Hypersurface foliation approach to renormalization of ADM formulation of gravity

I. Y. Park

1 Department of Physics, Hanyang University, Seoul 133-791, Korea
2 Department of Applied Mathematics, Philander Smith College, Little Rock, AR 72223, USA

Received: 4 July 2015 / Accepted: 3 September 2015 / Published online: 28 September 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract We carry out ADM splitting in the Lagrangian formulation and establish a procedure in which (almost) all of the unphysical components of the metric are removed by using the 4D diffeomorphism and the measure-zero 3D symmetry. The procedure introduces a constraint that corresponds to the Hamiltonian constraint of the Hamiltonian formulation, and its solution implies that the 4D dynamics admits an effective description through 3D hypersurface physics. As far as we can see, our procedure implies potential renormalizability of the ADM formulation of 4D Einstein gravity for which a complete gauge-fixing in the ADM formulation and hypersurface foliation of geometry are the key elements. If true, this implies that the alleged unrenormalizability of 4D Einstein gravity may be due to the presence of the unphysical fields. The procedure can straightforwardly be applied to quantization around a flat background; the Schwarzschild case seems more subtle. We discuss a potential limitation of the procedure when applying it to explicit time-dependent backgrounds.

1 Introduction

The quantization of 4D Einstein–Hilbert action has been a long-standing problem. (See [1] for a review.) One-loop renormalizability of pure 4D gravity (i.e., gravity without any matter field coupled) was established in [2]. However, the presence of matter fields does not preserve the one-loop renormalizability [3, 4]. Also, it was subsequently shown [5] that two-loop and higher order diagrams of the pure gravity require proliferation of counter terms, thereby leading to unrenormalizability. Needless to say, the lack of renormalizability has been a serious obstacle that has delayed (or even blocked) progress in many fundamental issues such as the black hole information paradox.

There have been several approaches to the quantization of gravity. The first was through the Hamiltonian formulation of gravity [6, 7]. (Earlier discussions can be found, e.g., in [8–11].) The obstacle in this approach was the complexity of the Hamiltonian constraint; a different approach based on a different set of variables was proposed and is now known as loop quantum gravity [12–14]. Still another approach is based on adding higher derivative terms in the action [15, 16].

In this work, we show by conducting the 3 + 1 splitting in the Lagrangian formulation that the dynamics of the pure 4D gravity is effectively reduced.1 to the dynamics of pure “3D gravity.” (As is well known, the genuine 3D gravity does not have a graviton. Our “3D” theory – originating from 4D gravity – is not the genuine 3D gravity; as we will see, it has two propagating degrees of freedom that are inherited from 4D. It holographically represents the original 4D system as shown in Fig. 1.) Since pure 3D gravity is known to be renormalizable [19, 20], our result implies that the 4D Einstein gravity is renormalizable: renormalizability of the 4D gravity is achieved essentially by removing all of the unphysical degrees of freedom in the particular manner which we will describe in detail in the main body. Some of the analyses in this work have been repeated from different angles in [21–23].

It may be good to clearly state at this point the task undertaken in the present work. The equivalence between the usual formulation and ADM formulation of general relativity was questioned in [24]. In our view, it is a legitimate concern given, e.g., that the ADM formulation may not generally be applicable to an arbitrary spacetime but is most useful in dealing with a globally hyperbolic spacetime. The ADM formulation should be applicable to a “locally hyperbolic” spacetime

1 To our pleasant surprise, we recently have found the work of [17] where a Hamiltonain reduction led to lower dimensional volume. There also is an earlier related work [18] We will comment on these further in the conclusion.
when the issue under consideration is local in nature such as renormalizability. Furthermore, once one takes the gauge in which the shift vector is set to zero, the analysis becomes essentially that of the usual formulation in the synchronous gauge. It is certainly possible to repeat all of the analysis in this work without referring to the ADM formulation but rather within the usual Lagrangian and Hamiltonian formulations with the synchronous gauge. Since gauge-fixing is a subtle issue in general, we clearly state the task undertaken in this work: quantization associated with the physical states dictated by the ADM formulation with the synchronous type gauge.

The focus of the present work is the possibility that the alleged unrenormalizability may be due to the presence of unphysical degrees of freedom in the 4D Einstein action. There are eight unphysical degrees of freedom, and four of them are associated with the 4D diffeomorphism. The other four degrees of freedom are non-dynamical, and they may explicitly be removed. In principle, there is a chance that such a removal may eliminate the need for some of the counter terms, and we confirm this anticipation in a dramatic way.

Basically, we remove the unphysical fields by gauge fixings. (It may appear that the procedure depends on the gauge-fixing too sensitively. We will have more to say on this in the conclusion.) The field equations of the lapse function and shift vector do not have any time derivative acting on them. Therefore, it should be unnecessary to use the bulk gauge symmetry to gauge-fix them. Instead, one may use the residual symmetry after the bulk fixing to gauge-fix these non-dynamic fields. (This idea may be related to the one in [10,11].) Gauge-fixing of the bulk diffeomorphism will be carried out in the standard manner; it is the gauge-fixing of the non-dynamical fields that makes the difference.\(^2\) In terms of Dirac’s terminology [25], the non-dynamical fields introduce the first class constraints in the Hamiltonian formalism. In the spirit of the comment made in section 7.6 of [26] (in which it was stated that the first class constraints can be eliminated by a choice of gauge), we deal with the Lagrangian analog of the first class constraints simply by additional gauge-fixings. In the Hamiltonian approach, the first class constraints are associated with gauge symmetries. The necessity of such a gauge-fixing was discussed in [1,27]. Since the non-dynamical fields do not evolve in time and are thus virtually three-dimensional, they should be viewed as generating three-dimensional (as opposed to four-dimensional) symmetries. Because of this, a measure-zero symmetry will be used to remove the non-dynamical fields. (See [21] for an explicit and quantitative analysis on the nature of the residual symmetry and ghost-term related issues.)

In spite of the unphysical nature of the non-dynamical fields, the conventional procedure includes them in the renormalization program: they appear as external lines of loop diagrams in the Green function computations. They also run around the loops of various loop diagrams. The main reason for not gauge-fixing them should be maintaining the covariance. (Moreover, such a procedure was successful in the cases of gauge theories.) However, if there exists a (non-covariant) renormalization procedure that does not include the non-dynamical fields and is still controllable, it would be worthwhile to examine the procedure. We show that there indeed exists such a procedure that maintains 3D covariance.

Since we claim in this work that a different (and complete) gauge-fixing may render the 4D gravity renormalizable,\(^3\) one may raise a question how that could happen given that physics should not depend on a gauge choice. Firstly, let us recall that only the correlator of gauge invariant quantities are independent of the gauge-fixing; in a renormalization procedure, one considers correlators of the elementary fields – which may or may not be gauge singlets (for example, the metric is not a gauge singlet). Therefore, renormalization procedure does depend on the gauge choice, and we have demonstrated this in the Appendix. What we observe in this work is that the dependence can be drastic. A second clue to the answer to the question above can be found in [28]. We will come back to this work later but briefly the authors noted that the presence of an unphysical mode makes the path integral divergent and ill-defined. The divergence here is something that one should worry about before one examines renormalizability of the theory, and is worse in that sense than unrenormalizability. Therefore, it seems reasonable to expect that removing all, as opposed to four (as in, e.g., the de Donder gauge) out of eight, of the unphysical modes can, in principle, make a drastic difference. Therefore, gauge-choice independence of the physics including the renormalizability would be expected for different gauges as far as they are complete fixings.

\(^2\) In this work we adopt the approach in which all of the unphysical degrees of freedom are explicitly fixed. Alternatively one may keep the off-shell unphysical degrees of freedom that are associated with the lapse and shift. See [23] for this approach.

\(^3\) The analogous phenomenon does not occur in gauge theories since gauge theories are renormalizable in the Lorenz gauge that keeps the non-dynamical time component. It seems that gravity is just different from gauge theories in this aspect.
There exists a gauge, called the radiation gauge [29, 30], in which all of the unphysical fields are removed. The gauge-fixing procedure that we adopt keeps only the physical degrees of freedom as in the radiation gauge, but is different from the radiation gauge in that it leads to effective reduction of the dynamics onto the hypersurface. The hypersurface foliation approach in this work provides a convenient arena for examining the renormalizability after removal of the unphysical degrees of freedom. Let us consider the $3 + 1$ splitting of the 4D Einstein–Hilbert action. Compared with the existing covariant approach, the ADM formulation employed in this manuscript has two advantages: firstly, the ADM formulation is effective in organizing the degrees of freedom for easy isolation of the unphysical degrees of freedom. In other words, the formulation readily identifies the non-dynamical fields thereby setting the stage for their removal. Secondly, the ADM splitting brings out the utility of the measure-zero gauge symmetry for removal of the non-dynamic fields. This feature will play a crucial role in Sect. 2.

The rest of the paper is organized as follows. In Sect. 2, we discuss the removal of unphysical degrees of freedom. After starting off by recalling the gauge-fixing procedure of a YM type gauge theory, we note that it should be possible to gauge-fix the lapse function and shift vector by the measure-zero diffeomorphisms after fixing the bulk gauge symmetry (the de Donder gauge will be adopted for the bulk fixing). As is well known (see, e.g. [35]), the choice of the lapse and shift is arbitrary, hence gauge-fixing the lapse and shift should be a legitimate procedure in any case. What is important is that it should be possible to gauge-fix them by using the measure-zero gauge symmetry, not the bulk ones. The result of these gauge-fixings is the projection of 4D dynamics onto the 3D hypersurface: the 3D system has two physical degrees of freedom inherited from 4D. (The reduction is limited to a pure Einstein system; a matter field, if present, will not be reduced, at least not in any simple way.) By invoking the logic of [5, 20] one arrives at renormalizability. This will be pointed out in Sect. 3, in which we will also comment on precisely what physics the reduced theory should describe. The procedure should be viewed as a generalization of the holography idea of [36]. A flat spacetime will be considered throughout. It should be possible to apply the procedure to a Schwarzschild black hole background with relatively minor modifications. However, it is not clear whether the procedure can be applied to a more complex background such as an explicit time-dependent black hole background, and this is potentially a limitation of the procedure. For the case of globally hyperbolic spacetimes, the reduction can easily be understood from a different and more mathematical perspective [37]. (The condition of global hyperbolicity is not strictly required, though.) This result is summarized in Sect. 3.2. In the conclusion, we comment on several issues such as recovery of the 4D covariance in the present context or the fully 4D-covariant formulation of the quantization. (Progress has been recently made in [23].) In the Appendix, we illustrate the gauge-fixing dependence of a renormalization procedure by taking a system of metric coupled with a scalar. The sole purpose of considering this system is to demonstrate the dependence: we consider only the pure gravity system in the main body.

2 Removal of unphysical degrees of freedom

It will be useful for what follows to recall the quantization procedure in Maxwell’s theory. The vector field has four components to start with but only two of them are physical degrees of freedom. The system has gauge symmetry; it reduces the number of degrees of freedom to three. The time component is non-dynamical, leading to the further reduction of the number of physical fields to two. Let us consider temporal gauge to be specific. It turns out that temporal gauge does not entirely use up the gauge freedom but leaves gauge symmetry associated with the hypersurface of a fixed time [38]. For the perspective of our gravity analysis, what is important is that the non-dynamical time component can be gauge-fixed without using the full bulk symmetry but instead by using measure-zero lower dimensional gauge symmetry. Below we will show that there is an analogous procedure in general relativity.

As in the Maxwell case, it is the close conspiracy between non-dynamism and gauge symmetry that brings complete removal of the unphysical degrees of freedom. In due course of the analysis below, the lapse function $n$ will get fixed to $n = 1$. (This choice is suitable because we are considering expansion of the theory around a flat spacetime to be specific.) This introduces a constraint that corresponds to Hamiltonian constraint of the Hamiltonian formulation. As we will see, the constraint can be solved, and its solution implies, among other things, the effective projection of the 4D dynamics onto the 3D hypersurface.

The gauge-fixings of the measure-zero symmetries will be carried out following the spirit of section 15.4 of [39] in which the first class constraint was eliminated by fixing the corresponding symmetry. In essence, what we do here is fix the gauge and explicitly solve the resulting constraints. Therefore the procedure does not introduce any ghosts at the bulk level. (In [39], quantization in the axial gauge was analyzed, and it was shown that the axial gauge quantization does
not introduce any ghosts. For the actual perturbation computations, a covariant gauge – which does introduce ghosts – was used. See [21] for the ghost-related issues.

2.1 Isolation of unphysical degrees of freedom

Consider the 4D Einstein–Hilbert action (see, e.g. [40] for a review)

\[ S = \int d^4x \sqrt{-g} \ R. \]  

(1)

To illustrate the procedure with a specific example, we separate out the time coordinate and split the coordinates into 

\[ x^\mu \equiv (t, y^a) \]  

(2)

where \( \mu = 0, \ldots, 3 \) and \( a = 1, 2, 3 \). (The resulting 3D system will be Euclidean, thus non-dynamical. For the study of dynamics, one should consider a different setup by separating out, say, the \( y^3 \) coordinate from the rest. We will come back to this issue below.) By parameterizing the 4D metric [6,35]

\[ \hat{g}_{\mu\nu} = \begin{pmatrix} -n^2 + \gamma^{ab} N_a N_b & \gamma^{ab} N_a \\ \gamma^{ab} N_b & \gamma_{ab} \end{pmatrix}, \]  

(3)

the 3 + 1 splitting yields (the boundary terms will not be kept tract of)

\[ S = \int d^4x \ n \sqrt{-\gamma} (R^{(3)} - K^2 + K_{ab} K^{ab}) \]  

(4)

with

\[ K_{ab} = \frac{1}{2n} (\mathcal{L}_\gamma \gamma_{ab} - \nabla_a N_b - \nabla_b N_a), \quad K = \gamma^{ab} K_{ab} \]  

(5)

where \( \mathcal{L}_\gamma \) denotes the Lie derivative along the time coordinate \( t \) and \( \nabla_a \) is the 3D covariant derivative (namely, its \( a \) index denotes the 3D Christoffel symbol. As is well known, the de Donder gauge leaves a residual symmetry (see 6)

In the conventional perturbative analysis, one proceeds and adds the corresponding ghost term in the action. However, the measure-zero gauge fixing will affect, as we will show now, the bulk fixing simply because the bulk fixing (9) contains the non-dynamical fields that get fixed by the 3D gauge-fixing. At the end, only the 3D ghosts will be required after the system is reduced. In terms of the ADM variables, (9) translates to

\[ \hat{g}^{\mu\rho} \Gamma^\mu_{\rho\sigma} = 0. \]  

(9)

where \( \Gamma^\mu_{\rho\sigma} \) denotes the 3D Christoffel symbol. As is well known, the de Donder gauge leaves a residual symmetry (see 6) that this choice will be good for quantization around a flat background. If one has a black hole background in mind with the radial coordinate separated out, one will have to introduce a radial coordinate transformation such that \( g_{rt} \) does not depend on \( r \). As we will further comment later, this is where one may encounter difficulties when trying to apply the procedure to an explicitly time-dependent background because it is not clear whether such a coordinate transformation will always be available without subtleties. This also reveals the background dependence of our procedure: although the methodology is background independent, the detailed steps will depend on the background under consideration. This aspect also makes a connection with precisely what physics the reduced theory describes, a point that we will discuss in Sect. 3 and the conclusion.

7 As a matter of fact, one only needs the \( \mu = 0 \) part of this for the reduction:

\[ \hat{g}^{\rho\sigma} \Gamma^0_{\rho\sigma} = 0. \]  

(8)

The rest will serve as the 3D de Donder gauge in the reduced theory. Improved versions of the reduction can be found in [22,23]. The account in [23] is especially simple.
for example [45]; a nonlinear level discussion can be found in [21] although it is not manifest in (10). It is thus possible, by using the residual diffeomorphism, to set
\[ N_\alpha = 0 \]  
initially on the hypersurface of the fixed time. As mentioned above, this equation can then be taken as valid in the entire bulk due to the non-dynamism of \( N_\alpha \). Substituting \( N_\alpha = 0 \) into (6), which now serves as a constraint, one gets
\[ \nabla^a \left[ \frac{1}{n} (\mathcal{L}_\partial \gamma_{ab} - \gamma_{ab} \mathcal{L}_\partial \gamma) \right] = 0. \]  
(12)

When the covariant derivative acts on objects other than \( n \), it yields zero because the covariant derivative and the Lie derivative commute in the present case [46]. This step requires the advanced-level knowledge in differential geometry of Ch.6 of [46]; for further details see [22,37]. The constraint (12) can be solved and implies
\[ \partial_a n = 0. \]  
(13)

Since \( n \) is non-dynamical, namely, \( n = n(y^\alpha) \), (13) implies that \( n \) should be a constant\(^9\): the \( N_\alpha \) constraint implies \( n = n(t) \), but since \( n \) is non-dynamical, one can set \( n = 1 \)

\[ K_{ab} = \frac{1}{2} \mathcal{L}_\partial y_{ab} \]  
(16)

once (11) and (14) are substituted into (5). The constraint (15) allows one to rewrite the action
\[ S = \int d^4 x \sqrt{-\gamma} \, R^{(3)} \]  
(17)

where the overall factor 2 has been absorbed. The form of the action (17) does not yet imply that the system becomes three-dimensional. This is because one should still consider the constraint (15). We will now argue that (17) with (15) implies projection of the 4D dynamics onto 3D. But before we proceed, a cautionary remark is in order. As is well known, 3D gravity does not have a graviton. In this sense, our reduced system is not the genuine 3D gravity: it has two propagating degrees of freedom, the physical degrees of freedom of the original 4D gravity projected onto the hypersurface. (We will have more to say on this below.)

Let us split the constraint (15) into two parts: the \( R^{(3)} \) term and the \( K \) terms. They vanish separately: to see that \( R^{(3)} \) vanishes we just need to recall that it is an on-shell constraint. In other words, we define the physical states to be annihilated by (15), in which the fields are now viewed as the corresponding operators acting on the Fock space. One may now consider the full field equation for the 3D metric (i.e., the one that follows from (17) and first obtain the mode expansion at the linear level. Then the linear level expression can be substituted into the full field equation, from which one can obtain the full expansion by iteration. (For the definition of the physical states in the actual perturbative computations, one may expand the metric to the linear order and neglect the higher interactions. This is in the same spirit as the comment in one of the footnotes in section 15.7 of [39].) The full expansion of \( y_{ab} \) should yield zero once substituted into the \( R^{(3)} \) part of the constraint (15). With this, only the \( K \) terms remain. As can be seen from the first equation in (10) \( K \) vanishes,
\[ K = 0, \]  
(18)

leaving only
\[ K_{ab}^2 = 0. \]  
(19)

(The same conclusion has been reached in a complementary analysis in [22,23].) Since this is a positive definite metric, this implies
\[ \mathcal{L}_\partial y_{ab} = 0. \]  
(20)

In other words, the physical states of the original 4D system are fully reduced to 3D. The projected system has two physical degrees of freedom: the metric of the hypersurface has six components. The first equation of (10) imposes a constraint and the second equation of (10) imposes the 3D de Donder gauge, effectively removing three degrees of freedom. (We will have to say more on this in the next section.)

In the remainder of this section, we discuss the case of separating out one of the spatial directions, say, \( y^3 \):
\[ x^\mu \equiv (\epsilon^m, y^3) \quad \text{where} \quad m = 0, 1, 2 \]  
(21)

The procedure goes through almost identically, except for several sign changes. For example, the split form of the action now takes the form
\[ S = \int d^4 x \, n \sqrt{-\gamma} (R^{(3)} + K^2 - K_{mn}K^{mn}) \]  
(22)

with
\[ K_{mn} = \frac{1}{2n} (\mathcal{L}_\partial h_{mn} - \nabla_m N_n - \nabla_n N_m), \quad K = h^{mn} K_{mn}. \]  
(23)
The only non-trivial difference compared with the $t$-splitting is the condition that corresponds to (19):

$$K_{mn}^2 = 0. \quad (24)$$

Unlike in (19), the contractions of indices in this equation are done with the 3D metric with $(-+++)$ signature. In general, a Wick rotation must be considered in field theories with non-positive definite metrics, since otherwise the path integral is not well defined. In the gravitational case, the procedure has a subtlety addressed in [28,47,48]. As far as we can see, it is not a subtlety associated with the Wick rotation but with the presence of an unphysical component of the metric. We will take up this issue in the next section. Once the Wick rotation is carried out, Eq. (24) leads to

$$K_{mn} = 0, \quad (25)$$

namely, an on-shell reduction to a dynamical 3D system.

3 Quantization and renormalization

The goal of the previous section was to remove unphysical degrees of freedom. In this section we tackle the very issue of renormalizability of 4D Einstein gravity. Afterwards, we note that the holographic reduction admits a nice mathematical perspective for the cases of globally hyperbolic spacetimes.

For quantization, the $t$-separation case and $y^3$-separation case do not have an essential difference. Both cases are subject to the subtlety observed in [28,47,48], namely, the divergence associated with the trace part of the metric.

3.1 4D quantization through hypersurface

The upshot of the analysis in the previous section is that, upon fixing $n = 1$, $N_a = 0$, by using a lower dimensional diffeomorphism after the bulk fixing, the dynamics of the original 4D system is projected onto “3D”:

$$S = \int d^4x \sqrt{-\gamma} \, R^{(3)}. \quad (26)$$

The quotation mark around 3D above is written because of the presence of the 4D integration and 4D coordinate dependence of the off-shell metric $\gamma_{ab}$. We now show that the integral over the 4D spacetime is effectively reduced to 3D. For the discussion in this section, we consider separating out the spatial coordinate $y^3$. This action should be supplemented by the nonlinear form of the 3D de Donder gauge,

$$\gamma_{mn} \Gamma_{mn}^p = 0. \quad (27)$$

With the genuine 3D gravity, one can use the residual symmetry after imposing (27) to gauge away three non-dynamical components, thereby arriving at the well-known absence of propagating degrees of freedom. In the current case, however, it is impossible to do the same simply because the metric in (26) still has the fourth coordinate dependence off-shell. The residual symmetry after (27) should be the symmetry within the hypersurface. This can be seen by considering the residual symmetry condition stated in [45]. Namely, it is a partial differential equation of the gauge parameter $\epsilon_m$ on the hypersurface and therefore cannot be used to gauge away additional components. Alternatively, one can consider the entire procedure of the gauge-fixings in the path-integral setup. The result (26) with (27) has been obtained as a result of this complete gauge-fixing by implicitly following the steps given in section 15.4 of [39], and therefore one should not further gauge-fix (26) other than by (27).

Remarks on precisely what physics the reduced system (26) with (27) is supposed to describe are in order. The focus of the present work is perturbative quantization around a flat background, exactly the goal that we set out for. Although the procedure should be applicable to a Schwarzschild background, it is not clear whether the procedure can be used to study more complex backgrounds. Related to this, there is an issue of non-perturbative corrections. We will not pursue these issues here.

For perturbation, one can use the usual linear form of the 3D de Donder gauge (27). The 3D gauge-fixing and ghost terms will have to be introduced, and with that one can carry out the BRST quantization. Here we focus on the graviton part of the computation. Since the procedure maintains only the 3D covariance, one should worry about compatibility with the (yet unavailable) fully 4D covariant approach. We will further remark on this in the conclusion.

The perturbation series with (26) can be set up as follows. (The following analysis has been substantially expanded in [211].) For the perturbative analysis, the usual linearized version of (27) can be imposed. The propagator is then given by

$$\langle q_{mn}(x_1)q_{pq}(x_2) \rangle = \frac{1}{2} \left( \eta_{mp} \eta_{nq} + \eta_{mq} \eta_{np} - 2 \frac{3}{m} \eta_{mn} \eta_{pq} \right)$$

$$\times \int \frac{d^4k}{(2\pi)^4} \frac{e^{i\langle x_1-x_2 \rangle}}{ik^2} \quad (28)$$

where $k$ is the 3-vector part of the 4-momentum $k\mu$.

The field $q_{ij}$ is the traceless part of the fluctuation around a flat metric (use of the traceless metric $\eta_{mn}$ is not essential for the reduction),

$$\gamma_{mn} \equiv \eta_{mn} + l_{mn}, \quad q_{mn} \equiv l_{mn} - \frac{1}{3} \eta_{mn} l, \quad l \equiv \eta^m^nl_{mn}. \quad (29)$$

This form of the propagator implies that one can carry out the renormalization program just as in the regular 4D case. After performing various momentum integrals, thereby removing momentum delta functions, one will be left with integration over the loop momenta in which the propagators take
the three-dimensional forms. The 3D integration can be carried out by the standard techniques. The integration over the reduced direction can be carried out by taking, e.g., a momentum cut-off regularization. In general, the momentum cut-off regularization does not respect gauge symmetry. However, adopting it for the fourth direction does not pose a problem here: dimensional regularization can be used for the 3D integrals. The divergence factor arising from the $\chi^0$ integration can be absorbed by rescaling the fields (i.e., wave-function renormalization). In more detail, let us introduce a rescaling of $w$ by a dimensional parameter $L$ (we have defined $w \equiv y^3$):

$$u = \frac{w}{L}$$

such that $u$ is dimensionless. With this rescaling, one can use a dimensionless momentum cut-off for the $u$ direction. The rescaling will lead to renormalization (i.e., rescaling by $L$) of Newton’s constant, which has been suppressed, and one achieves an effective reduction to 3D:

$$S = \int d^3z \sqrt{-\gamma} R^{(3)}.$$  

(31)

Since the 4D system (4) with which we have started has been reduced to (31) with the constraint (27), the former would be renormalizable if the latter is.

For renormalizability of (31) with the constraint (27), the result of [20] can mostly be borrowed with several cautions to which we will now turn. It was shown in [20] that 3D pure gravity is renormalizable even though it is power-counting non-renormalizable. To use the result of [20] in the present case, one should first understand that the theory that was considered in [20] is the genuine 3D Einstein gravity that has no propagating degrees of freedom. In the case of the 3D pure gravity, the author relied on certain kinematic features of 3D gravity instead of explicit computation in order to deduce renormalizability. However, the author considered propagating gravitons as can be seen in the gravity–matter coupling case, where explicit computation was carried out. The propagating degrees of freedom must be the non-dynamical ones; this is in line with the common practice in the literature before and after [20], in which non-dynamical fields are allowed to propagate in loops. In the current setup, those non-dynamical fields are dynamical in the 4D standpoint, and they are justified to be present in the loops.

The second cautionary step one should take is the fact that the result of [20] was obtained up to the observation in [28], which addressed the issue raised in [47,48]. The constraint shows that the trace part of the original 4D metric is non-dynamical. Upon the fixing of $n = 1, N_a = 0$, the constraint implies that the trace part of the 3D metric is non-dynamical. It was shown in [28] that the 4D trace part is non-dynamical in the 4D sense. Since our reduced system should represent the 4D system, the trace part of the metric (3D one after the fixing) should be excluded from the dynamics.

3.2 Perspective of mathematics of foliation

The analysis of the previous section does not need any assumption on the causality of the starting 4D manifold. It is possible to have a complementary understanding of the finding of the previous section through abstract foliation theory. We will assume for simplicity, although it is not necessary, that the background is globally hyperbolic. This section is a brief summary of [37], in which more details can be found. (See [22] for further development of the idea along these lines.)

A globally hyperbolic spacetime – which covers most of the cosmologically interesting spacetimes – admits a codimension-1 foliation through a family of hypersurfaces. The condition (13) obtained by the shift vector constraint can be written precisely as the condition for the foliation to be Riemannian:

$$\mathcal{L}_\partial n = 0.$$  

(32)

In mathematics, it is known that a codimension-3 Riemannian foliation admits a “dual” totally geodesic foliation. The duality involved is a mathematical one and operates between two different foliations: the Riemannian and totally geodesic.

Now the starting manifold can be viewed as a principal bundle of 1D abelian fibration over the 3D base through the totally geodesic foliation. Then one may associate the action of the 1D group fibration with the gauge symmetry; modding out the gauge symmetry will correspond to casting the 4D manifold into its 3D base. The presence of the totally geodesic foliation has an enlightening implication for the apparent gauge-choice sensitivity, as we will discuss in the conclusion.

4 Conclusion

A metric in four dimensions has ten components, four of which are non-dynamical, and another four components are associated with 4D gauge symmetry. In the conventional covariant renormalization program, only the modes associated with 4D gauge symmetry are removed by a gauge-fixing such as the de Donder gauge; the 4D Einstein gravity is power-counting unrenormalizable and indeed turned out to be unrenormalizable. Even though it is only the measure-zero diffeomorphism that remains unixed in the conventional covariant approach, the non-dynamical fields circulate the loop acting as bulk fields, thereby apparently ruining the renormalizability.

In this work, we have contemplated the possibility that additional removal of the non-dynamic fields may lead to
renormalizability of the ADM formulation of 4D Einstein gravity. We have shown that there exists a way to remove all (or most) of the unphysical degrees of freedom and at the same time to set up a convenient stage for examining renormalizability. (As stated in one of the footnotes in the Introduction, one may keep the four of the eight unphysical fields at the off-shell level and carry out renormalization [23].) After removal of the unphysical fields, the 4D dynamics admits the effective description in terms of the 3D language: 4D renormalizability is achieved based on the 3D renormalizability.

Compared with AdS/CFT, it is notable that the boundary theory in the present case is not a gauge theory but takes the form of a lower dimensional gravity theory. We do not believe that this has anything to do the hypersurface foliation approach, because the approach (combined with manual dimensional reduction) has led to a boundary gauge theory in the IIB supergravity setup [31,49].

There is an interesting implication of the present work for the previous works of [32–34], wherein explicit dimensional reduction was carried out in order to avoid the quantization issues. The analyses there were for certain sub-sectors (i.e., the sectors associated with the hypersurfaces selected) of the whole moduli space. The result of the present work implies that the analyses in those works are much more complete than originally believed.

We end with several directions for future research.

One of the subtle points is that the whole procedure of reduction seems to heavily depend on the elaborate gauge-fixings. (It would thus be worthwhile to explore whether there exists a bulk gauge other than the de Donder gauge that would lead to the same conclusion. Perhaps an easier task that will nevertheless shed light on this issue is to study how the form of the de Donder gauge should be modified in the case of a Schwarzschild background.) Part of the reason for the gauge sensitivity should be that the ratio of the physical versus unphysical fields is much higher than in YM theory. It seems, therefore, natural to a certain extent that certain things depend on the gauge choice more sensitively than in YM theory. The mathematical approach in [22,37] – which takes advantage of the presence of the dual totally geodesic foliation – strongly suggests that the reduction should be a rather general phenomenon (associated with relatively simple backgrounds). This is because in that approach the only thing one needs is the very natural gauge-fixing $\tilde{\xi}_{\mu\nu} = 0$ (which of course is $N_a = 0$). Judging from this and other pieces of evidence, the gauge sensitivity of the present approach should not be such a serious problem.

The present work was carried out for fluctuations around a flat spacetime. For example, it would definitely be worthwhile to carry out a similar analysis for a Schwarzschild black hole background. Although the methodology would be basically the same, the detailed steps would have to be modified: the present approach has a moderate background dependence. To this end, it would be better to consider separating out the radial coordinate than the time coordinate. We believe that this is a task that can be carried out with only relatively minor changes in the present analysis. With this achieved, one would be in a good position to tackle the Firewall [50] (see [57] for an earlier related work) and related issues, which was the main motivation for the present work.

At some point, one must face the difficult issue of non-perturbative contributions to the path integral. There are two kinds of non-perturbative contributions. The first kind is the contribution coming from summing over all orders of the coupling constant in the perturbation theory. A 3D version of the techniques of the asymptotically safe gravity [51–55] may be useful along this line. There will be several important issues such as the Gribov problem (see, e.g. [56] and references therein for a recent account in the context of gravity) that would have to be addressed eventually. The other kind is the contribution analogous to the instanton contribution in YM theory. It is not clear whether the present procedure is adequate for that, given the limitation mentioned in the main text. (Unlike the instantons in YM theory, much less is known about the gravitational analogs anyway.)

One of the utmost important directions of research concerns the 4D covariance, since the present approach maintains only the 3D covariance. There exist several different ways of ensuring equivalence between a non-covariant approach and a covariant approach. (See [23] for recent progress.) Let us recall by a few examples how the equivalence is accomplished in quantum field theory and string theory. One may attempt to construct a 4D-covariant operator formulation without using any constraint directly in the action. In other words, one would attempt 4D-covariant removal of all or most of the unphysical degrees of freedom. Perhaps it could be done along the line of the gauge invariant quantization [39]. Some of the observations of [24] may be useful. While these programs could succeed, an easier approach will presumably be along the line of the gauge invariant quantization analysis in section 9.6 of [26]. It was shown that the Coulomb gauge canonical quantization of electrodynamics can be related to the Lorenz gauge path integral. When both the non-covariant and the covariant formulations were independently available, one could establish the equivalence by simply comparing the amplitudes of the physical states (and this is the way the equivalence between the light-cone quantization and covariant quantization was achieved in string theory). Computations of various Feynman diagrams must be preceded to this end. Both of these approaches were discussed in [21], at least to some extent.

The observation on the renormalizability in this work is, to some extent, up to the issue of reduced space quantization vs. Dirac quantization. In general, those two approaches do not lead to the same physics [58,59]; see also the refer-
ences therein. This was demonstrated by taking an example in which the first constraint was not associated with the gauge symmetry. It was shown that, in general, both quantization methods lead to different but consistent theories. We are not aware of any work in which a similar conclusion was drawn for a theory with a gauge symmetry inducing a first class constraint. We believe that those two approaches are more likely than not to lead to the same results at the end as was the case in string quantization: light-cone, old covariant, and modern BRST quantizations all led to the same results.

Finally, we have recently become aware of the work of York [18], Moncrief [60], and Fischer-Moncrief [17] wherein certain Hamiltonian reductions were carried out on a certain class of 4D manifolds. It was a pleasant surprise to discover that the reduced Hamiltonians were given by the volumes of the hypersurfaces. Given that the action for a particle is given by its length and that of a string by its surface, the appearance of a volume for the hypersurface seems natural, except that the volume appears as the Hamiltonian instead of the Lagrangian. Presumably this is due the fact that the reduced direction is the time-direction, so that the hypersurface is Euclidean. In our case, it was the 3D Einstein action that has emerged. Perhaps, the 3D hypersurfaces admit dual descriptions, one through the volume and the other through the Einstein action. It would be interesting to understand the potential relation between the work of [17,18,60] and the case in string quantization: light-cone, old covariant, and modern BRST quantizations all led to the same results.

We will report on progress made with some of these issues in the near future.10

Acknowledgments I acknowledge the hospitality of S.-J. Sin and B.-H. Lee at Hanyang University and CQUeST during my stay. I thank the members of Hanyang University and KIAS for their interest and criticisms. This work benefited from the discussions I had with them. I thank V. Moncrief, G. 't Hooft, and R. Woodard for their communications. I also thank H. Lee, E. Poisson, S. Weinberg for reading the main idea of the work and making encouraging comments.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Funded by SCOAP3.

Appendix: Gauge-fixing dependence of renormalization

Let us demonstrate the complications caused by the presence of the unphysical modes in carrying out a renormalization program. We show that the renormalization procedure substantially depends on whether or not the undynamical metric component, the trace, is fixed.11 As we will see, whether or not one gauges away the trace part makes the procedures drastically different. What we observed in the main body was that the difference can be so drastic as to render the 4D pure Einstein action renormalizable.

The gauge-fixing dependence of a renormalization procedure can easily be illustrated by a coupled system of a metric and scalar.12 Consider the 4D Einstein action coupled with a scalar,

\[
S = \int d^4x \sqrt{-g} \left( \mathcal{R} - \frac{1}{2} \dot{g}^{\mu
u} \partial_\mu \phi \partial_\nu \phi \right),
\]

and the weak-field expansion around the Minkowski metric,

\[
\dot{g}_{\mu\nu} = \eta_{\mu\nu} + g_{\mu\nu}.
\]

Let us impose the de Donder gauge by adding the following gauge-fixing term to the action:

\[
\mathcal{L}_{g.f.} = -\frac{1}{2} \left( \partial_\rho g^{\mu\rho} - \frac{1}{2} \dot{g}^{\mu\nu} g \right)^2
\]

where \( g \equiv \eta^{\alpha\sigma} g_{\rho\sigma} \). Multiplying an overall factor 2, the action is given by

\[
\mathcal{L} = \left( -\frac{1}{2} \partial_\alpha g_{\mu\nu} \partial^\alpha g^{\mu\nu} + \frac{1}{4} \partial_\alpha g^{\mu\nu} \partial^\alpha g_{\mu\nu} - \partial^\nu \bar{C}^\mu \partial_\nu C^\mu \\
\frac{1}{2} g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial^\nu g_{\mu\rho} + 2 g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial_\nu g_{\rho\gamma} + \frac{1}{2} g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial_\nu g_{\rho\gamma} - g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial_\nu g_{\rho\gamma} \\
+ \frac{1}{4} g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial_\nu g_{\rho\gamma} - g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial_\nu g_{\rho\gamma} + \frac{1}{2} g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial_\nu g_{\rho\gamma} - g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial_\nu g_{\rho\gamma} \\
- \frac{1}{2} g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial^\nu g_{\rho\gamma} + g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial^\nu g_{\rho\gamma} + \frac{1}{4} g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial^\nu g_{\rho\gamma} \\
- g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial^\nu g_{\rho\gamma} + g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial^\nu g_{\rho\gamma} - g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial^\nu g_{\rho\gamma} - g^{\alpha\beta} \partial_\alpha g_{\nu\beta} \partial^\nu g_{\rho\gamma} \\
- \partial^\nu \bar{C}^\mu \partial_\nu \bar{C}^\nu g_{\rho\gamma} - \partial^\nu \bar{C}^\mu \partial_\nu C^\nu g_{\rho\gamma} - \partial^\nu \bar{C}^\mu C^\nu g_{\rho\gamma} + \partial_\mu \bar{C}^\nu \partial^\nu g_{\rho\gamma} + \frac{1}{2} \partial_\mu \bar{C}^\nu C^\nu g_{\rho\gamma} \\
- \frac{1}{2} \left[ \eta^{\mu\nu} - g^{\mu\nu} + \frac{1}{2} \dot{g}^{\mu\nu} g + g^{\mu\nu} g \right] \partial_\mu \phi \partial_\nu \phi \right)
\]

where \( C \) is the ghost field that corresponds to the gauge-fixing (A.3). The trace part is not a dynamical degree of freedom

11 Of course, keeping the trace part makes the path integral ill-defined as observed in [28,47,48]; for the sake of the discussion in this section, we set this observation aside and formally proceed for now. (This problem has been addressed recently in [23].)

12 The reduction to 3D that we observed in the main body was for the pure Einstein gravity; we do not claim the same for the coupled system that we consider in this section.
and the traceless condition can explicitly be enforced by using the traceless propagator as follows. Let us explicitly introduce the traceless metric

\[ h_{\mu\nu} \equiv g_{\mu\nu} - \frac{1}{4} \eta_{\mu\nu} g \]  

(A.5)

and choose

\[ L_{\text{kin}} = -\frac{1}{2} \partial_\mu h_{\mu\nu} \partial^\nu h^{\mu\nu} \]  

(A.6)

as the kinetic term. The propagator then is traceless:

\[
\langle h_{\mu\nu}(x) h_{\rho\sigma}(y) \rangle = \frac{1}{\sqrt{2}} \left( \eta_{\mu\rho} \eta_{\nu\sigma} + \eta_{\mu\sigma} \eta_{\nu\rho} - \frac{1}{2} \eta_{\mu\nu} \eta_{\rho\sigma} \right) 
\times \int \frac{d^4k}{(2\pi)^4} e^{ik(x-y)}. 
\]

(A.7)

It follows from this that

\[
\langle h_{\mu\nu}(x) h(y) \rangle = 0 = \langle h(x) h_{\rho\sigma}(y) \rangle; 
\]

(A.8)

any vertex term that contains a factor \( h \) will lead to zero in a correlator computation. All the quadratic terms containing a \( h \)-factor will not contribute to any correlator, and thereby be effectively dropped; the action becomes substantially simplified:

\[
L = -\frac{1}{2} \partial_\mu h_{\mu\nu} \partial^\alpha h^{\mu\nu} - \partial_\nu \bar{C}_\mu \partial^\nu C^\mu \\
- \frac{1}{2} \left[ h^{\alpha\beta} \partial_\alpha h_{\beta\gamma} \partial_\delta h_{\gamma\delta} + 2h^{\alpha\beta} \partial_\alpha h^{\gamma\delta} \partial_\delta h_{\gamma\beta} + h^{\alpha\beta} \partial_\alpha h_{\beta\gamma} \partial_\gamma h_{\beta\gamma} + h^{\alpha\beta} \partial_\alpha h_{\beta\gamma} h^{\gamma\delta} \partial_\delta h_{\beta\gamma} - h^{\alpha\beta} \partial_\alpha h_{\beta\gamma} \partial_\gamma h_{\beta\gamma} - h^{\alpha\beta} \partial_\alpha h_{\beta\gamma} \partial_\gamma h_{\beta\gamma} + h^{\alpha\beta} \partial_\alpha h_{\beta\gamma} \partial_\gamma h_{\beta\gamma} \right] \\
- \partial^\nu \bar{C}_\mu \partial_\nu C^\mu h_{\nu\nu} - \partial^\nu \bar{C}_\mu \partial_\nu C^\mu ; h_{\mu\rho} - \partial^\nu \bar{C}_\mu \partial_\nu C^\rho h_{\nu\nu} \\
+ \partial_\mu \bar{C}_\nu \partial^\nu C^\rho h_{\nu\nu} - \frac{1}{2} \left[ \eta_{\mu\nu} - h_{\mu\nu} \right] \\
+ h^{\mu\nu} h^-_{\rho\sigma} - \frac{1}{4} h^{\mu\nu} h_{\rho\sigma} h^{\rho\sigma} \right] \partial_\mu \phi \partial_\nu \phi + \cdots 
\]

(A.9)

One can check that \( h = 0 \) is a legitimate gauge. At one loop (see Fig. 2a), \((\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4))\), which is a gauge singlet, is the same whether one uses (A.4) or (A.9).

The see the gauge dependence, let us consider the scalar–scalar–metric amplitudes drawn in Fig. 2b, c. One can easily see that the tree-level correlators \( \langle \phi \phi g_{\mu\nu} \rangle \) with the action (A.4) and \( \langle \phi \phi h_{\mu\nu} \rangle \) with the action (A.9) lead to different results. The same is true for the one-loop correlators. The computations are simplified in the latter case because the number of relevant vertices is smaller. Although we have not explicitly carried out, the diagram in Fig. 3 clearly shows the drastic simplification in the computation with \( h_{\mu\nu} \) as can be seen by counting the number of relevant vertices: 13 in (A.4) vs. 5 in (A.9).

Carrying out a normalization procedure for \( g_{\mu\nu} \) means that we consider the amplitudes in which the unphysical component \( g \) appears, e.g., as an external line. It seems highly likely that the exclusion of \( g \) will reduce the number of counter terms. We have confirmed this expectation in the main text in a very dramatic way.

Fig. 3 Diagram with cubic metric coupling

References

1. S. Carlip, Quantum gravity: a progress report, Rept. Prog. Phys. 64, 885 (2001) arXiv:gr-qc/0108040
2. G. ’t Hooft, M.J.G. Veltman, One loop divergencies in the theory of gravitation. Annales Poincare Phys. Theor. A 20, 69 (1974)
3. S. Deser, P. van Nieuwenhuizen, Nonrenormalizability of the quantized Dirac–Einstein system. Phys. Rev. D 10, 411 (1974)
4. S. Deser, P. van Nieuwenhuizen, One loop divergences of quantized Einstein–Maxwell fields. Phys. Rev. D 10, 401 (1974)
5. M.H. Goroff, A. Sagnotti, The ultraviolet behavior of einstein gravity. Nucl. Phys. B 266, 709 (1986)
6. R.L. Arnowitt, S. Deser, C.W. Misner, The Dynamics of general relativity, Gen. Rel. Grav. 40, 1997 (2008). arXiv:gr-qc/0405109
7. E. Gourgoulhon, 3+1 formalism and bases of numerical relativity. arXiv:gr-qc/0703035
8. P.A.M. Dirac, Fixation of coordinates in the Hamiltonian theory of gravitation. Phys. Rev. 114, 924 (1959)
9. P.G. Bergmann, A. Komar, The coordinate group symmetries of general relativity. Int. J. Theor. Phys. 5, 15 (1972)
10. C.J. Isham, K.V. Kuchar, Representations of space–time diffeomorphisms. 1. Canonical parametrized field theories. Ann. Phys. 164, 288 (1985)
11. C.J. Isham, K.V. Kuchar, Representations of space–time diffeomorphisms. 2. Canonical geometrodynamics. Ann. Phys. 164, 316 (1985)
12. A. Ashtekar, New variables for classical and quantum gravity. Phys. Rev. Lett. 57, 2244 (1986)
13. C. Rovelli, Loop quantum gravity. Living Rev. Rel. 1, 1 (1998). arXiv:gr-qc/9710008
14. T. Thiemann, Modern Canonical Quantum General Relativity (Cambridge Univ. Pr., Cambridge, 2007). arXiv:gr-qc/0101003
15. K.S. Stelle, Renormalization of higher derivative quantum gravity. Phys. Rev. D 16, 953 (1977)
16. I. Antoniadis, E.T. Tomboulis, Gauge invariance and unitarity in higher derivative quantum gravity. Phys. Rev. D 33, 2756 (1986)
17. A.E. Fischer, V. Moncrief, Hamiltonian reduction of Einstein’s equations of general relativity. Nucl. Phys. Proc. Suppl. 57, 142 (1997)
18. J.W. York Jr, Role of conformal three geometry in the dynamics of gravitation. Phys. Rev. Lett. 28, 1082 (1972)
19. E. Witten, (2+1)-dimensional gravity as an exactly soluble system. Nucl. Phys. B 311, 46 (1988)
