Gravity as a diffeomorphism invariant gauge theory

Kirill Krasnov
School of Mathematical Sciences
University of Nottingham

January 2011

Abstract

A general diffeomorphism invariant SU(2) gauge theory is a gravity theory with two propagating polarisations of the graviton. We develop this description of gravity, in particular for future applications to the perturbative quantisation. Thus, the linearised theory, gauge symmetries, gauge fixing are discussed in detail, and the propagator is obtained. The propagator takes a simple form of that of Yang-Mills theory with an additional projector on diffeomorphism equivalence classes of connections inserted. In our approach the gravitational perturbation theory takes a rather unusual form in that the Planck length determined from the self-coupling of the graviton is no longer fundamental but becomes a derived quantity.

1 Introduction

The field theory approach to gravity, see e.g. [1], tells us that gravity is not a gauge theory. Indeed, the carriers of force in a gauge theory (such as e.g. Maxwell electrodynamics) are spin one particles. For this reason there are two types of charged objects interacting by exchange of carriers of force - those negatively and positively charged. Like particles repel and unlikes attract. In contrast, there is only one type of charge in gravity and everything attracts everything. Thus, gravity is not a gauge theory, see [1] for a further discussion.

This simple argument forbids a direct gauge theory description of gravity. It says nothing, however, about less direct possible relationships. And indeed, a relation of a completely different type is now being very popular. This has its origin in the open-closed string duality, which implies that amplitudes for closed strings are squares of those for open strings. Since the low energy limit of the closed string theory is gravity, and that for open strings is gauge theory, this implies that scattering amplitudes for gravitons must be expressible as squares of amplitudes for gluons, see e.g. a recent paper [2] and references therein. The relationship is not direct, and it is in particular not easy to find a Lagrangian version of the correspondence. However, it has recently led to some very interesting developments on loop divergences in $N = 8$ supergravity.

The aim of the present paper is to develop yet another relationship between gravity and gauge theory. To the best of this author's understanding, there is no relation between the present story and that of [2]. The relationship of interest for us here has its origins in the discovery of Plebanski [3] that certain triple of self-dual two-forms can be used as the basic variables for gravity[1]. The same "self-dual" formulation of general relativity (GR) has been rediscovered a decade later by Ashtekar [5] via a completely different path of a canonical transformation on the phase space of GR. The two discoveries were later linked in [6], and the outcome was a realisation that gravity can be reformulated as a theory whose phase space coincides with that of an SU(2) gauge theory. This gravity/gauge theory relationship was taken one step further in [7]. Thus, it was realised that the two-form fields of

[1]Similar ideas has appeared in the literature much earlier, see e.g. [4] for historical remarks, but it was Plebanski who proposed to reformulate general relativity without the metric, with only two-forms as dynamical variables.

Plebanski formulation of GR [3] can be integrated out to obtain a "pure connection" formulation of general relativity, where the only dynamical field is an SU(2) connection. The result was a completely new perspective on general relativity, in which GR becomes reformulated as a novel type of a theory of the gauge field — a *diffeomorphism invariant gauge theory.*

The work on "pure connection" formulation of GR [7] has led to some further advances in that it was realised in [8] that there is not a single diffeomorphism invariant gauge theory, but an infinite parameter class of them. All these theories share the same key properties with GR, as they have the same number of propagating degrees of freedom (DOF). Thus, for any theory in the class the phase space is that of an SU(2) gauge theory. However, in addition to the usual SU(2) gauge rotations, there are also diffeomorphisms acting on the phase space variables, which reduce the number of propagating DOF from 6 of SU(2) gauge theory to 2 of GR.

Unfortunately, the new viewpoint on GR originating in [7] (and having its roots in Plebanski's key insight [3]) has not been significantly developed. The phase space version [5] of this story has formed the foundation of the approach of loop quantum gravity [9], but the new pure connection formulation of GR [7] and of the infinite-parameter family [8] of "neighbours of GR" has not had any significant applications, as far as the author is aware.

The author's current interest in the subject started in [10] from a simple power counting argument describing how the non-renormalisability of GR manifests itself in the Plebanski formulation [3]. The outcome was an infinite-parameter family of Plebanski-like theories, where the constraint term of the Plebanski action was replaced by a "potential" term for the would-be Lagrange multipliers. Each of the new theories is just the familiar from discussions of non-renormalisability counterterm corrected GR (in disguise), and so the interpretation of the infinite number of new parameters is that they are related to coefficients in front of counterterms constructed from the curvature and its derivatives in the usual metric description of gravity. It was very quickly realised [11] that the new infinite-parameter family of theories [10] is essentially the same as the one introduced and studied by Begtsson and collaborators a decade earlier [8], with the difference being that [8] worked at the level of a "pure connection" formulation, while the theories [10] are formulated as Plebanski-like theories with two-form fields as the basic variables.

The class of gravity theories [8], [10] can be thought of as summing at least some of the quantum corrections that arise in the process of renormalisation of GR, and in [12] this was confirmed by directly exhibiting the familiar GR counterterms as appearing from [10]. The work [10] also conjectured that this class of gravity theories sums up *all* the arising quantum corrections; in other words, it was conjectured that the class [10] is closed under the renormalisation, and that the arising renormalisation group flow is that in the space of "potential" functions defining the theory.

At the time of writing [10] the only motivation for this conjecture was the author's optimism — the conjecture did not contradict anything one knew about the non-renormalisability of GR, and was the most optimistic scenario for how the divergences of GR might organise themselves. The remark [11] relating the Plebanski-like theories [10] to the pure connection theories [8] brought with it an additional justification. Thus, a closer look at these theories made it clear that they are just the most general diffeomorphism invariant gauge theories. The class of such theories should therefore be closed under the renormalisation, because any counterterm that can be needed for cancelling the arising quantum divergences is already included into the action, see [13] for the first spell-out of this argument.

The aim of the present paper is to set the stage for a systematic study of the quantum perturbation theory for the gravitational theories introduced in [10] (and previously in [8]). Our final goal is to settle the status of the conjecture of [10] that these class of theories is closed under the renormalisation, and then to compute the resulting renormalisation group flow. However, it would be impractical to try to write up all the necessary calculations in a single paper. For this reason, in the present paper we develop the classical theory to the extent that the propagating degrees of freedom (gravitons) are manifest. We also do some preliminary steps necessary for the perturbative loop computations in that the gauge fixing is discussed in details and the propagator is obtained. It is then straightforward to start to compute loop diagrams. This is however not attempted in the present paper, and the task to develop a sufficiently economical way to study the renormalisation is left to future work.

Apart from just setting up the stage for future quantum calculations, a somewhat unexpected outcome of this work is a completely new viewpoint on the gravitational perturbation theory. As we shall see, in the present diffeomorphism invariant gauge theoretic approach to gravity, the fundamental scale is set not by the Newton's constant, which does not appear in the original formulation of the theory at all, but rather by the radius of curvature of the background that is used to expand the theory around. Thus, the natural fundamental length scale is set by the cosmological constant, or, more precisely by the combination Λ/G (of mass dimension 4). This has the effect that in our theory the strength with which gravitons interact with each other — the Newton's constant — is given by a certain combination of the radius of curvature of the background and some dimensionless coupling constant of the theory. Then the fact that the multiplicative factor required to go from one scale to the other is so humongous has a possible explanation in our theory. Thus, the dimensionless quantity in question turns out to be related to the derivative of the delta-function which the defining function of our theory must be close to in order to reproduce GR at low energies, see [7]. Thus, the present gauge-theoretic approach to gravity gives the famous "cosmological constant problem" a new form, to be further discussed in the last sections.

To summarise, the main aim of this work is to develop a new approach to the gravitational perturbation theory, for future use in particular in the quantum loop calculations. What makes this paper distinct from previous works (in particular of this author) is that here for the first time the "pure connection" formalism close in spirit to the formulation in [8] is used as a starting point for the gravitational perturbation theory. Thus, all previous works on theories [10] used the two-form formulation. The gravitational perturbation theory in the two-form formulation is similar to that in the usual metric approach, see [12]. In particular, the fundamental scale that determines the self-coupling of the gravitons and sets the scale of the strong coupling regime is, as in the usual metric case, the Planck scale. However, the number of field components one has to works with in the two-form formulation is quite large — it is that of an SU(2) Lie algebra-valued two-form field. Moreover, there are second class constraints that require the path integral measure to be somewhat non-trivial. For all these reasons it proved to be rather difficult to set up an economical perturbation theory in the two-form formalism. At the same time, for a long time it seemed that the "pure connection" formulation is ill suited for being a starting point of a perturbative description, as it was not at all clear how one can expand the theory around the Minkowski spacetime background which corresponds to a zero connection, see e.g. remarks in [15].

In this paper the prejudices about the "pure connection" formulation of gravity are put aside and this formulation is used as a starting point for the gravitational perturbation theory. And, as we hope to convince the reader, this formulation can be used rather effectively, in that the arising perturbation theory is reasonably economical. In particular, the linearised theory is rather simple, and the propagator can be obtained without too much difficulty. As we shall see, in the "pure connection" formalism developed here gravity becomes not too dissimilar to SU(2) gauge theory, the main difference being that a certain additional projector on diffeomorphism equivalence classes is inserted into the standard $1/k^2$ propagator of the gauge theory. This gives hope that the renormalisation in this class of gravity theories will eventually become manageable. As we have already mentioned, this is left to the future work.

What is new in this work as compared to previous works on the "pure connection" formulation, in particular the work [8] and works by Bengtsson and collaborators that followed, is that our treatment uses in an essential way the formulation in terms of a homogeneous potential function applied to a matrix-valued 4-form. This was developed in earlier works of the author, and first spelled out in [14] for the version of the theory that uses a two-form field, and in [13] for the pure connection formulation. This formulation makes it possible to set up the perturbation theory without too much difficulty.

Before we proceed with a description of the theory, there are a few things that ought to be emphasised to avoid misunderstanding. In our gauge-theoretic approach to gravity the theory (or any of the class of theories that we study) remains as non-renormalisable as GR in the usual metric-based treatment. Thus, as we shall explicitly see below, the coupling constant of our theory has a negative mass dimension, which signals non-renormalisability by power counting. Thus, the final goal of our enterprise is not to show that the theory is renormalisable — it is not — but rather to show that the infinite-parameter class of theories that we study is closed under the renormalisation, and then to compute the arising renormalisation group flow. In other words, we are not after the renormalisability in the usual sense of quantum field theory, which is that a Lagrangian with a finite number of couplings is closed under the renormalisation. Rather, we are after the renormalisability in the effective field theory sense of Weinberg, see e.g. [16] for a recent discussion, where any theory is renormalisable once all possible counterterms are added to the action. Our aim is then to show that in the case of gravity in four spacetime dimensions it is sufficient to consider only those counterterms (infinite in number) that can be compactly summed up into our diffeomorphism invariant gauge theory Lagrangian. Should this indeed be the case, the renormalisation group flow in the infinite dimensional space of gravity theories will be just a flow in the space of defining functions, and will become manageable. Note once again, however, that the quantum theory, while being our main motivation, is not the subject of the present work.

We would also like to explain at the outset how a gauge theory (with spin one excitations) can describe gravity with its spin two excitations. This is a version of the story "spin one plus spin one is spin two", of relevance for the gauge theory gravity relationship [2]. There are, however, also significant differences. Thus, the main dynamical field of our theories is an SU(2) connection A^i_μ, where μ is a spacetime index, and $i = 1, 2, 3$ is a Lie algebra one. Let us recall that in the usual gauge theories in Minkowski spacetime the temporal component A^i_0 of the connection field becomes a Lagrange multiplier — the generator of the gauge rotations. Then of the spatial components A^i_a, where $a = 1, 2, 3$ is a spatial index, some components are pure gauge in that they can be set to any desired value by a gauge transformation. The physical propagating degrees of freedom of the theory can be described as the gauge equivalence classes of the spatial projection of the connection. In the case of gauge group SU(2), the gauge invariance removes 3 of the 9 components of the spatial connection A^i_a, leaving two propagating polarisations per each Lie algebra generator.

As we shall see below, in the case of our gravitational theories the situation is very similar, with the exception that the Lagrangian is in addition invariant under diffeomorphisms. The way this is realised in our theories is that the Lagrangian is simply independent of certain 4 combinations of the connection field A^i_μ. This is where the spin two comes from. Thus, consider once again the spatial projection of the connection A^i_a. We shall see that (using the background) it will be possible to identify two types of indices — the spatial and the internal Lie algebra ones. Once this is done, the spatial connection can be thought of as a 3×3 matrix, or, in representation theoretic terms, it constitutes the spin one tensor spin one representation. This decomposes as spin two plus spin one plus spin zero. On the other hand, the temporal component of the connection A^i_0 forms the spin one (adjoint) representation of SU(2). The diffeomorphism invariance projects out the spin zero components of the spatial connection A^i_a, as well as a certain combination of the spin one component of A^i_a and A^i_0, leaving only one of these spin one components in the game. Thus, after the projection induced by the diffeomorphisms, the Lagrangian depends only on the spin two component of A^i_a, as well as on the spin one set of Lagrange multipliers — generators of SU(2) rotations. These make the 3 longitudinal components of the 5 component spin two field unphysical, leaving only 2 propagating physical modes. To summarise, in our version of gauge theory/gravity correspondence the spin two also comes from the tensor product of two spin one representations. As in any gauge theory in Minkowski space, one of these spin one representations is supplied by the spatial projection of the connection field. The other spin one is provided by the adjoint representation of the SU(2) Lie algebra in which the connection field takes values.

With these preparatory remarks having been made, we can proceed to describe how gravity can be reformulated as a diffeomorphism invariant gauge theory. The organisation of the paper is as follows. In Section 2 we define an action principle for our theories, explain how their parameterisation by a homogenous function works, derive the field equations and verify gauge invariances of the action. Section 3 studies the theory linearised around a constant curvature background. In particular, a simple quadratic in the gauge field fluctuations action is obtained, and its Hamiltonian analysis is performed. This confirms the outlined above picture of how the spin two nature of excitations comes about. Section 4 is central to our analysis. It discusses the gauge-fixing appropriate to the situation at hand, and inverts the gauge-fixed quadratic form to obtain the propagator. In Section 5 we obtain the cubic interaction term, and show how the Planck constant measuring the graviton self-interaction strength becomes a derived quantity in our theory. We conclude with a brief discussion.

2 Diffeomorphism invariant gauge theories

2.1 Gravity as a gauge theory

In the pure connection formulation gravity becomes the most general diffeomorphism invariant gauge theory. In the case of a purely gravitational theory[2] the gauge group is (complexified) SU(2). The action is a functional of an SU(2) connection $A^i, i = 1, 2, 3$ on a spacetime manifold M. Let $F^i = dA^i + (1/2)\epsilon^{ijk} A^j \wedge A^k$ be the curvature of A^i. The action is given by the following gauge and diffeomorphism invariant functional of the connection:

$$S[A] = (1/\mathrm{i}) \int_M f(F^i \wedge F^j). \tag{1}$$

Here $\mathrm{i} = \sqrt{-1}$ is a factor introduced for future convenience, and f is a function with properties to be spelled out below.

We shall refer to f as the defining function of our theory. It is a holomorphic, homogeneous of degree one and gauge invariant function of its matrix (and 4-form) valued argument. Thus, let $X^{ij} \in \mathfrak{su}(2) \otimes_S \mathfrak{su}(2)$ be a matrix valued in the second symmetric power of the Lie algebra. The gauge group SU(2) $\sim$ SO(3) acts in the space of such matrices via $X \to gXg^T$, where T is the operation of the transpose. We first consider scalar valued functions $f : \mathfrak{su}(2) \otimes_S \mathfrak{su}(2) \to \mathbb{C}$ that are holomorphic, gauge-invariant $f(gXg^T) = f(X)$ and homogeneous of degree one $f(\alpha X) = \alpha f(X)$. A convenient for practical computations parameterisation of such functions is as follows. Consider the following 3 SU(2) invariants of X^{ij}:

$$\mathrm{Tr}(X), \qquad \mathrm{Tr}(X^2), \qquad \mathrm{Tr}(X^3), \tag{2}$$

where the traces (and powers of X) are computed using the Killing metric on the Lie algebra for which we take δ^{ij}. When $\mathrm{Tr}(X) \neq 0$ we can parameterise the defining function f as follows:

$$f(X) = \mathrm{Tr}(X)\, \chi \left(\frac{\mathrm{Tr}(X^2)}{(\mathrm{Tr}(X))^2}, \frac{\mathrm{Tr}(X^3)}{(\mathrm{Tr}(X))^3} \right), \tag{3}$$

where χ is now an arbitrary holomorphic function of its two arguments.

Given f with the properties as spelled out above, e.g. one parameterised as in (3), it can be seen that this function can be applied to a matrix valued 4-form, with the result being a 4-form. Indeed, consider $F^i \wedge F^j$, which is a $\mathfrak{su}(2) \otimes_S \mathfrak{su}(2)$ valued 4-form. Choose a reference volume form on M (we assume that M is orientable), and denote it by (vol). Of course, (vol) is only defined modulo the multiplication by a nowhere zero function. Using this reference volume form we can

[2] One can also consider unified Yang-Mills-gravity theories of the same sort, see [17].

write $F^i \wedge F^j = X^{ij}(\text{vol})$, where X^{ij} is again defined only modulo rescalings. We can now use the homogeneity of f to write

$$f(F^i \wedge F^j) = (\text{vol}) f(X). \tag{4}$$

It is moreover clear that the result on the right-hand-side does not depend on which reference volume form is used in this argument. This is again due to the homogeneity of f. This shows that the integrand in (1) is a well-defined 4-form that can be integrated to obtain the action. This finishes the formulation of our theory.

We note that, as formulated, there are no dimensionful parameters in our theory. Indeed, we assume the connection field A^i to have the usual mass dimension one, so that the curvature has the mass dimension two, and the matrix of the wedge products $[X] = 4$. The defining function f is essentially the function χ of ratios of powers of X^{ij} that are dimensionless, and so does not contain any dimensionful parameters (but contains an infinite number of dimensionless "coupling constants", once expanded appropriately). Thus, due to the homogeneity of f, its mass dimension is the same as that of X (in the parameterisation (3) the mass dimension is carried by the first term $\text{Tr}(X)$, while the function χ is dimensionless). The function f can then be integrated to produce a dimensionless action (as usual we work in the units $c = \hbar = 1$). As we shall see, the fact that there are no dimensionful coupling constants in our theory has profound implications for the structure of its perturbation theory.

Classically (1) is a theory that can be shown to propagate two (complex for the time being, reality conditions will be discussed below) degrees of freedom. We will see a version of the argument that leads to this conclusion below when we consider the perturbation theory. The theory (1) is a gravity theory, in spite of the fact that no metric is present anywhere. However, it can be reformulated explicitly as a theory of metrics via a sequence of transformations. The main idea is to note that declaring the 3 two-forms F^i to span the space of (anti-) self-dual two-forms determines a conformal metric on M whenever the matrix X^{ij} defined from the wedge product of curvatures is non-degenerate. One can then rewrite the theory (1) explicitly as the theory of this metric, see [12] for details. However, in this paper, we shall not need this relation to metric theories. Our plan is to study (1) as is. We shall set the stage for its perturbative quantisation and a study of its renormalisation. The main justification for this undertaking is that a whole class of gravity theories (for varying defining functions f) can be treated in one go. Moreover, our theories are theories of a connection, and we can hope to use the expertise that was accumulated in quantum field theory for dealing with quantum gauge theories.

2.2 First variation and field equations

The first variation of the action (1) gives us field equations. To write these down, let us give a parameterisation of the matrix X^{ij} useful for practical computations. Thus, let $\tilde{\epsilon}^{\mu\nu\rho\sigma}$ be a completely anti-symmetric rank 4 vector, which is a density of weight one (as is indicated by the tilde over its symbol). This object exists on any orientable manifold and does not need a metric for its definition. Consider:

$$\tilde{X}^{ij} := \frac{1}{4}\tilde{\epsilon}^{\mu\nu\rho\sigma} F^i_{\mu\nu} F^j_{\rho\sigma}, \tag{5}$$

where as before $F^i_{\mu\nu}$ is the curvature two-form, with its spacetime indices now indicated explicitly. The quantity $\tilde{X}^{ij}$ is a $\mathfrak{s}u(2) \otimes_S \mathfrak{s}u(2)$ valued matrix, and a density of weight one. One takes the defining function f to be a function of $\tilde{X}^{ij}$ given by the same expression as in (3). With convention $dx^\mu \wedge dx^\nu \wedge dx^\rho \wedge dx^\sigma = \tilde{\epsilon}^{\mu\nu\rho\sigma} d^4x$ we can write the action (1) as

$$S[A] = (1/\text{i}) \int_M d^4x \, f(\tilde{X}^{ij}) \, . \tag{6}$$

The first variation of the action can now be easily computed and reads:

$$\delta S[A] = (1/\mathrm{i}) \int_M d^4x \, \frac{\partial f}{\partial \tilde{X}^{ij}} \frac{1}{2} \tilde{\epsilon}^{\mu\nu\rho\sigma} F^i_{\mu\nu} D_{A\,\rho} \delta A^j_\sigma \,. \tag{7}$$

Integrating by parts, we see that the field equations for (1) can be written as

$$D_A B^i = 0, \tag{8}$$

where we have used the form notations again, and the two-form B^i is defined via

$$B^i := \frac{\partial f}{\partial \tilde{X}^{ij}} F^j. \tag{9}$$

We note that the matrix of first derivatives that appear on the right-hand-side of this expression is a symmetric matrix, and has density weight zero (as the ratio of the density weight one function $f(\tilde{X})$ and the density weight one quantity $\tilde{X}$). Thus, (9) is a well-defined two-form.

2.3 Symplectic structure

The computation of the first variation in the previous subsection also gives us the symplectic structure of the theory. Thus, the phase space of the theory is the space of all solutions of (8), and the symplectic structure can be obtained by considering the boundary term that was neglected in passing from (7) to (8). The integral of the boundary term gives rise to an integral over the spatial slice Σ of the following quantity

$$\Theta := \frac{1}{2\mathrm{i}} \int_\Sigma B^i \wedge \delta A^i, \tag{10}$$

where B^i is as in (9). This is a one-form on the phase space of the theory. Its exterior derivative produces the symplectic two-form. We see that the significance of the quantity B^i defined by (9) is that its spatial projection plays the role of the momentum canonically conjugate to the spatial projection of the connection A^i. We emphasise that in the present "pure gauge" formulation, the two-form B^i is not independent and is a function of the connection field. A formulation that "integrates in" the two-form field as an independent variable is possible, and has been studied in previous works by the author, but will not be considered here.

2.4 Gauge invariance

Let us now verify by an explicit computation that our theory is invariant under diffeomorphisms as well as $\mathrm{SO}(3,\mathbb{C})$ rotations. This is of course expected, because the action was constructed in the way that these invariances should hold. However, an explicit verification of this fact will allow us to establish some identities for use in what follows. The gauge transformations act on the connection field as follows

$$\delta_\xi A^i_\mu = \xi^\alpha F^i_{\mu\alpha}, \qquad \delta_\phi A^i_\mu = D_{A\,\mu} \phi^i. \tag{11}$$

The first of these transformations can be seen to be a diffeomorphism corrected by a gauge transformation, while the second one is the usual gauge rotation with the parameter ϕ^i.

It is not too difficult to prove the invariance of our action (1) under these transformations. Let us first consider the diffeomorphisms. The variation of the action (7) becomes proportional to

$$\int_M d^4x \, \frac{\partial f}{\partial \tilde{X}^{ij}} \tilde{\epsilon}^{\mu\nu\rho\sigma} F^i_{\mu\nu} D_{A\,\rho} \xi^\alpha F^j_{\sigma\alpha}. \tag{12}$$

We now need some identities. First we note that one can write the Bianchi identity $D_A F^i = 0$ as

$$D_{A\,[\mu} F^i_{\nu]\rho} = -\frac{1}{2} D_{A\,\rho} F^i_{\mu\nu}. \tag{13}$$

Another identity that we need is

$$\tilde{\epsilon}^{\mu\nu\rho\sigma} F^{(i}_{\mu\nu} F^{j)}_{\sigma\alpha} = -\frac{1}{4}\delta^\rho_\alpha \tilde{\epsilon}^{\mu\nu\gamma\delta} F^i_{\mu\nu} F^j_{\gamma\delta} = -\delta^\rho_\alpha \tilde{X}^{ij}, \tag{14}$$

where δ^ρ_α is the Kronecker delta. Note that the symmetrisation is taken on the left hand-side. The above two identities, as well as the definition (5) of the matrix $\tilde{X}^{ij}$, allow us to rewrite (12) as

$$-\int_M d^4x \, \frac{\partial f}{\partial \tilde{X}^{ij}} \left(\tilde{X}^{ij} \partial_\alpha \xi^\alpha + \frac{1}{2} \tilde{\epsilon}^{\mu\nu\rho\sigma} F^i_{\mu\nu} \xi^\alpha D_{A\,\alpha} F^j_{\rho\sigma} \right) = -\int_M d^4x \, \frac{\partial f}{\partial \tilde{X}^{ij}} D_{A\,\alpha}(\xi^\alpha \tilde{X}^{ij}). \tag{15}$$

Integrating by parts, this becomes equal to

$$\int_M d^4x \, \xi^\alpha \tilde{X}^{ij} D_{A\,\alpha} \frac{\partial f}{\partial \tilde{X}^{ij}}. \tag{16}$$

We should now see that the integrand here is zero. This follows from the homogeneity of the function f. Indeed, we have

$$\tilde{X}^{ij} \frac{\partial f}{\partial \tilde{X}^{ij}} = f \tag{17}$$

from the fact that f is a homogeneous function of degree one. Let us now apply the operator of partial derivative ∂_μ to both sides of this equation. We get

$$(\partial_\mu \tilde{X}^{ij}) \frac{\partial f}{\partial \tilde{X}^{ij}} + \tilde{X}^{ij} \partial_\mu \frac{\partial f}{\partial \tilde{X}^{ij}} = \partial_\mu f = \frac{\partial f}{\partial \tilde{X}^{ij}} \partial_\mu \tilde{X}^{ij}. \tag{18}$$

Comparing the two sides we see that

$$\tilde{X}^{ij} \partial_\mu \frac{\partial f}{\partial \tilde{X}^{ij}} = 0, \tag{19}$$

which is almost the integrand in (16), except for the fact that we have the covariant derivative in (16). Let us now consider the difference between the covariant and the usual derivatives. We have

$$\tilde{X}^{ij} (D_{A\,\mu} - \partial_\mu) \frac{\partial f}{\partial \tilde{X}^{ij}} = 2 \tilde{X}^{ij} \epsilon^{ikl} A^k_\mu \frac{\partial f}{\partial \tilde{X}^{lj}}. \tag{20}$$

The quantity here is zero in view of the gauge invariance of the function f. Indeed, under infinitesimal gauge transformations an $\mathfrak{su}(2) \otimes_S \mathfrak{su}(2)$-valued matrix $\tilde{X}^{ij}$ transforms as

$$\delta_\phi \tilde{X}^{ij} = \epsilon^{ikl} \phi^k \tilde{X}^{lj} + \epsilon^{jkl} \phi^k \tilde{X}^{il}. \tag{21}$$

Then the statement that f is an $\mathrm{SO}(3, \mathbb{C})$ invariant function becomes

$$\epsilon^{ikl} \tilde{X}^{kj} \frac{\partial f}{\partial \tilde{X}^{lj}} = 0, \tag{22}$$

which can be expressed in words by saying that the commutator of the matrix $\tilde{X}^{ij}$ with the matrix $\partial f / \partial \tilde{X}^{ij}$ of the first derivatives of the defining function is zero.

The identity (22) immediately implies that the difference of the derivatives in (20) is zero and thus

$$\tilde{X}^{ij} D_{A\,\mu} \frac{\partial f}{\partial \tilde{X}^{ij}} = 0, \tag{23}$$

which proves the invariance of the action (1) under diffeomorphisms.

Let us now prove the invariance of (1) under the gauge rotations. The variation of the action in this case becomes proportional to

$$\int_M d^4x \, \frac{\partial f}{\partial \tilde{X}^{ij}} \tilde{\epsilon}^{\mu\nu\rho\sigma} F^i_{\mu\nu} D_{A\,\rho} D_{A\,\sigma} \phi^j . \tag{24}$$

Expressing the commutator of the covariant derivatives as the commutator with the curvature, and recalling the definition (5) of the matrix $\tilde{X}^{ij}$ we get

$$4 \int_M d^4x \, \frac{\partial f}{\partial \tilde{X}^{ij}} \epsilon^{jkl} \tilde{X}^{ik} \phi^l , \tag{25}$$

which is zero in view of (22). This proves the invariance of the action (1) under the $\mathrm{SO}(3, \mathbb{C})$ rotations.

2.5 Second variation

We can now compute the second variation, in preparation for the next section treatment. We have

$$\delta^2 S[A] = (1/\mathrm{i}) \int_M d^4x \, \left(\frac{\partial^2 f}{\partial \tilde{X}^{ij} \partial \tilde{X}^{kl}} \delta \tilde{X}^{ij} \delta \tilde{X}^{kl} + \frac{\partial f}{\partial \tilde{X}^{ij}} \delta^2 \tilde{X}^{ij} \right) . \tag{26}$$

Here the first variation of $\tilde{X}^{ij}$ was already computed above and reads

$$\delta \tilde{X}^{ij} = \frac{1}{2} \tilde{\epsilon}^{\mu\nu\rho\sigma} F^{(i}_{\mu\nu} D_{A\,\rho} \delta A^{j)}_\sigma . \tag{27}$$

The second variation reads

$$\delta^2 \tilde{X}^{ij} = \frac{1}{2} \tilde{\epsilon}^{\mu\nu\rho\sigma} D_{A\,\mu} \delta A^i_\nu D_{A\,\rho} \delta A^j_\sigma + \frac{1}{2} \tilde{\epsilon}^{\mu\nu\rho\sigma} F^{(i}_{\mu\nu} \epsilon^{j)kl} \delta A^k_\rho \delta A^l_\sigma . \tag{28}$$

3 Constant curvature background

In this and the next section, to get a better feel for our theory and also to prepare for its quantisation, we consider the action (1) expanded around a specific background connection A^i.

3.1 Second order action around a general background

We now write our connection as the background A^i plus a fluctuation $\mathcal{A}^i$, and obtain the part of the action quadratic in $\mathcal{A}^i$ directly from (26). Thus, we divide the second variation by 2, replace δA^i_μ by $\mathcal{A}^i_\mu$, and get the following Lagrangian

$$\begin{aligned}(8\mathrm{i}) \mathcal{L}_{\mathcal{A}} = {} & \frac{\partial^2 f}{\partial \tilde{X}^{ij} \partial \tilde{X}^{kl}} (\tilde{\epsilon}^{\mu\nu\rho\sigma} F^i_{\mu\nu} D_{A\,\rho} \mathcal{A}^j_\sigma)(\tilde{\epsilon}^{\alpha\beta\gamma\delta} F^k_{\alpha\beta} D_{A\,\gamma} \mathcal{A}^l_\delta) \\ & + 2 \frac{\partial f}{\partial \tilde{X}^{ij}} \tilde{\epsilon}^{\mu\nu\rho\sigma} \left(D_{A\,\mu} \mathcal{A}^i_\nu D_{A\,\rho} \mathcal{A}^j_\sigma + F^i_{\mu\nu} \epsilon^{jkl} \mathcal{A}^k_\rho \mathcal{A}^l_\sigma \right) .\end{aligned} \tag{29}$$

In the following works this action will be used for a background field method one-loop computation, but here we specialise to a particular background.

3.2 The background

The background that we take is a constant curvature one and can be defined as follows. Let ds^2 be the interval for a constant curvature metric in 4 spacetime dimensions. It does not matter whether we take a positive or negative curvature background, as later we are going to take the limit when the curvature goes to zero, or, equivalently, work at energy scales larger than the energy scale set by the curvature of the background. Let $\theta^I, I = 0, 1, 2, 3$ be a collection of tetrads for the corresponding metric. We then define the following triple of two-forms:

$$\Sigma^i := \mathrm{i}\theta^0 \wedge \theta^i - \frac{1}{2}\epsilon^{ijk}\theta^j \wedge \theta^k, \tag{30}$$

where, as before $i = 1, 2, 3$. As is not hard to check, these two-forms are anti-self-dual

$$\frac{\mathrm{i}}{2}\epsilon_{\mu\nu}{}^{\rho\sigma}\Sigma^i_{\rho\sigma} = \Sigma^i_{\mu\nu} \tag{31}$$

with respect to the Hodge star operation on two-forms defined by the metric $ds^2 = \theta^I \otimes \theta^J \eta_{IJ}$, where $\eta_{IJ} = \mathrm{diag}(-1, 1, 1, 1)$. It can be shown that the constant curvature condition can be rewritten as follows. Let us introduce an SU(2) connection A^i_0 such that the covariant exterior derivative of Σ^i with respect to A^i_0 is zero. In other words:

$$0 = D_{A_0}\Sigma^i = d\Sigma^i + \epsilon^{ijk}A^j_0 \wedge \Sigma^k. \tag{32}$$

It is not hard to solve this equation for A^i_0 explicitly (in terms of derivatives of tetrads θ^I), but we will not need the corresponding expression. Then the constant curvature condition can be written as

$$F^i(A_0) = M_0^2\Sigma^i, \tag{33}$$

where we have introduced a dimensionful parameter $M_0 = 1/L_0$, where $1/L_0^2$ is the (constant) curvature of the background. We take such a constant curvature connection A^i_0 as the background for a perturbative expansion of (1). We shall soon see that for the linearised theory around this background it will be possible to take the limit $M_0 \to 0$ without any difficulty, and so we will be effectively considering a (linearised) gauge theory in Minkowski spacetime. An alternative way to think about the limit $M_0 \to 0$ is to say that one works at energies $E \gg M_0$ (or at length scales much smaller than the radius of curvature L_0 of the background).

3.3 The cosmological constant

It is not hard to see that the dimensionful parameter M_0 that was introduced by the background is related to the cosmological constant. For this we need some preliminary results. Using

$$\tilde{\epsilon}^{\mu\nu\rho\sigma}\theta^0_\mu\theta^i_\nu\theta^j_\rho\theta^k_\sigma = \sqrt{-g}\epsilon^{ijk}, \tag{34}$$

where $\sqrt{-g}$ is the square root of the determinant of the metric $ds^2 = \theta^I \otimes \theta^J \eta_{IJ}$, we easily get

$$\Sigma^i \wedge \Sigma^j = -2\mathrm{i}\sqrt{-g}\,\delta^{ij}\,d^4x\,, \tag{35}$$

where Σ^i are the anti-self-dual forms (30). Thus, the matrix $\tilde{X}^{ij}$ at the background is equal to

$$\tilde{X}^{ij}_0 = -2\mathrm{i}M_0^4\sqrt{-g}\,\delta^{ij}\,, \tag{36}$$

i.e., is proportional to the identity matrix. Thus, the value of the action (1) at the background is

$$S[A_0] = -2M_0^4 f_0 \int \sqrt{-g}\,d^4x, \tag{37}$$

where $f_0 := f(\delta)$ is the value of the defining function at the identity matrix $X^{ij} = \delta^{ij}$.

On the other hand, the Einstein-Hilbert action for the signature $(-,+,+,+)$ reads

$$S_{\rm EH}[g] = -\frac{1}{16\pi G}\int (R-2\Lambda)\sqrt{-g}\, d^4x. \tag{38}$$

On a constant curvature background (in 4 dimensions) $R = 4\Lambda$ and we get

$$S^0_{\rm EH} = -\frac{\Lambda}{8\pi G}\int \sqrt{-g}\, d^4x. \tag{39}$$

Comparing to (37) we see that

$$M_0^4 f_0 = \frac{\Lambda}{8\pi G}. \tag{40}$$

For the observed value $\Lambda \sim 10^{-35}(1/s^2)$ of the cosmological constant the quantity on the right-hand-side, if measured in electron-volts, is equal to approximately $(10^{-2}eV)^4$ (again, our units are $c = \hbar = 1$). Thus, if the value f_0 is not too far from unity, then the natural energy scale for the background to expand the theory about is of the order $M_0 \sim 10^{-2}eV$. Below we will see that the usual Newton's constant that in the perturbative approach to gravity determines the strength of the graviton self-coupling becomes in this theory a derived quantity, built from M_0 as well as the derivative of the defining function f at the identity. We shall see that this derivative must be very large.

3.4 Substituting the background

We now check that the constant curvature background is a solution of (8) and then evaluate the second variation of the action (26) at the background.

The derivatives of (3) at the identity matrix are easily computed. Let us first write down the general expression for the first derivative. We omit the tilde from X for brevity (we can always pull out the density weight factor from the function f using the homogeneity). We have

$$\frac{\partial f}{\partial X^{ij}} = \delta^{ij}\chi(X) + {\rm Tr}(X)\chi_1'(X)\left(\frac{2X^{ij}}{({\rm Tr}(X))^2} - \frac{2{\rm Tr}(X^2)}{({\rm Tr}(X))^3}\delta^{ij}\right)$$
$$+{\rm Tr}(X)\chi_2'(X)\left(\frac{3(X^2)^{ij}}{({\rm Tr}(X))^3} - \frac{3{\rm Tr}(X^3)}{({\rm Tr}(X))^4}\delta^{ij}\right), \tag{41}$$

where $\chi'_{1,2}(X)$ are the derivatives of the function χ with respect to the first and second arguments, evaluated at X. It is easy to check that for $X_0^{ij} \sim \delta^{ij}$ the second and third terms on the right are zero, and we have:

$$\left.\frac{\partial f}{\partial X^{ij}}\right|_{X_0} = \delta^{ij}\chi(X_0) = \frac{f_0}{3}\delta^{ij}. \tag{42}$$

We note that this is M_0 independent. We remind the reader that the background value of X_0 matrix is given by (36) above.

Let us now compute the matrix of second derivatives of the defining function. Since the expressions in brackets in (41) become zero when evaluated on X_0, the only way to get a non-zero result in the second derivative is to act by a derivative on these expressions. We get

$$\left.\frac{\partial^2 f}{\partial X^{ij}\partial X^{kl}}\right|_{X_0} = \frac{2(\chi_1'(X_0)+\chi_2'(X_0))}{{\rm Tr}(X_0)}P^{ij|kl}, \tag{43}$$

where

$$P^{ij|kl} := I^{ij|kl} - \frac{1}{3}\delta^{ij}\delta^{kl}, \qquad I^{ij|kl} := \frac{1}{2}\left(\delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk}\right). \tag{44}$$

We have introduced a special notation $P^{ij|kl}$ for the matrix that appeared in (43), as this is just the projector on the symmetric traceless part $P^{ij|kl}\delta_{ij} = 0$, and similarly for the contraction with δ_{kl}.

Having evaluated the derivatives of the defining function at the background, we are ready to specialise (29) for our constant curvature background. However, let us first check that our chosen background is indeed a solution of field equations (8). With the quantity $\chi(X_0)$ being a constant, the background $B_0^i \sim F^i(A_0)$, and thus the field equations (8) are satisfied (in view of the Bianchi identity).

Let us now consider the second term in (26). Since the matrix of the first derivatives is proportional to the identity matrix (42) with a constant proportionality coefficient we need to consider the integral of $\delta_{ij}\delta^2\tilde{X}^{ij}$ over the manifold. Let us see that this is a total derivative. We have:

$$\int_M d^4x\, \delta_{ij}\delta^2\tilde{X}^{ij} = \frac{1}{2}\int_M \Big(D_A\delta A^i \wedge D_A\delta A^i + F^i(A) \wedge \epsilon^{ijk}\delta A^j \wedge \delta A^k \Big), \tag{45}$$

where we wrote everything in terms of forms (our form convention is $F = (1/2)F_{\mu\nu}dx^\mu \wedge dx^\nu$). Integrating by parts in the first term (and neglecting the total derivative term), the first term becomes

$$\frac{1}{2}\int_M \delta A^i \wedge D_A D_A \delta A^i = \frac{1}{2}\int_M \delta A^i \wedge \epsilon^{ijk}F^j(A) \wedge \delta A^k, \tag{46}$$

which is minus the second term in (45), and so (45) is a total derivative.

We therefore only need to consider the first term in (26). Let us write this directly in terms of the two-forms Σ^i by substituting the expression (33) for the background curvature. Using the anti-self-duality (31) of Σ^i we have the following compact expression for the second variation

$$\delta^2 S\Big|_{A_0} = -g_0 \int_M d^4x\sqrt{-g}\, P^{ij|kl}(\Sigma^{i\,\mu\nu}D_{A_0\,\mu}\delta A^j_\nu)(\Sigma^{k\,\rho\sigma}D_{A_0\,\rho}\delta A^l_\sigma), \tag{47}$$

where we have introduced a notation

$$g_0 := \frac{\chi_1'(X_0) + \chi_2'(X_0)}{3}. \tag{48}$$

Note that the factors of M_0 have cancelled from this result. The combination (48) of the first derivatives of the defining function plays an important role below. Thus, we shall see that the Newton's constant will be built from M_0 and g_0.

3.5 Taking the flat limit and the linearised action

We could have continued to work on a general constant curvature background. However, most of the quantum field theory technology is developed for the Minkowski background. For this reason it is very convenient to take the limit $M_0 \to 0$ and consider the flat (zero) background connection. Alternatively, since we are after the UV behaviour of our theory, we are interested in its behaviour at energies $E \gg M_0$. In this case we can neglect the fact that the background is curved, and consider an effective theory in Minkowski space.

In the limit $M_0 \to 0$ the background matrix $X_0^{ij} \sim M_0^4\delta^{ij}$ goes to zero. However, the values of the two arguments of the function χ in (3) remain finite. Indeed, in this limit they are just $1/3, 1/9$. If the function χ is differentiable at this point, which we assume, then the constant g_0 introduced in (48) is also finite. Absorbing this constant into the linearised fields by rescaling we and taking the limit and obtain the following linearised action in Minkowski space

$$S_{\rm lin}[a] = -\frac{1}{2}\int_M d^4x\, P^{ij|kl}(\Sigma^{i\,\mu\nu}\partial_\mu a^j_\nu)(\Sigma^{k\,\rho\sigma}\partial_\rho a^l_\sigma), \tag{49}$$

where the (rescaled) linearised field is now called $a^i_\mu = \sqrt{g_0}(\delta A^i_\mu)$, and we have divided the second variation of the action by 2 to get the correct linearised action. The two-forms $\Sigma^i_{\mu\nu}$ are now those corresponding to the Minkowski spacetime

$$\Sigma^i = \mathrm{i} dt \wedge dx^i - \frac{1}{2}\epsilon^{ijk} dx^j \wedge dx^k. \tag{50}$$

A quick note about dimensions of all the fields. As we have already mentioned, we take the connection to have the mass dimension one, as is appropriate for a field that can be combined into a derivative operator. Then the curvature has mass dimension two, the matrix X^{ij} has mass dimension 4, the matrix of first derivatives of the defining function is dimensionless, and the matrix of second derivatives has dimension minus 4. The two-forms Σ^i that are constructed from the dimensionless metric are dimensionless. The constant g_0 introduced in (48) is a sum of derivatives of a function of dimensionless arguments, and thus is dimensionless. Overall, we see that the mass dimension of the integrand in (49) is 4, as needed.

3.6 Symmetries

We have started from a diffeomorphism invariant action (1) and linearised it around the constant curvature (and then zero curvature) background. We should check that the linearised action that we have obtained is still diffeomorphism invariant. As before, the diffeomorphisms can be lifted to the SU(2) bundle as follows:

$$\delta_\eta A^i_\mu = \eta^\alpha F^i_{\mu\alpha}(A). \tag{51}$$

Here η^μ is the vector field (of mass dimension minus one) - generator of an infinitesimal diffeomorphism, and $F^i_{\mu\nu}(A)$ is the curvature of A^i_μ. It can be checked that the above formula is a diffeomorphism corrected by a gauge transformation. Replacing the background curvature by its value (33) we get the following formula for an infinitesimal variation

$$\delta_\eta a^i_\mu = M_0^2 \eta^\alpha \Sigma^i_{\mu\alpha}. \tag{52}$$

This suggests that we consider vector fields $\xi^\mu = M_0^2\eta^\mu$ of mass dimension one that are finite in the limit $M_0 \to 0$. Thus, let us consider the following variations

$$\delta_\xi a^i_\mu = \xi^\alpha \Sigma^i_{\mu\alpha}, \tag{53}$$

which will play the role of an infinitesimal diffeomorphism for the theory (49).

Another set of transformations that we have to consider are gauge symmetries. An infinitesimal gauge transformation is given by

$$\delta_\phi a^i_\mu = \partial_\mu \phi^i. \tag{54}$$

Let us now verify that the linearised action is invariant under (53) and (54). For this we will need the following identity

$$\Sigma^{i\,\mu\nu}\Sigma^j_{\nu\rho} = -\delta^{ij}\eta^\mu{}_\rho + \epsilon^{ijk}\Sigma^{k\,\mu}{}_\rho\,, \tag{55}$$

which can be verified by a direct computation. Here $\eta^{\mu\nu}$ is the Minkowski metric. Let us first consider diffeomorphisms. Thus, consider the quantity

$$\Sigma^{i\,\mu\nu}\partial_\mu \delta_\xi a^j_\nu = \Sigma^{i\,\mu\nu}\partial_\mu \xi^\alpha \Sigma^j_{\nu\alpha}. \tag{56}$$

Using (55) we see that ij-symmetric part of this quantity is proportional to δ^{ij}. However this, when contracted with the projector in (49) gives zero. Thus, the invariance under infinitesimal changes of coordinates is established. The invariance under gauge transformations (54) follows by noting that the quantity $\Sigma^{i\,\mu\nu}$ is anti-symmetric and therefore $\Sigma^{i\,\mu\nu}\partial_\mu \delta_\phi a^j_\nu = 0$.

Since our gauge theory action (49) is both diffeomorphism and gauge invariant we can already make a suspected count of the number of propagating DOF. Indeed, the configurational variable of the theory should be the spatial projection of the connection. This has $3 \times 3 = 9$ components. Subtracting 4 diffeomorphisms and 3 gauge DOF leaves us with 2 suspected propagating DOF. Let us confirm this count by the Hamiltonian analysis of the linearised theory. This will also help us to see the gravitons explicitly.

3.7 Hamiltonian analysis

In this subsection we give a more detailed demonstration of the spin two nature of our theory given in the introduction.

To obtain the action in the Hamiltonian form let us expand the quantity that appears as the main building block of the linearised action (49). We have

$$\Sigma^{\mu\nu}_i \partial_\mu a^j_\nu = \mathrm{i}\partial_i a^j_0 - \mathrm{i}\dot{a}^j_i - \epsilon_i^{\ kl}\partial_k a^j_l. \tag{57}$$

Here we have identified the spatial a and internal i indices using e.g. the component $\delta^i_a := \Sigma^i_{0a}$ of the background two-form, and ∂_i are the partial derivatives with respect to spatial coordinates. We raise and lower spatial indices freely using δ^{ij} metric.

It is now easy to compute the conjugate momenta. Since the time derivatives that appear in the action are those of the spatial projection of the connection, it is clear that only these components can have non-zero momenta. However, since the projector is involved in (49), we see that only the symmetric tracefree part of a^j_i has non-zero momenta. These are

$$\pi^{ij} = P^{ij|kl}\left(\dot{a}_{kl} - \partial_k a_{0\,l} - \mathrm{i}\epsilon_{kmn}\partial_m a_{n\,l}\right). \tag{58}$$

We note that the action (49) does not at all depend on the trace part of the spatial connection a^j_i. However, there is a dependence on the anti-symmetric (and of course symmetric) parts. Let us separate the trace, symmetric and anti-symmetric parts of a^j_i and write

$$a_{ij} = a^s_{ij} + b\delta_{ij} + \epsilon_{ijk}c_k. \tag{59}$$

Here a^s_{ij} is the symmetric and tracefree part, and b, c_i parameterise the trace and anti-symmetric parts respectively. Let us now rewrite the expression for the momentum using this decomposition. We have

$$\pi^{ij} = \dot{a}^{s\,ij} - \mathrm{i}\epsilon^{ikl}\partial_k a^{s\,j}_l + P^{ij|kl}\partial_k(\mathrm{i}c_l - a_{0\,l}). \tag{60}$$

We note that the second term here is automatically symmetric and tracefree. On the other hand, it is clear that the Lagrangian density in (49) is

$$\mathcal{L} = \frac{(\pi^{ij})^2}{2}. \tag{61}$$

We see that the Lagrangian (density) is independent of b. This has a simple interpretation. Indeed, computing the infinitesimal diffeomorphism action on the temporal and spatial projections of the connection we find

$$\delta a^i_0 = \mathrm{i}\xi^i, \qquad \delta_\xi a^i_j = -\mathrm{i}\xi^0\delta^i_j - \epsilon^i_{\ jk}\xi^k. \tag{62}$$

This in particular means that the trace part b of the matrix a^j_i is a pure gauge quantity that can be set to zero by a temporal diffeomorphism. We also see that the Lagrangian depends on the anti-symmetric part of spatial and temporal components of the connection only in the combination $\mathrm{i}c_i - a_{0\,i}$. Indeed, it is easy to check that precisely this combination is invariant under spatial diffeomorphisms, as the anti-symmetric component transforms as $\delta_\xi c_i = \xi_i$. Let us denote the invariant combination by ϕ_i. As we shall soon see, it will become a generator of infinitesimal gauge rotations in our theory. Thus, we finally rewrite the momentum as

$$\pi^{ij} = \dot{a}^{s\,ij} - \mathrm{i}\epsilon^{ikl}\partial_k a_l^{s\,j} + P^{ij|kl}\partial_k \phi_l, \tag{63}$$

and compute the Hamiltonian density as $\mathcal{H} = \pi^{ij}\dot{a}^s_{ij} - \mathcal{L}$. We get

$$\mathcal{H} = \frac{(\pi^{ij})^2}{2} + \mathrm{i}\pi^{ij}\epsilon_i{}^{kl}\partial_k a_{lj} + \phi_i \partial_j \pi^{ij}, \tag{64}$$

where we have dropped the index s from a^s_{ij} for brevity. Thus, now all the dynamical fields appearing in the Hamiltonian are symmetric tracefree tensors. The quantity ϕ_i is the Lagrange multiplier, which serves as a generator of SU(2) rotations on the connection. Indeed, the Poisson bracket of the integrated last term with the connection gives

$$\delta_\phi a_{ij} = \partial_{(i}\phi_{j)}, \tag{65}$$

which is just the (symmetrised) gauge transformation. To see the structure of the arising Hamiltonian it is convenient to fix the gauge and require the connection to be transverse

$$\partial_i a_{ij} = 0. \tag{66}$$

The momentum is required to be transverse by the condition obtained varying the action with respect to the Lagrange multipliers ϕ_i. So, it is now clear that the reduced phase space of our linearised system is parameterised by two symmetric, tracefree and transverse matrices a_{ij} and π_{ij}. This corresponds to two propagating DOF.

Let us now see what the dynamics becomes. To unravel the structure of the arising expression for the (reduced) Hamiltonian let us further rewrite it as

$$\mathcal{H} = \frac{1}{2}(\pi^{ij} + \mathrm{i}\epsilon^{ikl}\partial_k a_l^j)^2 + \frac{1}{2}(\partial_k a_{ij})^2. \tag{67}$$

Up to this point no reality conditions for the fields were specified. We can now deduce the linearised theory reality conditions from the Hamiltonian (67). Indeed, declaring the symmetric tracefree transverse connection field a_{ij} to be real, and defining a new real momentum field

$$p^{ij} := \pi^{ij} + \mathrm{i}\epsilon^{ikl}\partial_k a_l^j, \qquad p^{ij} \in \mathbb{R} \tag{68}$$

we can rewrite the linearised Hamiltonian in an explicitly positive definite form

$$\mathcal{H} = \frac{1}{2}(p^{ij})^2 + \frac{1}{2}(\partial_k a_{ij})^2. \tag{69}$$

The field equations that follow are now the usual

$$\Box\, a_{ij} = 0, \tag{70}$$

which is just the wave equation for the two components of the connection field a_{ij}. This is how gravitons are described by our gauge theory approach. We note that one can recognise in the analysis of this section the linearised version of the new Hamiltonian formulation of gravity [18]. In particular, the arising reality conditions for the phase space fields are the same as in this formulation. Thus, even though our starting point of a gauge theory is a bit unconventional, the linearised theory mimics constructions familiar from other formulations.

What is different about our linearised theory (49) from the more familiar treatment in [18] is that no diffeomorphism constraints are left in the final result. Instead, our linearised action is simply independent of certain components of the connection field, so the theory is formulated on a smaller configuration space to start with.

4 Propagator

In this section we invert the quadratic form that we have obtained by expanding the theory around the Minkowski spacetime background. In doing this we must decide on the gauge fixing.

4.1 Gauge fixing

We have seen that the action (49) is invariant under both gauge and diffeomorphism transformations, but we have also seen above that this invariance is manifested very differently in the two cases. Thus, in the case of the gauge invariance the situation is completely standard in that some of the field components have zero momenta and are thus Lagrange multipliers — generators of gauge symmetries. In the case of diffeomorphisms the situation is very different — we have seen that the action is simply independent of some components of the field, exactly those components that can be freely changed by performing a diffeomorphism. Thus, while there is very little choice for dealing with the gauge rotations — we have to treat them in the usual way by fixing the gauge and thus making the unphysical components of the gauge field propagate — we will need a different procedure for dealing with those components of the connection that gets affected by diffeomorphisms.

A useful analogy here is as follows. Let us consider a theory of two scalar fields ϕ, ψ with the Lagrangian

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu(\phi - \psi))^2. \tag{71}$$

It is clear that the Lagrangian is invariant under a simultaneous shift of both of the fields by some function. The way this is realised is that the Lagrangian is simply independent of a certain combination of the fields, namely of $\phi + \psi$, being only a function of the combination $\phi - \psi$. A natural quantisation strategy in this case is to introduce a new field $\phi - \psi$ and rewrite the Lagrangian in terms of the new field only. Then only this combination of the fields is a propagating field, while the other combination $\phi + \psi$ is a fiction.

In the case of the simple Lagrangian above it is very easy to see what the propagating field is. In our case (49) this is much harder. Thus, we will not be able to rewrite the Lagrangian in a way that has only diffeomorphism invariant combinations of the connection components appearing. However, an appropriate strategy is as follows. We can consider the quadratic form (49) as a form on the space of diffeomorphism invariant classes of connections a^i_μ, i.e. connections related via

$$a^i_\mu \sim a^i_\mu + \xi^\nu \Sigma^i_{\mu\nu}. \tag{72}$$

The quadratic form in (49) is degenerate on this space because there is still the usual gauge invariance to be taken care of. However, this gauge invariance can be dealt with in the usual way, by fixing the gauge. As we shall see below, it will be possible to find a gauge-fixing condition that is invariant under (72). After doing this we obtain a non-degenerate quadratic form on the space of diffeomorphism classes (72). It can be inverted, to obtain a propagator on the space of diffeomorphism classes of connections. As is standard for gauge-fixing, this procedure will make the temporal and longitudinal components of the connection propagate (and will add ghosts that will offset the effect of making this components propagating). At the same time, the components of the connection that are identified in (72) will not be propagating, as the propagator will involve a projector on the space of diffeomorphism equivalence classes. This way of dealing with the gauge symmetries of our theory is a bit non-standard, but is quite natural given that the two gauge symmetries are realised differently.

Having explained the logic of our procedure it remains to find a gauge-fixing condition that is diffeomorphism invariant. After some trial and error we found the following gauge-fixing condition to be useful:

$$\partial^\mu \Pi^{\mu i|\nu j} a_{\nu j} = \frac{2}{3}\partial^\mu \left(a^i_\mu + \frac{1}{2}\epsilon^{ijk}\Sigma^{k\nu}_\mu a^j_\nu \right) = 0, \tag{73}$$

where

$$\Pi^{\mu i|\nu j} := \eta^{\mu\nu}\delta^{ij} + \frac{1}{3}\Sigma^{i\,\mu\rho}\Sigma^{j\,\nu}_{\rho} = \frac{2}{3}\left(\eta^{\mu\nu}\delta^{ij} + \frac{1}{2}\epsilon^{ijk}\Sigma^{k\,\mu\nu}\right) \tag{74}$$

is a projector operator whose meaning is to be clarified below. The projector property

$$\Pi^{\mu i|\nu j}\Pi_{\nu j}{}^{\rho k} = \Pi^{\mu i|\rho k}, \tag{75}$$

can be checked by an elementary computation. It is easy to see that our gauge-fixing condition is diffeomorphism invariant. Indeed, consider

$$\xi^{\nu}\Sigma^{i}_{\mu\nu} + \frac{1}{2}\epsilon^{ijk}\Sigma^{k\,\nu}_{\mu}\xi^{\rho}\Sigma^{j}_{\nu\rho}. \tag{76}$$

Using the algebra (55) of $\Sigma^{i}_{\mu\nu}$ matrices we see that the last term here equals

$$\frac{1}{2}\epsilon^{ijk}\epsilon^{kjl}\xi^{\rho}\Sigma^{l}_{\mu\rho} = -\xi^{\nu}\Sigma^{i}_{\mu\nu}. \tag{77}$$

Thus, the quantity in (76) is zero, and the gauge-fixing condition (73) is diffeomorphism-invariant. It is also clear that as far as the gauge transformations are concerned the last term in (73) is inessential, for it is zero for any a^{i}_{μ} that is a pure gauge $a^{i}_{\mu} = \partial_{\mu}\phi^{i}$. Thus, (73) is the usual gauge theory gauge-fixing condition, corrected by a term that is inessential as far as the behaviour under the gauge transformations is concerned.

Let us now confirm that the projector $\Pi^{\mu i|\nu j}$ is just that on diffeomorphism equivalence classes of connections, and so it is natural to apply it before the usual gauge-fixing condition is imposed (to make this condition diffeomorphism invariant). We compute the action of the projector on the connection a^{j}_{ν} decomposed as in the previous subsection

$$a^{j}_{\nu} = a^{j}_{0}(dt)_{\nu} + (a^{s}_{ij} + b\delta_{ij} + \epsilon_{ijk}c_{k})(dx^{i})_{\nu}. \tag{78}$$

The result is

$$\Pi^{\mu i|\nu j}a^{j}_{\nu} = \frac{2}{3}\left(\delta^{ij}\left(\frac{\partial}{\partial t}\right)^{\mu} + \frac{i}{2}\epsilon^{ijk}\left(\frac{\partial}{\partial x^{k}}\right)^{\mu}\right)(a^{j}_{0} - \mathrm{i}c^{j}) + a^{s}_{ij}\left(\frac{\partial}{\partial x^{j}}\right)^{\mu}. \tag{79}$$

We note the the quantity b got projected out, and the projected connection only depends on the temporal and the anti-symmetric spatial components of the connection in the combination $a^{i}_{0} - \mathrm{i}c^{i}$, as expected from the previous section. Thus, the projector $\Pi^{\mu i|\nu j}$ is indeed just that on the diffeomorphism invariant subspace, and selects the components $a^{i}_{0} - \mathrm{i}c^{i}$, which play the role of the generators of the Gauss constraints, as well as a^{s}_{ij}, which are the two propagating DOF plus three longitudinal modes of the connection. As usual for a gauge theory we shall make the components generators of the Gauss constraints as well as the longitudinal components of the connection propagating by adding a gauge fixing term, and then offset their effects by adding ghosts.

We now add the gauge-fixing condition squared with some parameter to the Lagrangian. Thus, we consider the following gauge-fixed Lagrangian on the space of diffeomorphism equivalence classes of connections

$$\mathcal{L}_{\rm gf} = -\frac{1}{2}P^{ij|kl}(\Sigma^{i\,\mu\nu}\partial_{\mu}a^{j}_{\nu})(\Sigma^{k\,\rho\sigma}\partial_{\rho}a^{l}_{\sigma}) - \frac{\alpha}{2}\left(\partial^{\mu}a^{i}_{\mu} - \frac{1}{2}\epsilon^{ijk}\Sigma^{j\,\mu\nu}\partial_{\mu}a^{k}_{\nu}\right)^{2}, \tag{80}$$

where we have changed the order of indices jk in the gauge-fixing term for convenience, and absorbed the $(2/3)^2$ factor into the gauge-fixing parameter α. As in the case of Yang-Mills theory, the idea is now to select the gauge-fixing parameter α so that the gauge-fixed action is as simple as possible.

4.2 The algebra of gauge-fixing

In this subsection we will simplify the expression for the gauge-fixed Lagrangian and find a useful value for the gauge-fixing parameter α. To this end, let us first write the Lagrangian in the momentum space. Omitting the argument $\pm k$ from the Fourier components $a^i_\mu(k)$ of a^i_μ for brevity we have the following expression

$$\mathcal{L}_{\rm gf} = -\frac{1}{2}P^{ij|kl}(\Sigma^{i\,\mu\nu}k_\mu a^j_\nu)(\Sigma^{k\,\rho\sigma}k_\rho a^l_\sigma) - \frac{\alpha}{2}\left(k^\mu a^i_\mu - \frac{1}{2}\epsilon^{ijk}\Sigma^{j\,\mu\nu}k_\mu a^k_\nu\right)^2. \tag{81}$$

Let us expand the last term. Introducing a compact notation $(ka^i) := k^\mu a^i_\mu$ and expanding the product of two ϵ's we have

$$\begin{aligned}\left((ka^i) - \frac{1}{2}\epsilon^{ijk}\Sigma^{j\,\mu\nu}k_\mu a^k_\nu\right)^2 &= (ka^i)^2 - (ka^i)\epsilon^{ijk}\Sigma^{j\,\mu\nu}k_\mu a^k_\nu \\ &+\frac{1}{4}\Sigma^{i\mu\nu}\Sigma^{i\,\rho\sigma}k_\mu k_\rho a^j_\nu a^j_\sigma - \frac{1}{4}(\Sigma^{i\mu\nu}k_\mu a^j_\nu)(\Sigma^{j\,\rho\sigma}k_\rho a^i_\sigma).\end{aligned} \tag{82}$$

Let us now expand the first term of the Lagrangian. We have

$$\begin{aligned}P^{ij|kl}(\Sigma^{i\,\mu\nu}k_\mu a^j_\nu)(\Sigma^{k\,\rho\sigma}k_\rho a^l_\sigma) &= \frac{1}{2}\Sigma^{i\mu\nu}\Sigma^{i\,\rho\sigma}k_\mu k_\rho a^j_\nu a^j_\sigma \\ +\frac{1}{2}(\Sigma^{i\mu\nu}k_\mu a^j_\nu)(\Sigma^{j\,\rho\sigma}k_\rho a^i_\sigma) &- \frac{1}{3}(\Sigma^{i\,\mu\nu}k_\mu a^i_\nu)(\Sigma^{j\,\rho\sigma}k_\rho a^j_\sigma).\end{aligned} \tag{83}$$

We can now use the following two identities

$$\Sigma^{i\mu\nu}\Sigma^{i\,\rho\sigma} = \eta^{\mu\rho}\eta^{\nu\sigma} - \eta^{\nu\rho}\eta^{\mu\sigma} - \mathrm{i}\epsilon^{\mu\nu\rho\sigma} \tag{84}$$

and

$$\Sigma^{i\mu\nu}\Sigma^{j\,\rho\sigma} - \Sigma^{j\mu\nu}\Sigma^{i\,\rho\sigma} = \epsilon^{ijk}\left(\Sigma^{k\,\mu\sigma}\eta^{\nu\rho} - \Sigma^{k\,\nu\sigma}\eta^{\mu\rho} - \Sigma^{k\,\mu\rho}\eta^{\nu\sigma} + \Sigma^{k\,\nu\rho}\eta^{\mu\sigma}\right). \tag{85}$$

We can now use the identity (84) to rewrite the first term in (83), and the identity (85) to rewrite the last term as a multiple of the second plus some extra terms. We get

$$\frac{1}{2}(k^2(a^i_\mu)^2 - (ka^i)^2) + \frac{1}{6}(\Sigma^{i\mu\nu}k_\mu a^j_\nu)(\Sigma^{j\,\rho\sigma}k_\rho a^i_\sigma) + \frac{1}{3}\left(k^2\epsilon^{ijk}\Sigma^{i\,\mu\nu}a^j_\mu a^k_\nu + 2(ka^i)\epsilon^{ijk}\Sigma^{j\,\mu\nu}k_\mu a^k_\nu\right). \tag{86}$$

We now note that if we make a choice

$$\alpha = \frac{2}{3} \tag{87}$$

then the terms $(\Sigma^{i\mu\nu}k_\mu a^j_\nu)(\Sigma^{j\,\rho\sigma}k_\rho a^i_\sigma)$, as well as $(ka^i)^2$ and $(ka^i)\epsilon^{ijk}\Sigma^{j\,\mu\nu}k_\mu a^k_\nu$ cancel out and we get the following simple gauge-fixed action

$$\mathcal{L}_{\rm gf} = -\frac{k^2}{3}\left((a^i_\mu)^2 + \frac{1}{2}\epsilon^{ijk}\Sigma^{k\,\mu\nu}a^i_\mu a^j_\nu\right) = -\frac{k^2}{2}\Pi^{\mu i|\nu j}a_{\mu i}a_{\nu j}, \tag{88}$$

where $\Pi^{\mu i|\nu j}$ is the projector (74). Because the projector on diffeomorphism equivalence classes appears here explicitly, it is obvious that this action is still invariant under the diffeomorphisms (72), and so is now a non-degenerate quadratic form on the space of diffeomorphism equivalence classes.

4.3 Propagator

We now invert the quadratic form in (88). Thus, we add a current term to the action

$$S_{\rm gf} = \int \frac{d^4k}{(2\pi)^4}\left[-\frac{k^2}{2}\Pi^{\mu i|\nu j}a_{\mu i}(-k)a_{\nu j}(k) + J^{\mu\, i}(-k)a^i_\mu(k)\right], \tag{89}$$

and then integrate the field a^i_μ out. This can be easily done in the space of diffeomorphism equivalence classes, and we immediately see that the action with the original connection field integrated out is given by

$$S[J] = \int \frac{d^4k}{(2\pi)^4}\frac{1}{2k^2}\Pi^{\mu i|\nu j}J^i_\mu(-k)J^j_\nu(k). \tag{90}$$

In other words, the propagator of our theory is given by

$$\langle a^{\mu i}(-k)a^{\nu j}(k)\rangle = (1/\mathrm{i})^2\frac{\delta}{\delta J^i_\mu(-k)}\frac{\delta}{\delta J^j_\nu(k)}e^{\mathrm{i}S[J]}\Big|_{J=0} = (1/\mathrm{i})\frac{1}{k^2}\Pi^{\mu i|\nu j}, \tag{91}$$

which is just the usual $1/k^2$ term times the projector onto the space of diffeomorphism equivalence classes of connections, times the (convention dependent) 1/i factor.

This finishes our discussion of the free theory of gravitons on the Minkowski spacetime background (or gravitons with energy $E \gg M_0$ much greater than the energy scale of our constant curvature background). We refrain from considering ghosts that are irrelevant for our purely classical purposes in this paper. Instead, let us now consider the lowest order interactions.

5 Interactions

In this section we consider graviton self-interactions.

5.1 Third variation of the action

The third variation of the action is easily computed from (26). We get

$$\delta^3 S[A] = (1/\mathrm{i})\int_M d^4x\,\left(\frac{\partial^3 f}{\partial\tilde X^{ij}\partial\tilde X^{kl}\partial\tilde X^{pq}}\delta\tilde X^{ij}\delta\tilde X^{kl}\delta\tilde X^{pq} + 3\frac{\partial^2 f}{\partial\tilde X^{ij}\partial\tilde X^{kl}}\delta^2\tilde X^{ij}\delta X^{kl} + \frac{\partial f}{\partial\tilde X^{ij}}\delta^3 X^{ij}\right). \tag{92}$$

We have already computed the first and second variations of the matrix $\tilde X^{ij}$ in (27), (28). The third variation is given by

$$\delta^3\tilde X^{ij} = \frac{3}{2}\tilde\epsilon^{\mu\nu\rho\sigma}D_{A\,\mu}\delta A^{(i}_\nu\epsilon^{j)kl}\delta A^k_\rho\delta A^l_\sigma. \tag{93}$$

We also note that the fourth variation, of relevance for higher-order interaction vertices, is zero, which follows by expanding the product of two ϵ's and noting that there is always a δ^{ij}-contraction of two variations of the connection. On the other hand, spacetime indices of all 4 variations of the connection are contracted with $\tilde\epsilon^{\mu\nu\rho\sigma}$, and so the result is zero.

5.2 Cubic interaction

We have already computed the first and second derivatives of the defining function at the identity matrix in (42), (43). Let us now compute the third derivative. We get:

$$\frac{\partial^3 f}{\partial X^{ij}\partial X^{kl}\partial X^{pq}}\Big|_{X_0} = -\frac{2g_0}{3(-2\mathrm{i}M_0^4)^2}\left(\delta^{ij}P^{kl|pq} + \delta^{kl}P^{ij|pq} + \delta^{pq}P^{ij|kl}\right), \tag{94}$$

where g_0 is the dimensionless constant given by (48), and $P^{ij|kl}$ is the projector on the symmetric traceless part that we already encountered above.

We now compute the cubic interaction term. We evaluate (92) at the constant curvature background connection A_0. The last term in (92) is then seen to be a total derivative. We can also note that of the two terms coming from $\delta^2 X^{ij}$ one term is proportional to $(D\delta A)^2$, while the other is of the order $M_0^2(\delta A)^2$. For energies $E \gg M_0$ we can neglect the term $M_0^2(\delta A)^2$. Then, after some rewriting we get

$$\delta^3 S\Big|_{A_0} = \frac{g_0}{2M_0^2}\int d^4x\sqrt{-g}\,P^{ij|kl}(\Sigma^{i\,\mu\nu}D_{A_0\,\mu}\delta A^j_\nu)\Big[(\Sigma^{k\,\rho\sigma}D_{A_0\,\rho}\delta A^l_\sigma)(\Sigma^{m\,\alpha\beta}D_{A_0\,\alpha}\delta A^m_\beta) \\ -3\mathrm{i}\,\epsilon^{\alpha\beta\gamma\delta}D_{A_0\,\alpha}\delta A^k_\beta\,D_{A_0\,\gamma}\delta A^l_\delta\Big]. \tag{95}$$

Now passing to the high-energy limit $E \gg M_0$ we replace the covariant derivatives by the usual coordinate ones, and then rewrite the interaction term in terms of the connection field $a^i_\mu = \sqrt{g_0}(\delta A^i_\mu)$, for which the kinetic term (49) is canonically normalised. We also need to divide the third variation by 3! to get the correct cubic interaction term. We get:

$$S^{(3)} = \frac{1}{12\sqrt{g_0}M_0^2}\int d^4x\,P^{ij|kl}(\Sigma^{i\,\mu\nu}\partial_\mu a^j_\nu)\Big[(\Sigma^{k\,\rho\sigma}\partial_\rho a^l_\sigma)(\Sigma^{m\,\alpha\beta}\partial_\alpha a^m_\beta) - 3\mathrm{i}\,\epsilon^{\alpha\beta\gamma\delta}\partial_\alpha a^k_\beta\partial_\gamma a^l_\delta\Big]. \tag{96}$$

This is trivially invariant under (54), and can be checked to be diffeomorphism invariant if one takes into account the second-order contribution to the connection transformation law and the resulting order three contribution from the quadratic part of the action.

To summarise, schematically, the cubic interaction is of the form

$$\mathcal{L}^{(3)} \sim \frac{1}{\sqrt{g_0}M_0^2}(\partial a)^3. \tag{97}$$

We learn that our theory of gravity has a negative mass dimension coupling constant, and so is non-renormalisable in the usual sense of the word, as it should be. We can also determines the Newton's constant and thus the Planck mass from the graviton self-coupling. We see that M_p is a derived quantity in our theory, given by

$$M_p^4 \sim g_0 M_0^4, \tag{98}$$

where g_0 is the dimensionless parameter (48).

Before we turn to a discussion of implications of (98), let us briefly compare the obtained interaction term (97) with that in the usual metric description. Note, however, that the precise form of the interaction vertex depends on details of the gauge-fixing (and can be simplified by a careful choice of the latter, see e.g. [19]). But we can compare the general structure of our cubic interaction vertex with that in the metric approach. One immediately notes that the cubic interaction in our gauge theory approach contains three derivatives, while that in the metric approach contains only two. This is why the dimension of the coefficient in front of our cubic term is $1/M^2$, while that in the usual case is $1/M$. The reason for this seeming discrepancy is that the relation between our gravitons and those of the metric treatment is quite non-trivial. Essentially, we describe gravitons using the components of the gravitational spin connection, which are given by the derivative of the metric perturbations. Thus, a derivative operator is involved in the relation between our gauge theory description and the usual metric one. This is why the power of the derivative operator in the vertices can be different in the two theories.

An attempt at a detailed comparison between the two perturbation theories would take us too far. Let us just say that one possible way to do this is to "integrate in" the two-form field as an independent field, and then re-express the perturbation theory in the language of the two-form perturbations, as was done in [12]. Thus, it is clear that the relation between our description in terms of a connection A and the usual metric one is that of duality, in the sense that a new field, say B, is introduced, such that when B is integrated out the original "pure-A" formulation results. Then one can instead integrate out the field A of the original formulation, and obtain a dual formulation in terms of B. This trick is used in theoretical physics on many occasions, and the relationship between our gauge-theoretic formulation and the usual metric one is of the same sort.

5.3 The cosmological constant problem

In order for our gauge theory (around the background chosen) to describe the usual gravitons whose self-interaction is controlled by the coupling constant of the known value $M_p \sim 10^{19} GeV$, the dimensionless parameters of our theory f_0, see (40), and g_0, see (48), must satisfy

$$g_0 \sim 10^{120} f_0. \tag{99}$$

Note that this relation is independent of the chosen value of the background curvature controlling parameter M_0. In words, the ratio of the sum of the two first derivatives of the defining function evaluated at the identity matrix, to its value at this point must be of the order 10^{120}. Thus, the defining function that would reproduce the known graviton self-interaction must be very special indeed!

Recalling that the constants g_0, f_0 appear as the dimensionless parameters multiplying the matrices of second and first derivatives of the defining function, we can rephrase our condition (99) by saying that the matrix of the second derivatives of the defining function must be much larger than that of the first derivatives:

$$\left| \mathrm{Tr}(X_0) \left(\frac{\partial^2 f}{\partial X^{ij} \partial X^{kl}} \right)_{X_0} \right| \sim 10^{120} \left| \left(\frac{\partial f}{\partial X^{ij}} \right)_{X_0} \right|. \tag{100}$$

In other words, the defining function that reproduces GR at low energies should be very steep at the point corresponding to a constant curvature connection.

It is clear that the famous cosmological constant problem, which is to explain how such a large dichotomy of scales as in (99) can exist in Nature, has received in our theory a new form. Indeed, the famous ratio of scales present in (99) has been encoded into properties of the low energy defining function of the theory. Recall now that the defining function for this theory should be expected to run with energy. Thus, once the corresponding renormalisation group flow is obtained, the theory must explain how the condition (99) can arise at low energies. The only currently imaginable to the author explanation of this phenomenon is that the fine-tuning (99) at low energies is necessary in order for the renormalisation group flow to arrive at some desired (fixed) point at high energies. But at this point of the development of the theory it remains a speculation. A derivation and analysis of the RG flow for our theories is left to future work.

We also note that the need for a defining function with property (99) will seem much less strange if we recall the result obtained in [7]. Indeed, this work showed that general relativity can be written in the form of a diffeomorphism invariant gauge theory provided the defining function is chosen to be a δ-function imposing the condition

$$\mathrm{Tr}(X)^2 = \frac{1}{2}(\mathrm{Tr}(X))^2. \tag{101}$$

In the paper [7] this condition has been imposed via an extra Lagrange multiplier field. However, we can also impose it in our framework, by choosing the function χ appropriately. Then the first derivative of the function χ with respect to the argument $\mathrm{Tr}(X)^2/(\mathrm{Tr}(X))^2$, evaluated at the value 1/3 of this argument corresponding to the identity matrix, can be quite large for the function χ being close to a δ-function. We refrain from making any estimates here because at this stage of the development of the theory it is hard to guess what kind of functions χ are natural from the point of view of the renormalisation group flow. But the fact that general relativity (in Minkowski spacetime) arises for a very special δ-function-like choice [7] of the defining function makes the property (99) that must hold at low energies less surprising.

6 Discussion

In this paper we have proposed a new approach to the gravitational perturbation theory. While our main motivation was the quantum theory (renormalisation), in the present paper we remained in the classical domain. We have recalled how a diffeomorphims invariant gauge theory can be formulated using a homogeneous degree one defining function, and how such a theory for the gauge group SU(2) is a gravity theory describing two propagating degree of freedom. Our main interest here was in the perturbation theory. Hence, we expanded our general diffeomorphims invariant gauge theory Lagrangian around a constant curvature connection. The original theory does not have any dimensionful parameters, and we have seen that it is the choice of the background that brings in a dimensionful quantity into the game, in our case the radius of curvature of the background. We then took a limit of the radius of curvature becoming very large (or working at energies such that the curvature of the background can be neglected). This way we obtained a theory on the Minkowski spacetime background.

The linearised action (49) we obtained is quite simple, and can be seen to be a natural construct involving the linearised connection, as well as the basic (anti-) self-dual two-forms $\Sigma^i_{\mu\nu}$. Indeed, as is sometimes done in the literature, one can introduce the derivative operators $\partial^{\mu\, i} := \Sigma^{\mu\nu\, i}\partial_\nu$. The basic building block of our linearised action is then $\partial^{\mu\, i} a^j_\mu$, where this quantity is symmetrised and then its tracefree part is squared to form the action. Note that the projector $P^{ij|kl}$ on the symmetric tracefree part is just that on the spin two part of the tensor product of two spin one representations, and this is another manifestation of how the spin two appears in the game. Indeed, one could rewrite our linearised gauge theory action using the spinor notation as a multiple of $(\Sigma^{\mu\nu\,(AB}\partial_\mu a_\nu^{CD)})^2$, where the brackets denote the symmetrisation. A completely symmetrised rank 4 spinor is the standard realisation of the spin two representation.

What we are seeing here is the appearance of a new type of gauge theory actions in Minkowski space, constructed using the two-forms $\Sigma^i_{\mu\nu}$ instead of the metric. We have also seen that the effect of the spin two projector $P^{ij|kl}$ in the action is that the linearised action is independent of some of the components of the gauge field, and this is how the diffeomorphism invariance is realised in our theory. The Hamiltonian analysis of the linearised theory then confirmed the count of the number of propagating modes, and also exhibited the standard spin two graviton Hamiltonian.

A bit non-standard point of our construction was our strategy of dealing with the diffeomorphisms. Thus, we did not fix the gauge for them as is usual in field theory (making the unphysical modes propagate and adding ghosts). Instead, we decided to simply project these components out. The justification for this procedure is that the linearised action is independent of the projected out components, and so it is certainly a legitimate procedure at the linearised level. We have then dealt with the other gauge symmetry, the usual SU(2) gauge rotations, in the standard way. Our procedure of projecting out the diffeomorphism equivalent components of the connection requires some thought if it is to be extended to the full interacting theory, but we believe that it is still a viable procedure for the theory linearised around an arbitrary background, and thus for the background field method. This is to be developed in the forthcoming works.

Another justification for our strategy of dealing with the diffeomorphisms is that the gauge-fixed action that we obtained is extremely simple. It is just the simplest possible $a\Box a$ type action, with an additional projector on diffeomorphism equivalence classes inserted. This simplification of the propagator should be of great help in the future quantum calculations. The obtained simple form of the gauge-fixed action gives one more justification for the name of our approach to gravity. Indeed, the gauge-fixed linearised action is quite clearly just the usual gauge-fixed gauge theory, with, in addition, a projector taking into account the diffeomorphisms inserted. Thus, our gravity theory is just a diffeomorphism invariant gauge theory, precisely as the title of this paper suggests.

Apart from the quantum aspects, which we purposefully decided to avoid here, we did not comment much on the subtle issue of the reality conditions for our theory. Indeed, these were discussed at the linearised level, where their treatment is no different from that in the Ashtekar formulation, see [18]. It is clear, however, that the full interacting action will require a much more sophisticated choice of the reality conditions. For the quantum calculations to be carried out with this formalism this is not much of an issue, because all loops are computed via the trick of the analytic continuation, and under this all factors of $\sqrt{-1}$ in our formulas disappear and fields become real. However, these issues do matter for the questions of the unitarity of the arising quantum theory. We expect that these subtle issues will take some time to be settled, and refrain from trying to address them in this work.

Let us close by making some further comments on the type of gravitational perturbation theory that we have seen arising in this approach. Thus, (98) tell us that the strength of the graviton self-interactions is determined in this theory by a dimensionful parameter that arises as a combination of the radius of curvature of the background with a certain dimensionless coupling constant g_0 of the theory. This implies that the Newton's constant is no longer a fundamental parameter of our theory. We have also seen that the cosmological constant (or rather the combination Λ/G) can be expressed in terms of the radius of curvature and a dimensionless parameter as well, see (40). If the expectation that it is the defining function f that runs with energy in this theory is correct, then both Λ/G and $1/G^2$ would run as f_0 and g_0 parameters respectively (provided the background curvature can be chosen to be constant). Given the dimensionless nature of the latter couplings, this may be quite different from the running expected from the mass dimensions of Λ/G and $1/G^2$. However, at this stage it is hard to say anything concrete about how the resulting flow might look like. It is only clear from the estimate (99) that this renormalisation group flow might have some important cosmological constant problem message to tell us.

To conclude, we hope to have convinced the reader that the present gauge-theoretic approach to gravity brings with itself many rather exciting opportunities that are simply unavailable, or impractical in the usual metric setting. It now seems within reach that, with the new tools developed here, the renormalisation group flow for an infinite parametric class of gravity theories can be computed. Once this is achieved, ideas about the ultra-violet behaviour of gravity, e.g. the asymptotic safety conjecture [16], can be explicitly tested.

Acknowledgements

The author is grateful to J. Fine and D. Panov for making a draft of the paper on "A gauge theoretic approach to the anti-self-dual Einstein equations" available to him prior to publication. Some of its constructions made the author take the "pure connection" formulation of diffeomorphism invariant gauge theories more seriously, which resulted in the present paper.

References

[1] R. P. Feynman, F. B. Morinigo, W. G. Wagner and B. Hatfield, "Feynman lectures on gravitation," *Reading, USA: Addison-Wesley (1995) 232 p. (The advanced book program)*

[2] Z. Bern, T. Dennen, Y. T. Huang and M. Kiermaier, "Gravity as the Square of Gauge Theory," Phys. Rev. D **82**, 065003 (2010) [arXiv:1004.0693 [hep-th]].

[3] J. F. Plebanski, "On the separation of Einsteinian substructures," J. Math. Phys. **18**, 2511 (1977).

[4] K. Krasnov, "Plebanski Formulation of General Relativity: A Practical Introduction," arXiv:0904.0423 [gr-qc].

[5] A. Ashtekar, "New Hamiltonian Formulation of General Relativity," Phys. Rev. D **36**, 1587 (1987).

[6] T. Jacobson and L. Smolin, "Covariant Action for Ashtekar's Form of Canonical Gravity," Class. Quant. Grav. **5**, 583 (1988).

[7] R. Capovilla, T. Jacobson and J. Dell, "General Relativity Without the Metric," Phys. Rev. Lett. **63**, 2325 (1989).

[8] I. Bengtsson, "The Cosmological constants," Phys. Lett. B **254**, 55 (1991).

[9] C. Rovelli, "Loop quantum gravity," Living Rev. Rel. **11**, 5 (2008).

[10] K. Krasnov, "Renormalizable Non-Metric Quantum Gravity?," arXiv:hep-th/0611182.

[11] I. Bengtsson, "Note on non-metric gravity," Mod. Phys. Lett. A **22**, 1643 (2007) [arXiv:gr-qc/0703114].

[12] K. Krasnov, "Effective metric Lagrangians from an underlying theory with two propagating degrees of freedom," Phys. Rev. D **81**, 084026 (2010) [arXiv:0911.4903 [hep-th]].

[13] K. Krasnov, "Metric Lagrangians with two propagating degrees of freedom," Europhys. Lett. **89**, 30002 (2010) [arXiv:0910.4028 [gr-qc]].

[14] K. Krasnov, "Gravity as BF theory plus potential," Int. J. Mod. Phys. A **24**, 2776 (2009) [arXiv:0907.4064 [gr-qc]].

[15] A. Torres-Gomez and K. Krasnov, "Gravity-Yang-Mills-Higgs unification by enlarging the gauge group," Phys. Rev. D **81**, 085003 (2010) [arXiv:0911.3793 [hep-th]].

[16] S. Weinberg, "Effective Field Theory, Past and Future," PoS C **D09**, 001 (2009) [arXiv:0908.1964 [hep-th]].

[17] A. Torres-Gomez, K. Krasnov and C. Scarinci, "A Unified Theory of Non-Linear Electrodynamics and Gravity," arXiv:1011.3641 [gr-qc].

[18] A. Ashtekar, "Lectures on nonperturbative canonical gravity," *Singapore, Singapore: World Scientific (1991) 334 p. (Advanced series in astrophysics and cosmology, 6)*

[19] A. E. M. van de Ven, "Two loop quantum gravity," Nucl. Phys. B **378**, 309 (1992).