmetries act on spacetime points because of the isomorphism given by e. In other words, the connection and gauge symmetries are intimately related to the structure of spacetime itself, and do not merely act on some 'internal space'. Nonetheless, a cautious reader might object: was it not the case that a gauge transformation was described as a vertical automorphism in (Obst-Gauge), and if so, how can we square this with the 'translation' gauge transformations of Cartan gauge theory? The answer is simple enough: every Cartan gauge theory with gauge group G can be reformulated as a principal G-bundle with connection (i.e., a standard gauge theory), for which the fiber really is the gauge group G and whose vertical automorphisms are the gauge transformations.[26] However, Cartan gauge theory is a particularly [**?**] formalism with which to understand how (Obst-Gauge) is overcome, because it makes manifest the relationship between the gauge symmetries and spacetime.

Since e is a linear isomorphism in the above definition, it will allow us to pull back the $SO(2,1)$ invariant metric of Minkowski space to obtain a *nondegenerate* Lorentzian metric on M. However, we can just as well allow e to be non-invertible, in which case this procedure only defines a *degenerate* or singular Lorentz metric on M. Furthermore, it would seem that from the purely gauge-theoretic point of view, there is no

[25] See p. 16 of [29] for more on how A is in this way related to the Maurer-Cartan form of H.

[26] We refer the reader to p. 21 of [29] for an explanation of this result, and for the technical condition on the Ehresmann connection (of the principal G-bundle) that makes this true.

reason to ask that e be invertible; indeed, in many applications to quantum gravity, it is important for e to be non-invertible. Thus, let us now define a (generalized) Cartan geometry as including non-invertible e. We shall return to this point in §4; but till then, we shall be understood as working with Cartan geometry in this generalized sense.

Earlier we said that the Cartan connection allows us to approximate M by 'rolling' G/H along it. It will be instructive to work out what this amounts to in our case of interest, viz. where $G = ISO(2,1)$ and $H = SO(2,1)$ and so our 'model Klein geometry' is none other than Minkowski space. The Cartan geometry that describes this situation is a principal $SO(2,1)$ bundle

$$\pi : P \to M$$

along with a Cartan connection

$$A : TP \to \mathfrak{iso}(2,1).$$

The reader should think of P as describing an $ISO(2,1)/SO(2,1) \cong \mathbb{R}^{2,1}$ space 'rolling' around on M. A point of P consists of (i) the position of an $\mathbb{R}^{2,1}$ model space on M, and (ii) the 'Lorentz state' of the model space, which is given by an element of $SO(2,1)$.

To adapt a cute illustration due to Wise [29], we can think of this point as specifying the configuration of a 'Lorentz hamster' who is in some particular 'Lorentz state' as he stands on the model space $\mathbb{R}^{2,1}$ at a point of M. Our intrepid special relativistic hamster explores M by stepping forward on the model space, and thus rolling it while he remains perched atop its point of tangency to M. This can be formalized by saying that an infinitesimal change to the hamster's configuration (i.e. an element of TP) consists of (i) a 'transvection', i.e. a tiny pure translation of the point of tangency, and (ii) a tiny Lorentz transformation that changes the 'Lorentz state' of the hamster. We are now in a position to get clear on exactly what role the Cartan connection A plays: it takes as its input infinitesimal changes in the hamster's configuration, and produces as its output an infinitesimal $ISO(2,1)$ transformation. But since $ISO(2,1)$ is the symmetry group of $\mathbb{R}^{2,1}$, the output is simply telling us how the model space changes (infinitesimally) as our hamster explores M.

We have just seen the geometric background that underlies the heuristic constructions of §3.1. In order to say exactly how (Obst-Gauge) is overcome, however, we will need to note a special feature of the model Klein geometry $ISO(2,1)/SO(2,1) \cong \mathbb{R}^{2,1}$. First, we remind the reader that for a matrix Lie group G, the Ad representation of G acting on an element x of its Lie algebra $\mathfrak{g}$ is given by: $\mathrm{Ad}(g)x := gxg^{-1}$.[27]

[27]For a general Lie group G, one needs to first define the conjugation map $\phi(h) = ghg^{-1}$, where $g, h \in G$. $\mathrm{Ad}(g)$ is then defined as the push-forward ϕ_*.

The special feature of $ISO(2,1)/SO(2,1)$ that we are interested in is that it is a *reductive* Klein geometry, meaning (for our purposes) that the Lie algebra $\mathfrak{iso}(2,1)$ has the $\mathrm{Ad}(SO(2,1))$-invariant splitting

$$\mathfrak{iso}(2,1) = \mathfrak{so}(2,1) \oplus \mathbb{R}^{2,1},$$

and so the quotient algebra $\mathfrak{iso}(2,1)/\mathfrak{so}(2,1)$ can be identified with the subalgebra $\mathbb{R}^{2,1}$. The importance of having such an *invariant* splitting is that it ensures that even *globally*, this decomposition of the Lie algebra isn't 'mixed up' by $SO(2,1)$ transformations.

It follows that we can uniquely and globally decompose the Cartan connection A as

$$A = \omega + e,$$

where ω is an $\mathfrak{so}(2,1)$-valued 1-form on P and e is an $\mathbb{R}^{2,1}$-valued 1-form on P. In physics parlance, we would say that w takes values in the Lorentz generators, while e takes values in the momentum generators. Evidently, the local versions of w and e are precisely the spin connection and dreibein of §3.1 respectively. We thus see precisely how setting the gauge group of the theory to $ISO(2,1)$, and incorporating e and ω into the gauge field A, allows us to overcome (Obst-Gauge): ω plays the role of a $SO(2,1)$ gauge field on an abstract vector bundle, and e solders this bundle to the tangent space of M.

Finally, let us revisit from a more systematic point of view the issue of how to satisfy (Trace). Given the form of the action $CS(A)$ (15), we know that the trace function β has to be a nondegenerate bilinear form on $\mathfrak{iso}(2,1)$ and so we may be tempted to take it to be the familiar Killing form. However there are two problems with this. The first is that the Killing form is only unique for simple Lie algebras, and $\mathfrak{iso}(2,1)$ is not simple. In fact, it is easy to show that there exists a two-dimensional family of invariant inner products on $\mathfrak{iso}(2,1)$. Viz., a general inner product will be a linear combination of

$$\mathrm{tr}^\star(XY) = K(X \star Y) \tag{18}$$

and

$$\mathrm{tr}(XY) = K(XY), \tag{19}$$

where K is the usual Killing form.

Second, and more importantly, tr turns out to be *degenerate* when $\Lambda = 0$ and so cannot serve as a good inner product for obtaining the Einstein-Hilbert action from the Chern-Simons action.[28] If we instead let $\beta = \mathrm{tr}^\star$ and let $A = \omega + e$ (compare (12) and (14)), then we indeed obtain the Einstein-Hilbert action (2),

[28]Note however that $\mathrm{tr}^\star$ and tr are both non-degenerate in the $\Lambda \neq 0$ case – in fact, they lead to two different actions that have the same classical equations of motion, but have different quantizations!

thus showing that we have satisfied (Trace). We are now back in Witten's [30] footsteps, celebrating the 'minor miracle' (as we called it at the end of §2) that in three dimensions there is a non-degenerate inner product on $\mathfrak{iso}(2,1)$.

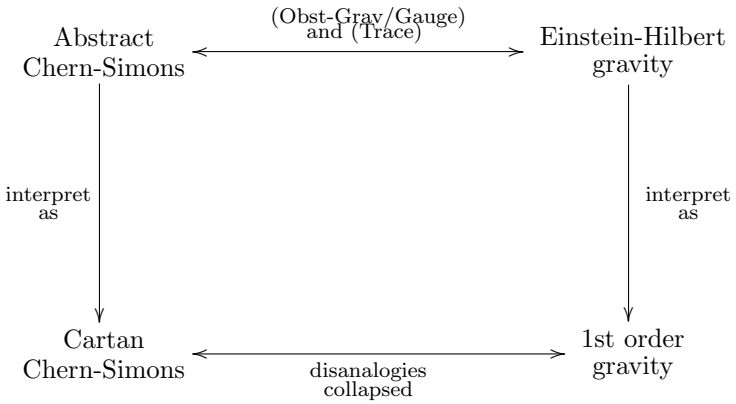
But why is this a miracle at all? Surely it would only be a miracle if such an inner product was unattainable in other dimensions, and indeed this is the case: if we were considering $\mathfrak{iso}(n,1)$ for arbitrary n, a Lorentz-invariant bilinear form would in general have to take the form

$$mJ_{ab}J^{ab} + nP_aP^a, \tag{20}$$

where m and n are constants. But since the inner product must commute with P_a, this forces $m = 0$ and so the inner product will be degenerate. The beauty of working in three dimensions is that, by virtue of Hodge duality, we can write a nondegenerate Lorentz-invariant inner product as (essentially): $\varepsilon_{abc}P^aJ^{bc}$!

3.3 Disanalogies collapsed

Let us take stock. We began by remarking on two disanalogies between gravity and gauge theories: one (Obst-Grav/Gauge) that is familiar from the philosophy of physics literature, and another (Trace) that is less familiar. We then showed how these two gaps could be closed, from the perspective of both gravity and gauge respectively. This progression can be summed up in the following diagram:



On the upper left and upper right-hand sides of the square, we have two theories – abstract Chern-Simons and Einstein-Hilbert gravity – that are prima facie very different. At least some of those differences are captured by (Obst-Grav/Gauge) and (Trace). By interpreting Einstein-Hilbert gravity as first-order gravity, on the one hand, and by interpreting abstract Chern-Simons theory as a Cartan gauge theory, on the other, we were able to overcome (Obst-Grav/Gauge) and (Trace).

It is noteworthy that (Obst-Grav/Gauge) was not overcome by rewriting all the small diffeomorphisms of gravity in terms of the formalism that we usually interpret as a gauge group, viz. a fiber of a principal bundle. Rather, the strategy that we used was somewhat more subtle than that: it turned on interpreting a gauge theory in such a way that only *some* of the $ISO(2,1)$ gauge symmetry needed to be written as a fiber of a principal bundle (viz. the $SO(2,1)$ subgroup of the gauge symmetry), whereas the rest could still be incorporated into the gauge field (and thus legitimately be said to be part of the gauge group), while playing the role of relating the fiber of the principal bundle to spacetime!

However, we still have not shown that Cartan Chern-Simons and first-order gravity are *identical* as theories. In order to show this, we would have to show that (Obst-Grav/Gauge) and (Trace) were the *only* dianalogies between gravity and gauge. But as we shall argue in the second half of this paper, there remain two, more subtle, disanalogies between gravity and gauge.

4 Two more disanalogies

This section discusses two more disanalogies between gravity and gauge, which were already hinted at in the preceding sections. The first stems from the difference between large gauge transformations and large diffeomorphisms. The second stems from including noninvertible dreibeins e as solutions to the gauge theory; (as we shall see, there is considerable pressure to do so). Unlike the disanalogies of the previous section, we shall argue that the first cannot in general be collapsed, and the second can only be collapsed at a considerable price. Furthermore, while the first disanalogy and the second are conceptually distinct, it will emerge that the first actually contrains the ways in which one might seek to escape the second disanalogy.

Earlier, we asked the reader to take it on faith that the small diffeomorphisms of gravity could be described by small gauge transformations. In §4.1, we explain why this is so and proceed to elaborate upon the notion of *large* gauge transformations as well as large diffeomorphism transformations. We shall argue that prima facie, the large gauge transformations do not coincide with the large diffeormorphisms; this then constitutes a serious disanalogy between gravity and gauge theory.

§4.2 discusses the relationship between the invertibility of e, on the one hand, and the equivalence of the respective phase spaces of first-order gravity and Cartan Chern-Simons, on the other hand. We shall argue that if the solutions of Cartan gauge theory include non-invertible e then the phase space of Cartan gauge theory and the phase space of first-order gravity are manifestly non-equivalent. Furthermore, we shall argue that this non-equivalence cannot be repaired by means of large symmetry transformations. If, on the other hand, the solution space should be restricted to only invertible e, then the phase spaces turn out to be

equivalent. However, we shall argue that this option can only be entertained at a considerable theoretical price – one that vitiates an important motivation for trying to identify gravity with a gauge theory in the first place.

4.1 What about the symmetries?

We shall consider the small gauge symmetries of Cartan Chern-Simons theory and argue that they coincide with the small diffeomorphisms of a gravity theory.

Recall that for a gauge theory with gauge field A, an infinitesimal symmetry of A is given by

$$\delta A = d_A u, \tag{21}$$

where u is a Lie algebra-valued scalar parameter, and the covariant derivative is taken in the adjoint representation, and so $d_A u = du + [A, u]$. (To calculate δA, one just plugs the infinitesimal gauge transformation $\exp(\alpha) \sim (1 + \alpha)$ into the formula for a finite gauge transformation of A.)

In the case of Cartan Chern-Simons theory, A has a global decomposition as $A = e^a P_a + \omega^a J_a$ and so our Lie algebra-valued parameter u takes the form:

$$u = \rho^a P_a + \tau^a J_a. \tag{22}$$

We thus see that the infinitesimal symmetry of A decomposes into infinitesimal symmetries of e and ω respectively:

$$\delta e = d\rho + [e, \tau] + [\omega, \rho], \tag{23}$$

$$\delta \omega = d\tau + [\omega, \tau]. \tag{24}$$

By using the usual Lie algebra bracket relations of $\mathfrak{iso}(2, 1)$ we can reinterpret this as saying that the infinitesimal Lorentz transformations are given by

$$\delta e = [e, \tau], \qquad \delta \omega = d\tau + [\omega, \tau], \tag{25}$$

whereas the infinitesimal translations are given by

$$\delta e = d\rho + [\omega, \rho]. \tag{26}$$

How are these infinitesimal gauge symmetries related to the infinitesimal diffeomorphism symmetries, which are given by the Lie derivative of a gauge field with respect to a vector field that generates that

particular diffeomorphism? The answer is simple enough: by making a judicious choice of $\mathfrak{iso}(2,1)$-valued infinitesimal parameters ρ and τ, we can show that these infinitesimal symmetries in fact coincide.

To see this, we begin by calculating the Lie derivatives of e and ω with respect to some vector field X. We obtain

$$\mathcal{L}_X e = d\iota_X e + [\omega, \iota_X e] + [e, \iota_X \omega], \tag{27}$$

where we have used (i) Cartan's magic formula $\mathcal{L}_X = d\iota_X + \iota_X d$, and the vanishing of the torsion (i.e. Cartan's first structural equation (9)), and (ii)

$$\mathcal{L}_X \omega = d\iota_X \omega + [\omega, \iota_X \omega], \tag{28}$$

for which we have again used Cartan's magic formula and the vanishing of the curvature (Cartan's second structural equation (10)).

By making the choice of infinitesimal parameters $\rho = \iota_X e$ and $\tau = \iota_X \omega$, we immediately see that (23) coincides with (27), and that (24) coincides with (28). Thus, the infinitesimal symmetries of Cartan gauge theory and first-order gravity are equivalent!

We have thus far been discussing infinitesimal symmetries (i.e. symmetries at the Lie algebraic level), and so our discussion also applies to all those *finite* symmetries that can be built up by exponentiating or 'integrating' infinitesimal symmetries: these finite symmetries are what we call the *small* symmetries. Viewed topologically, the small symmetries form the set of points in the space of symmetries that can be smoothly 'retracted' to the identity symmetry.

However, there is yet another class of (finite) symmetries, called the *large* symmetries, which are defined as those symmetries that are *not* small. In the case of Cartan Chern-Simons theory, they are called the large gauge transformations; and in the case of first-order gravity, they are called the large diffeomorphisms. Clearly, the argument used to prove the equivalence of small gauge transformations and small diffeomorphisms above will not apply to large symmetries. Indeed, nothing we have said so far constrains the large gauge transformations to (even partially) coincide with the large gauge transformations! This leads us to the first of our more serious and subtle disanalogies between gravity and gauge:

> (Large) The large gauge transformations of a gauge theory will in general be different from the large diffeomorphisms of gravity.[29]

[29] For a rigorous proof of this fact, see pp. 7-8 of [18], which shows that all (finite) local translations are small in Cartan Chern-Simons theory, whereas there exist *large* local translations in first-order gravity.

This insight can be seen as a rather more incisive version of the basic intuition that the gauge group of a gauge theory and the diffeomorphism group of gravity are very different beasts. That intuition was already operative in §2's formulation of (Obst-Grav/Gauge); but as we have seen, if one wants to formulate a really strong form of the disanalogy that holds even in (2+1) dimensions, it is insufficient to discuss small symmetries, as most theorists do.

Is it possible to make the large gauge transformations and the large diffeomorphisms coincide by fiat? That is to say, will simply laying it down that they give equivalent transformations generate any inconsistencies? Insufficient work has been done on this topic, and we hope to pursue it in a future investigation. But preliminary work ([18]) on adding large gauge transformations and large diffeomorphisms to the reduced phase space has already shown that certain consistency constraints have to be respected.

Before we go on to discuss the disanalogy that stems from the non-invertibility of e, let us pause briefly to reflect on the implications of large diffeomorphisms for what is commonly known as the 'hole argument'.[30] In the usual presentations of the hole argument, an active (small) diffeomorphism is made in some region of spacetime that lies to the future of some initial data hypersurface. It is then argued that if the diffeomorphed region (and the corresponding pullback of the metric) is not agreed to be physically equivalent to the original region, then we end up with a case of indeterminism. This argument can be re-run with *large* diffeomorphisms instead of small diffeomorphisms, and the nice feature that results from doing so is that, since a large diffeomorphism applies to the entire spacetime, what we end up with is not merely a case of indeterminism, but rather of underdetermination.

4.2 The phase spaces of the two theories

In Section §3.2, we noted that if one starts with a Cartan Chern-Simons theory whose basic dynamical variables are e and ω (combined into one gauge field A), there seems to be no good reason ('internal to the gauge theory') to require e to be invertible; we thus professed, in our formulation of Cartan Chern-Simons theory, to be agnostic about whether e was invertible (i.e. to use the generalized definition of Cartan geometry). Of course, if one is interested in trying to make the gauge theory equivalent to gravity then a reason might simply be: standard gravity uses a non-degenerate metric, and a non-degenerate metric requires an invertible dreibein! In a moment, we shall see that this move can be questioned, but in order to obtain a clearer view of the space of possible moves available to us here, let us now consider the phase spaces of Cartan Chern-Simons and first-order gravity respectively.

First some terminology: we shall take the configuration space Q of a field theory to be the space of its

[30] For an introduction to the 'hole argument' literature, see [**?**] and the references therein.

basic dynamical fields that live on a *space-time* manifold M.[31] In the case of Cartan Chern-Simons and first-order gravity, the respective configuration spaces Q_{CS} and Q_{GR} are the space of all pairs (e, ω) that are admissible in the theory. A configuration space Q contains a submanifold P of fields that satisfy the equations of motion of the theory. Now, a fruitful way to think of a theory's phase space is simply as the space of solutions to the theory's equations of motion. Thus, let us write P for phase space: P_{CS} in the case of Cartan Chern-Simons, and P_{GR} in the case of gravity.

Both P_{CS} and P_{GR} are infinite-dimensional. But this fact is misleading. It is misleading because the solutions (e, ω) that are connected by symmetries are considered to be 'physically equivalent' in each of these theories. The beauty of gravity in three dimensions is that the number of symmetries is so large relative to the solution space that once we have 'quotiented out' by these symmetries, we obtain a *finite-dimensional* phase space, called the reduced phase space R.

Let us now explain this idea in detail for small symmetries. The phase spaces P_{CS} and P_{GR} are both equipped with a canonical pre-symplectic structure that is defined by the action. They thus carry a canonical pre-symplectic 2-form Ω_P that is highly degenerate precisely because the solutions in P are transformed into 'physically equivalent' solutions by small symmetries. In other words, there is a large amount of redundancy in P, which is encoded in the degeneracy of Ω_P. If a small symmetry connects two solutions ϕ_1 and ϕ_2, then ϕ_1 and ϕ_2 are joined by a curve whose tangent vectors X lie in the kernel of Ω_P, i.e.

$$\Omega_P(X, \cdot) = 0.$$

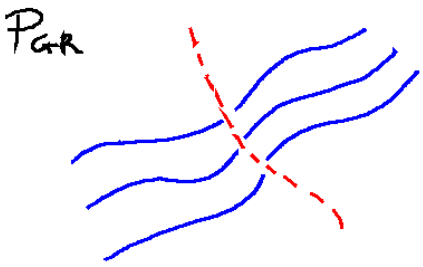
We shall call the equivalence class of solutions connected by such a curve a *null orbit*.[32] Of course, two solutions may also be physically equivalent because they are connected by a *large* symmetry, in which case a physical equivalence class may comprise not just one, but several null orbits. In order to go from the phase space P to the reduced phase space R, i.e. the space of physically equivalent solutions, we typically quotient out or 'reduce' the phase space P by both large and small symmetries. If all goes well, R will turn out to be a manifold that carries a unique symplectic (i.e. closed and *non-degenerate*) 2-form Ω_R.

Here is a cartoon of the null orbits in P_{GR}, which (by definition) does not include the manifold of non-invertible dreibein e (indicated by the dashed line), which we call the 'singular submanifold' for short (because the metric is singular on it):

In two-dimensions, it is clear that the singular submanifold will *separate* the space of null orbits into two *disconnected* components. However, this picture is misleading, because the space of fields is actually infinite-

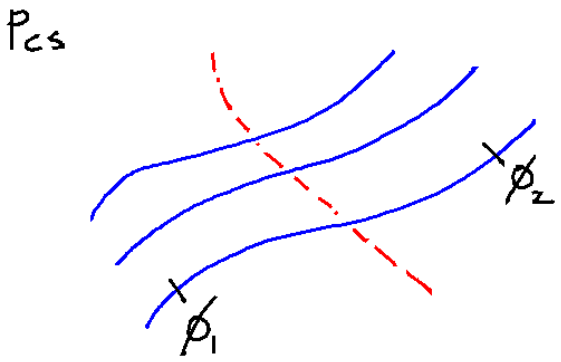
[31] Note that there is also a different and widespread (especially within the context of mechanics) use of configuration space, viz. the space of instantaneous state of the theory.

[32] This characterization of null orbits is merely heuristic. For a rigorous treatment, see Chapter 3 of [14].

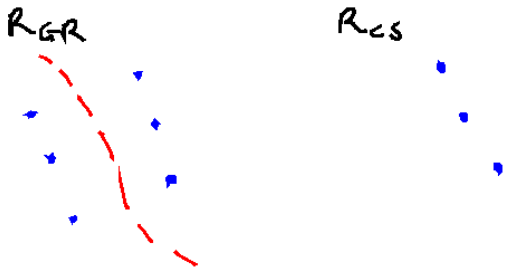


dimensional! Nonetheless, Matschull [18] has given a general argument that what we intuitively expect from the two-dimensional picture happens to also be true in the real, infinite-dimensional case.

What of the null orbits in P_{CS}? This is where our choice of whether or not to include non-invertible e in Cartan Chern-Simons theory becomes crucial. If we choose not to include the singular submanifold, then clearly P_{CS} is equivalent to P_{GR}. If, on the other hand, we include the singular submanifold, then P_{CS} looks as follows:



Clearly, P_{CS} with non-invertible dreibein differs from P_{GR}: P_{GR} did *not* include the 'points' of the dashed line (indeed its orbits were disconnected by the dashed line), whereas these 'points' are part of P_{CS} . This difference might not be thought to be so vast as they apparently only differ by a 'point' on each null orbit. However, in order to fully appreciate the force of this disanalogy between gravity and gauge, we should quotient out by small gauge/diffeomorphism transformations in order to obtain the (partially) reduced phase spaces of the respective theories:



We thus see that the (partially) reduced phase space R_{GR} now has twice as many physical states as R_{CS}, and this is a dramatic difference indeed!

What moves are available to a theorist who wishes to include non-invertible e in the solutions of Cartan Chern-Simons, and yet identify it with first-order gravity? One obvious avenue presents itself: the reason that the above phase spaces were only partially, and not fully, reduced is that we did not take into account the action of large symmetries on the solutions. Is it then possible that we might be able to use either the large gauge symmetries or the large diffeomorphisms to restore the equivalence of the reduced phase spaces? There are three relevant cases to consider:

- Case 1: We use large gauge transformations (more specifically: the finite local translations that are 'small' in Cartan Chern-Simons but 'large' in first-order gravity, due to the singular submanifold) to connect the null orbits in P_{GR} that are disconnected by the singular submanifold. If we now only consider a reduction of P_{GR} by small symmetries and these large gauge tranformations, this will result in a reduced phase space R_{GR} that is equivalent to R_{CS}.
- Case 2: We use large diffeomorphisms to connect the null orbits in P_{CS}. However, upon reduction, this does not bring us any closer to identifying R_{CS} with R_{GR}.

- Case 3: We add *both* the large gauge transformations and the large diffeomorphisms to P_{GR} and P_{CS}, after which we reduce both phase spaces by these (large and small) symmetries. This does give us equivalent reduced phase spaces for both theories, but as Matschull [18] convincingly argues, for a large class of static spacetimes, this move results in a gauge orbit that fills the plane densely, and thus the reduced phase space is no longer a manifold (indeed it is not even Hausdorff)!

From our discussion of the three cases, it is clear that only Case 1 might be a viable strategy: it results in a reduced phase space R_{GR} that is equivalent to R_{CS}, and furthermore, this reduced space is not pathological, as with Case 3.

However, upon closer examination, we see that this strategy is not really satisfactory for two reasons. First, Matschull [18] has shown that for a large class of static spacetimes, including these large local translations as symmetries of first-order gravity forces us to identify non-diffeomorphic spacetimes: this contravenes one of the fundamental assumptions of general relativity. Second, even if one is willing to bite that bullet, one must still take account of the fact that one has only reduced the phase space P_{GR} of first-order gravity by small symmetries and large gauge transformations. Thus, one still has to specify how to treat the large diffeomorphisms (which, we remind the reader, will in general be different from the large gauge transformations). There are two possibilities: (i) deny that one needs to quotient out by large diffeomorphisms to obtain the reduced phase space, or (ii) accept that solutions connected by large diffeomorphisms should be treated as physically equivalent. If one adopts (i), then one is in a 'hole argument' scenario, in which we have two putatively 'physically distinct' solutions connected by a large diffeomorphism, although we cannot in fact identify any physical difference between them. On the other hand, if one adopts (ii) then one is back in the situation of Case 3 (i.e. that of combining large diffeomorphisms with large gauge transformations) where the reduced phase space turns out to be pathological for a large class of examples. Both options are undesirable.

Ostensibly, including the singular submanifold in P_{CS} has led us to a dead end. Let us now return to the idea that the singular submanifold should be omitted from P_{CS} and see if we do any better.

If we view both gauge and gravity from a purely classical point of view, then P_{CS} (with the singular submanifold omitted) and P_{GR} will be equivalent, and so their reduced phase spaces (reduced by the small symmetries and the large diffeomorphisms) will also be equivalent. However, there are two problems with taking this tack. First, we have to omit the large gauge transformations from the phase space reduction if we are to avoid the plight of Case 3 above. And if we omit the large gauge transformations, then arguably this version of Cartan Chern-Simons is not really a gauge theory in the full sense of the word (since there will exist large gauge transformations between physically distinct states).

A second problem with omitting the singular submanifold from P_{CS} is that one of the most important motivations for identifying gravity with gauge, viz. learning how to quantize 3D gravity, is potentially vitiated by this move. As Witten points out in his [30], it is somewhat paradoxical that folklore (at least in the 1980s) seemed to hold that (i) gravity is unrenormalizable in $(2+1)$ dimensions, and (ii) 3D gravity is trivial, in the sense of having a finite-dimensional phase space. Witten goes on to resolve this paradox by arguing that non-invertible e must be allowed in order to make sense of the quantum theory, i.e. to prove renormalizability. Thus, from the point of view of [30], the move of omitting the singular submanifold from P_{CS} defeats the project of exploiting Chern-Simons theory in order to quantize 3D gravity.

Let us for the sake of argument suppose that the point of view adopted in [30] is the right approach to quantizing 3D gravity. We might then say that the quantum theory is telling us that the conventional interpretation of classical gravity says something false about the world, viz. that we should think of the metric as being everywhere non-degenerate. (Compare this with the Aharonov-Bohm effect, in which quantum mechanics coupled to classical electromagnetism tells us that the conventional interpretation of classical electromagnetism is saying something false when it claims that the fundamental dynamical quantity is the electromagnetic field.[33] Rather, the fundamental quantity is the holonomy of the gauge field, or equivalently, a point in the reduced moduli space of gauge fields.) From this point of view (that is to say, when we interpret classical gravity in light of quantum gravity), to omit the singular submanifold from P_{CS} is to commit ourselves to a false interpetation of 3D gravity!

Admittedly, the above point of view is not the only approach to quantizing 3D gravity. Indeed, in [32], Witten considers the view that Chern-Simons theory is not after all the right approach to quantizing 3D gravity, because it does not have sufficient degrees of freedom to make sense of the BTZ black hole. Furthermore, other considerations from the string theory perspective indicate that we should *exclude* non-invertible dreibein when we are quantizing gravity! It is beyond the scope of this paper to resolve these issues; here we can only point out that what turns out to be an aspect of the disanalogy between gravity and gauge is in fact subtly related to interpretive problems in quantum gravity. We hope to further clarify these relations, and resolve some of these problems, in a future investigation.

5 Summary and conclusion

In this paper, we have explained how the most frequently remarked disanalogy between gravity and gauge theory (i.e. the thought that the (small) gauge symmetry of gravity acts on spacetime, whereas the 'internal'

[33] Belot makes a similar point in [2].

gauge symmetry of gauge theory does not) can be collapsed, viz. by relating the 'internal' gauge symmetry to spacetime by means of a 'Cartan connection'. Furthermore, it turns out that, unlike other dimensions, three-dimensional space-time has special properties that allows one to write down the standard gravitational action as a gauge theory (Chern-Simons) action. Thus, prima facie, one might well be optimistic about the prospects for identifying the two theories in three dimensions! But perhaps surprisingly, yet more subtle obstacles to making this identification remain, such as the fact that (i) the large symmetries of the respective theories do not coincide, and (ii) the gauge-equivalent field configurations of the respective theories will in general be different. As we have argued above, the prospects for removing these obstacles are dim, and so we can be reasonably confident that, even in this best case scenario, gravity is not *identical* to a gauge theory. Even more important than this conclusion, however, is the possibility that grappling with such issues might lend us insight into the conceptual problems of quantum gravity.

References

[1] A. Achúcarro and P.K. Townsend, *A chern–simons action for three-dimensional anti-de sitter supergravity theories*, Phys. Lett. B **180** (1986), 89–92.

[2] Belot, *Understanding electromagnetism*, BJPS **49 (4)** (1998), 531–555.

[3] Belot and Earman, *From physics to metaphysics*, In From Physics to Philosophy, ed. Jeremy Butterfield & Constantine Pagonis. (1999).

[4] ———, *Pre-socratic quantum gravity*, Philosophy Meets Physics at the Planck Scale, CUP, 2001.

[5] Gordon Belot, *Whatever is never and nowhere is not*, Ph.D. thesis, University of Pittsburgh, 1996.

[6] C. Brans and R. H. Dicke, *Mach's Principle and a Relativistic Theory of Gravitation*, Physical Review **124** (1961), 925–935.

[7] J. D. Brown and Marc Henneaux, *Central charges in the canonical realization of asymptotic symmetries: An example from three dimensional gravity*, Communications in Mathematical Physics **104** (1986), 207–226, 10.1007/BF01211590.

[8] Steve Carlip, *(2+1) quantum gravity*, CUP, 2003.

[9] Oscar Chacaltana and Jacques Distler, *Tinkertoys for gaiotto duality*, Journal of High Energy Physics **2010** (2010), 1–54, 10.1007/JHEP11(2010)099.

[10] S. Deser, R. Jackiw, and G. 't Hooft, *Three-dimensional einstein gravity: Dynamics of flat space*, Ann. Phys. (N.Y.) **152** (1984), 220–235.

[11] M. J. Duff, *Kaluza-Klein theory in perspective*, (1994).

[12] Green, Schwarz, and Witten, *Superstring theory 2*, CUP, 1987.

[13] Richard Healey, *Gauging what's real*, OUP, 2007.

[14] Marc Henneaux and Claudio Teitelboim, *Quantization of gauge systems*, PUP, 1994.

[15] Gary T. Horowitz and Joseph Polchinski, *Gauge / gravity duality*, (2006).

[16] S. L. Lyakhovich and A. A. Sharapov, *Quantizing non-Lagrangian gauge theories: an augmentation method*, Journal of High Energy Physics **1** (2007), 47–+.

[17] S. W. MacDowell and F. Mansouri, *Unified geometric theory of gravity and supergravity*, Phys. Rev. Lett. **38** (1977), no. 14, 739–742.

[18] H.-J. Matschull, *On the relation between 2+1 Einstein gravity and Chern-Simons theory*, Classical and Quantum Gravity **16** (1999), 2599–2609.

[19] C. Meusburger and B. J. Schroers, *Poisson structure and symmetry in the Chern-Simons formulation of (2+1)-dimensional gravity*, Class. Quant. Grav. **20** (2003), 2193–2234.

[20] Georgios Papageorgiou and Bernd J. Schroers, *A Chern-Simons approach to Galilean quantum gravity in 2+1 dimensions*, JHEP **11** (2009), 009.

[21] H. Poincare, *Sur les hipothses fondamentales de la gomtrie*, Bulletin de la Socit mathmatique de France **15** (1887), 203–216.

[22] Finn Ravndal, *Scalar gravitation and extra dimensions*, arXiv:gr-qc/0405030v1 (2004).

[23] B. Russell, *An essay on the foundations of geometry*, CUP, 1897.

[24] R.W. Sharpe, *Differential geometry: Cartan's generalization of klein's erlangen program*, Springer, New York, 1997.

[25] A. Staruszkiewicz, *Gravitation theory in three-dimensional space*, Acta Phys. Pol. **6**, 734.

[26] Norbert Straumann, *General relativity with applications to astrophysics*, Springer Berlin / Heidelberg, 2004.

[27] David Wallace, *Time-dependent symmetries: the link between gauge symmetries and indeterminism*, ch. 9, pp. 163–174, CUP, 2003.

[28] Steven Weinstein, *Gravity and gauge theory*, Philosophy of Science **66 (3)** (1999), 155.

[29] Derek K Wise, *Macdowellmansouri gravity and cartan geometry*, Classical and Quantum Gravity **27** (2010), no. 15, 155010.

[30] Edward Witten, *2 + 1 dimensional gravity as an exactly soluble system*, Nuclear Physics B **311** (1988), no. 1, 46 – 78.

[31] Edward Witten, *Quantum field theory and the Jones polynomial*, Comm. Math. Phys. **121** (1989), no. 3, 351–399. MR 990772 (90h:57009)

[32] ———, *Three-Dimensional Gravity Revisited*, (2007).

[33] ———, *Geometric Langlands From Six Dimensions*, (2009).