

# The Asymptotic Safety Scenario in Quantum Gravity

Max Niedermaier

Laboratoire de Mathématiques et Physique Théorique  
CNRS/UMR 6083  
Université de Tours  
Parc de Grandmont  
37200 Tours, France  
email: Max.Niedermaier@lmpt.univ-tours.fr

Martin Reuter

Institute of Physics  
University of Mainz  
Staudingerweg 7  
55099 Mainz, Germany  
email: reuter@thep.physik.uni-mainz.de

Accepted on 23 October 2006

Published on 1 December 2006

*Living Reviews in Relativity*

Published by the  
Max Planck Institute for Gravitational Physics  
(Albert Einstein Institute)  
Am Mühlenberg 1, 14424 Golm, Germany  
ISSN 1433-8351

## Abstract

The asymptotic safety scenario in quantum gravity is reviewed, according to which a renormalizable quantum theory of the gravitational field is feasible which reconciles asymptotically safe couplings with unitarity. The evidence from symmetry truncations and from the truncated flow of the effective average action is presented in detail. A dimensional reduction phenomenon for the residual interactions in the extreme ultraviolet links both results. For practical reasons the background effective action is used as the central object in the quantum theory. In terms of it criteria for a continuum limit are formulated and the notion of a background geometry self-consistently determined by the quantum dynamics is presented. Self-contained appendices provide prerequisites on the background effective action, the effective average action, and their respective renormalization flows.

## How to cite this article

Owing to the fact that a *Living Reviews* article can evolve over time, we recommend to cite the article as follows:

Max Niedermaier and Martin Reuter,  
“The Asymptotic Safety Scenario in Quantum Gravity”,  
*Living Rev. Relativity*, **9**, (2006), 5. [Online Article]: cited [<date>],  
<http://www.livingreviews.org/lrr-2006-5>

The date given as <date> then uniquely identifies the version of the article you are referring to.

## Article Revisions

*Living Reviews* supports two different ways to keep its articles up-to-date:

**Fast-track revision** A fast-track revision provides the author with the opportunity to add short notices of current research results, trends and developments, or important publications to the article. A fast-track revision is refereed by the responsible subject editor. If an article has undergone a fast-track revision, a summary of changes will be listed here.

**Major update** A major update will include substantial changes and additions and is subject to full external refereeing. It is published with a new publication number.

For detailed documentation of an article’s evolution, please refer always to the history document of the article’s online version at <http://www.livingreviews.org/lrr-2006-5>.

# Contents

<b>1</b>	<b>Introduction and Survey</b>	<b>5</b>
1.1	Survey of the scenario	5
1.2	Evidence for asymptotic safety	14
1.3	Some working definitions	21
1.4	Relation to other approaches	25
1.4.1	Dynamical triangulations	25
1.4.2	Loop quantum gravity	26
1.4.3	String theory	28
1.5	Discussion of possible objections	29
<b>2</b>	<b>Renormalizing the Non-Renormalizable</b>	<b>33</b>
2.1	Perturbation theory and continuum limit	33
2.2	Functional flow equations and UV renormalization	37
2.3	Towards Quantum Gravidynamics	42
2.3.1	The role of Newton's constant	42
2.3.2	Perturbation theory and higher derivative theories	44
2.3.3	Kinematical measure	48
2.3.4	Effective action and states	51
2.3.5	Towards physical quantities	52
2.4	Dimensional reduction of residual interactions in UV	53
2.4.1	Scaling of fixed point action	53
2.4.2	Anomalous dimension at non-Gaussian fixed point	55
2.4.3	Strict renormalizability and $1/p^4$ propagators	55
2.4.4	Spectral dimension and scaling of fixed point action	56
<b>3</b>	<b>Asymptotic Safety from Dimensional Reduction</b>	<b>59</b>
3.1	2 + 2 truncation of Einstein gravity + matter	59
3.1.1	Gravity theories	59
3.1.2	2-Killing vector reduction	60
3.1.3	Hamiltonian formulation	64
3.1.4	Lapse and shift in 2D gravity theories	65
3.1.5	Symmetries and currents	67
3.2	Collinear gravitons, Dirac quantization, and conformal factor	68
3.2.1	Collinear gravitons	68
3.2.2	Dirac versus covariant quantization	71
3.2.3	Conformal factor	74
3.3	Tamed non-renormalizability	75
3.4	Non-Gaussian fixed point and asymptotic safety	83
<b>4</b>	<b>Asymptotic Safety from the Effective Average Action</b>	<b>93</b>
4.1	The effective average action for gravity and its FRGE	93
4.1.1	Properties of the effective average action	96
4.2	Geometries at different resolution scales	100
4.3	Truncated flow equations	102
4.4	Einstein–Hilbert and $R^2$ truncations	109
4.4.1	Phase portrait of the Einstein–Hilbert truncation	109
4.4.2	Evidence for asymptotic safety – Survey	110
4.4.3	Structure of the unstable manifold	111
4.4.4	Robustness of qualitative features	113

4.4.5	Comments on the full FRGE dynamics . . . . .	116
<b>5</b>	<b>Conclusions</b>	<b>119</b>
<b>6</b>	<b>Acknowledgements</b>	<b>120</b>
<b>A</b>	<b>Reminder on Kadanoff–Wilson Renormalization</b>	<b>121</b>
<b>B</b>	<b>The Background Effective Action</b>	<b>130</b>
B.1	Standard effective action and its perturbative construction . . . . .	130
B.2	Survey of background field formalisms . . . . .	133
B.2.1	Linear split, non-gauge theories . . . . .	134
B.2.2	Geodesic split, non-gauge theories . . . . .	135
B.2.3	Linear split, gauge theories . . . . .	139
B.2.4	Geodesic split, gauge theories . . . . .	141
B.3	Renormalization of Riemannian sigma-models . . . . .	141
<b>C</b>	<b>The Effective Average Action</b>	<b>149</b>
C.1	Definition and basic properties . . . . .	149
C.2	Flow equation . . . . .	151
C.3	Decoupling properties . . . . .	155
	<b>References</b>	<b>157</b>

# 1 Introduction and Survey

The search for a physically viable theory of quantized gravitation is ongoing; in part because the physics it ought to describe is unknown, and in part because different approaches may not ‘approach’ the same physics. The most prominent contenders are string theory and loop quantum gravity, with ample literature available on either sides. For book-sized expositions see for example [97, 177, 112, 199]. The present report and [157] describe a circle of ideas which differ in several important ways from these approaches.

## 1.1 Survey of the scenario

First, the gravitational field itself is taken seriously as the prime carrier of the relevant classical and quantum degrees of freedom. Second, a physics premise (“antiscreening”) is made about the self-interaction of these quantum degrees of freedom in the ultraviolet. Third, the effective diminution of the relevant degrees of freedom in the ultraviolet (on which morally speaking all approaches agree) is interpreted as universality in the statistical physics sense in the vicinity of an ultraviolet renormalization group fixed point. The resulting picture of microscopic geometry is fractal-like with a local dimensionality of two.

The concrete implementation of these ideas has begun only recently and led to a number of surprising results to be reviewed here. Part of the physics intuition, on the other hand, dates back to an 1979 article by Weinberg [227] (see also [94]). Motivated by the analogy to the asymptotic freedom property of non-Abelian gauge theories, the term “asymptotic safety” was suggested in [227], indicating that physical quantities are “safe” from divergencies as the cutoff is removed. Following this suggestion we shall refer to the above circle of ideas as the “asymptotic safety scenario” for quantum gravity. For convenient orientation we display the main features of the asymptotic safety scenario in an overview:

- Relate micro- and macro-physics of the gravitational field through a renormalization flow.
- As the basic physics premise stipulate that the physical degrees of freedom in the extreme ultraviolet interact predominantly antiscreening.
- Based on this premise benign renormalization properties in the ultraviolet are plausible. The resulting “Quantum Gravidynamics” can then be viewed as a peculiar quasi-renormalizable field theory based on a non-Gaussian fixed point.
- In the extreme ultraviolet the residual interactions appear two-dimensional.

This is a setting which places quantum gravity within the framework of known physics principles. It is not presupposed that continuous fields or distributions on a four dimensional manifold are necessarily the most adequate description in the extreme ultraviolet. However, since these evidently provide the correct dynamical degrees of freedom at ‘low’ (sub TeV) energies, a research strategy which focuses on the ‘backtracing’ of the origin of these dynamical degrees of freedom via renormalization group ideas seems most appropriate.

This amounts to a strategy centered around a functional integral picture, which was indeed the strategy adopted early on [144, 78], but which is now mostly abandoned. A functional integral over geometries of course has to differ in several crucial ways from one for fields on a fixed geometry. This led to the development of several formulations (canonical, covariant [64, 65, 66], proper time [212, 213], and covariant Euclidean [104, 92]). As is well-known the functional integral picture is also beset by severe technical problems [210, 63]. Nevertheless this should not distract attention from the fact that a functional integral picture has a physics content which differs from the physics content of other approaches. For want of a better formulation we shall refer to this fact by saying

that a functional integral picture “takes the degrees of freedom of the gravitational field seriously also in the quantum regime”.

Let us briefly elaborate on that. Arguably the cleanest intuition to ‘what quantizing gravity might mean’ comes from the functional integral picture. Transition or scattering amplitudes for nongravitational processes should be affected not only by one geometry solving the gravitational field equations, but by a ‘weighted superposition’ of ‘nearby possible’ off-shell geometries. The rationale behind this intuition is that all known (microscopic) matter is quantized that way, and using an off-shell matter configuration as the source of the Einstein field equations is in general inconsistent, unless the geometry is likewise off-shell. Moreover, relativistic quantum field theory suggests that the matter-geometry coupling is effected not only through averaged or large scale properties of matter. For example nonvanishing connected correlators of a matter energy momentum tensor should be a legitimate source of gravitational radiation as well (see [81]). Of course this does not tell in which sense the geometry is off-shell, nor which class of possible geometries ought to be considered and be weighed with respect to which measure. Rapid decoherence, a counterpart of spontaneous symmetry breaking, and other unknown mechanisms may in addition mask the effects of the superposition principle. Nevertheless the argument suggests that the degrees of freedom of the gravitational field should be taken seriously also in the quantum regime, roughly along the lines of a functional integral.

Doing so one has to face the before mentioned enormous difficulties. Nevertheless facing these problems and maintaining the credible physics premise of a functional integral picture is, in our view, more appropriate than evading the problems in exchange for a less credible physics premise. Of course in the absence of empirical guidance the ‘true’ physics of quantum gravity is unknown; so for the time being it will be important to try to isolate differences in the physics content of the various approaches. By physics content we mean here qualitative or quantitative results for the values of “quantum gravity corrections” to *generic* physical quantities in the approach considered. Generic physical quantities should be such that they in principle capture the entire invariant content of a theory. In a conventional field theory S-matrix elements by and large have this property, in canonical general relativity Dirac observables play this role [9, 219, 70]. In quantum gravity, in contrast, no agreement has been reached on the nature of such generic physical quantities.

Quantum gravity research strongly draws on concepts and techniques from other areas of theoretical physics. As these concepts and techniques evolve they are routinely applied to quantum gravity. In the case of the functional integral picture the transferral was in the past often dismissed as eventually inappropriate. As the concepts and techniques evolved further, the reasons for the original dismissal may have become obsolete but the negative opinion remained. We share the viewpoint expressed by Wilczek in [230]: “Whether the next big step will require a sharp break from the principles of quantum field theory, or, like the previous ones, a better appreciation of its potentialities, remains to be seen”. As a first (small) step one can try to reassess the prospects of a functional integral picture for the description of the quantized gravitational field, which is what we set out to do here. We try to center the discussion around the above main ideas, and, for short, call a quantum theory of gravity based on them **Quantum Gravidynamics**. For the remainder of Section 1.1 we now discuss a number of key issues that arise.

In any functional integral picture one has to face the crucial **renormalizability problem**. Throughout we shall be concerned exclusively with (non-)renormalizability in the ultraviolet. The perspective on the nature of the impasse entailed by the perturbative non-renormalizability of the Einstein–Hilbert action (see Bern [30] for a recent review), however, has changed significantly since the time it was discovered by ’t Hooft and Veltmann [210]. First, the effective field theory framework applied to quantum gravity (see [50] for a recent review) provides unambiguous answers for ‘low energy’ quantities despite the perturbative non-renormalizability of the ‘fundamental’ action. The role of an a-priori microscopic action is moreover strongly deemphasized when a Kadanoff–Wilson view on renormalization is adopted. We shall give a quick reminder on this

framework in Appendix A. Applied to gravity it means that the Einstein–Hilbert action should *not* be considered as the microscopic (high energy) action, rather the (nonperturbatively defined) renormalization flow itself will dictate, to a certain extent, which microscopic action to use and whether or not there is a useful description of the extreme ultraviolet regime in terms of ‘fundamental’ (perhaps non-metric) degrees of freedom. The extent to which this is true hinges on the existence of a fixed point with a renormalized trajectory emanating from it. The fixed point guarantees universality in the statistical physics sense. If there is a fixed point, any action on a renormalized trajectory describes identically the same physics on all energy scales lower than the one where it is defined. Following the trajectory back (almost) into the fixed point one can in principle extract unambiguous answers for physical quantities on all energy scales.

Compared to the effective field theory framework the main advantage lies not primarily in the gained energy range in which reliable computations can be made, but rather that one has a chance to properly identify ‘large’ quantum gravity effects at low energies. Indeed the (presently known) low energy effects that arise in the effective field theory framework, although unambiguously defined, are suppressed by the powers of energy scale/Planck mass one would expect on dimensional grounds. Conversely, if there are detectable low energy imprints of quantum gravity they presumably arise from high energy (Planck scale) processes, in which case one has to computationally propagate their effect through many orders of magnitudes down to accessible energies.

This may be seen as the challenge a physically viable theory of quantum gravity has to meet, while the nature of the ‘fundamental’ degrees of freedom is of secondary importance. Indeed, from the viewpoint of renormalization theory it is the universality class that matters, not the particular choice of dynamical variables. Once a functional integral picture has been adopted, even nonlocally and nonlinearly related sets of fields or other variables may describe the same universality class – and hence the same physics.

The arena on which the renormalization group acts is a space of actions or, equivalently, a space of measures. A typical action has the form  $\sum_{\alpha} u_{\alpha} P_{\alpha}$ , where  $P_{\alpha}$  are interaction monomials (including kinetic terms) and the  $u_{\alpha}$  are scale dependent coefficients. The subset  $u_i$  which cannot be removed by field redefinitions are called *essential* parameters, or couplings. Usually one makes them dimensionless by taking out a suitable power of the scale parameter  $\mu$ ,  $g_i(\mu) = \mu^{-d_i} u_i(\mu)$ . In the following the term “essential coupling” will always refer to these dimensionless variants. We also presuppose the principles according to which a (Wilson–Kadanoff) renormalization flow is defined on this area. For the convenience of the reader a brief reminder is included in Appendix A. In the context of Quantum Gravidynamics some key notions (unstable manifold and continuum limit) have a somewhat different status which we outline below.

Initially all concepts in a Wilson–Kadanoff renormalization procedure refer to a choice of **coarse graining** operation. It is part of the physics premise of a functional integral type approach that there is a description independent and physically relevant distinction between coarse grained and fine grained geometries. On a classical level this amounts to the distinction, for example, between a perfect fluid solution of the field equations and one generated by its  $10^{30}$  or so molecular constituents. A sufficiently large set of Dirac observables would be able to discriminate two such spacetimes. Whenever we shall refer later on to “coarse grained” versus “fine grained” geometries we have a similar picture in mind for the ensembles of off-shell geometries entering a functional integral.

With respect to a given coarse graining operation one can ask whether the flow of actions or couplings has a **fixed point**. The existence of a fixed point is the *raison d’être* for the universality properties (in the statistical field theory sense) which eventually are ‘handed down’ to the physics in the low energy regime. By analogy with other field theoretical systems one should probably not expect that the existence (or nonexistence) of a (non-Gaussian) fixed point will be proven with mathematical rigor in the near future. From a physics viewpoint, however, it is the high degree of universality ensued by a fixed point that matters, rather than the existence in the mathematical

sense. For example non-Abelian gauge theories appear to have a (Gaussian) fixed point ‘for all practical purposes’, while their rigorous construction as the continuum limit of a lattice theory is still deemed a ‘millennium problem’. In the case of quantum gravity we shall present in Sections 3 and 4 in detail two new pieces of evidence for the existence of a (non-Gaussian) fixed point.

Accepting the existence of a (non-Gaussian) fixed point as a working hypothesis one is led to determine the structure of its **unstable manifold**. Given a coarse graining operation and a fixed point of it, the stable (unstable) manifold is the set of all points connected to the fixed point by a coarse graining trajectory terminating at it (emanating from it). It is not guaranteed though that the space of actions can in the vicinity of the fixed point be divided into a stable and an unstable manifold; there may be trajectories which develop singularities or enter a region of coupling space deemed unphysical for other reasons and thus remain unconnected to the fixed point. The stable manifold is the innocuous part of the problem; it is the unstable manifold which is crucial for the construction of a continuum limit. By definition it is swept out by flow lines emanating from the fixed point, the so-called *renormalized trajectories*. Points on such a flow line correspond to actions or measures which are called *perfect* in that they can be used to compute continuum answers for physical quantities even in the presence of an ultraviolet (UV) cutoff, like one which discretizes the base manifold. In practice the unstable manifold is not known and renormalized trajectories have to be identified approximately by a tuning process. What is easy to determine is whether in a given expansion “sum over coupling times interaction monomial” a coupling will be driven away from the value the corresponding coordinate has at the fixed point after a sufficient number of coarse graining steps (in which case it is called *relevant*) or will move towards this fixed point value (in which case it is called *irrelevant*). Note that this question can be asked even for trajectories which are not connected to the fixed point. The dimension of the unstable manifold equals the number of independent relevant interaction monomials that are ‘connected’ to the fixed point by a (renormalized) trajectory.

Typically the unstable manifold is indeed locally a manifold, though it may have cusps. Although ultimately it is only the unstable manifold that matters for the construction of a continuum limit, relevant couplings which blow up somewhere in between may make it very difficult to successfully identify the unstable manifold. In practice, if the basis of interaction monomials in which this happens is deemed natural and a change of basis in which the pathological directions could simply be omitted from the space of actions is very complicated, the problems caused by such a blow up may be severe. An important issue in practice is therefore whether in a natural basis of interaction monomials the couplings are ‘safe’ from such pathologies and the space of actions decomposes in the vicinity of the fixed point neatly into a stable and an unstable manifold. This regularity property is one aspect of “asymptotic safety”, as we shall see below.

A second caveat appears in infinite-dimensional situations. Whenever the coarse graining operates on an infinite set of potentially relevant interaction monomials, convergence issues in the infinite sums formed from them may render formally equivalent bases inequivalent. In this case the geometric picture of a (coordinate independent) manifold breaks down or has to be replaced by a more refined functional analytic framework. An example of a field theory with an infinite set of relevant interaction monomials is QCD in a lightfront formulation [174] where manifest Lorentz and gauge invariance is given up in exchange of other advantages. In this case it is thought that there are hidden dependencies among the associated couplings so that the number of independent relevant couplings is finite and the theory is eventually equivalent to conventional QCD. Such a *reduction of couplings* is nontrivial because a relation among couplings has to be preserved under the renormalization flow. In quantum gravity related issues arise to which we turn later.

As an interlude let us outline the role of **Newton’s constant** in a diffeomorphism invariant theory with a dynamical metric. Let  $S[g, \text{matter}]$  be any local action, where  $g = (g_{\alpha\beta})_{1 \leq \alpha, \beta \leq d}$  is the metric and the “matter” fields are not scaled when the metric is. Constant rescalings of the



metric then give rise to a variation of the Lagrangian which vanishes on a shell:

$$\frac{d}{d\omega^2} S[\omega^2 g, \text{matter}] \Big|_{\omega=1} = \int dx \sqrt{g} g_{\alpha\beta} \frac{\delta S[g, \text{matter}]}{\delta g_{\alpha\beta}}. \quad (1.1)$$

As a consequence one of the coupling parameters which in the absence of gravity would be *essential* (i.e. a genuine coupling) becomes *inessential* (i.e. can be changed at will by a redefinition of the fields). The running of this parameter, like that of a wave function renormalization constant, has no direct significance. If the pure gravity part contains the usual Ricci scalar term  $\sqrt{g}R(g)$ , the parameter that becomes inessential may be taken as its prefactor  $Z_N$ . Up to a dimension dependent coefficient it can be identified with the inverse of Newton's constant  $Z_N^{-1} \sim G_{\text{Newton}}$ , the latter defined through the nonrelativistic force law. It is also easy to see that in a background field formalism  $\omega$  sets the overall normalization of the spectral/momentum values. Hence in a theory with a dynamical metric the three (conceptually distinct) inessential parameters – overall scale of the metric, the inverse of Newton's constant, and the overall normalization of the spectral/momentum values – are in one-to-one correspondence (see Section 2.3.1 for details). For definiteness let us consider the running of Newton's constant here.

Being inessential, the quantum field theoretical running of  $G_{\text{Newton}}$  has significance only relative to the running coefficient of some reference operator. The most commonly used choice is a cosmological constant term  $\Lambda \int dx \sqrt{g}$ . Indeed

$$G_{\text{Newton}} \Lambda^{\frac{d-2}{d}} =: \text{const} \times \tau(\mu)^{2/d}, \quad (1.2)$$

is dimensionless and invariant under constant rescalings of the metric [116]. The associated essential coupling  $\tau(\mu)$  is in the present context assumed to be asymptotically safe, i.e.  $\sum_{\mu_0 \leq \mu \leq \infty} \tau(\mu) < \infty$ ,  $\lim_{\mu \rightarrow \infty} \tau(\mu) = \tau_*$ , where here  $0 < \tau_* < \infty$ . Factorizing it into the dimensionless Newton constant  $g_N \sim \mu^{d-2} G_{\text{Newton}}$  and  $\lambda(\mu) = 2\mu^d g_N(\mu) \Lambda$ , there are two possibilities: One is that the scheme choices are such that both  $g_N$  and  $\lambda$  behave like asymptotically safe couplings, i.e. satisfy Equation (1.3) below. This is advantageous for most purposes. The second possibility is realized when a singular solution for the flow equation for  $g_N$  is inserted into the flow equation for  $\lambda$ . This naturally occurs when  $G_{\text{Newton}}$ , viewed as an inessential parameter, is frozen at a prescribed value, say  $[G_{\text{Newton}}]^{1/(d-2)} = M_{\text{Pl}} \approx 1.4 \times 10^{19}$  GeV, which amounts to working with Planck units [172]. Then the  $g_N$  flow is trivial,  $g_N(\mu) \sim (\mu/M_{\text{Pl}})^{d-2}$ , but the flow equation for  $\lambda$  carries an explicit  $\mu$ -dependence. By and large both formulations are mathematically equivalent (see Section 2.3.1). For definiteness we considered here the cosmological constant term as a reference operator, but many other choices are possible. In summary, the dimensionless Newton constant can be treated either as an inessential parameter (and then frozen to a constant value) or as a quasi-essential coupling (in which case it runs and assumes a finite positive asymptotic value).

The unstable manifold of a fixed point is crucial for the construction of a continuum limit. The fixed point itself describes a strictly scale invariant situation. More precisely the situation at the fixed point is by definition invariant under the chosen coarse graining (i.e. scale changing) operation. In particular any dependence on an ultraviolet cutoff must drop out *at* the fixed point, which is why fixed points are believed to be indispensable for the construction of a scaling limit. If one now uses a different coarse graining operation the location of the fixed point will change in the given coordinate system provided by the essential couplings. One aspect of universality is that all field theories based on the fixed points referring to different coarse graining operations have the same long distance behavior.

This suggests to introduce the notion of a **continuum limit** as an ‘equivalence class’ of scaling limits in which the physical quantities become independent of the UV cutoff, largely independent of the choice of the coarse graining operation, and, ideally, invariant under local reparameterizations of the fields.

In the framework of statistical field theories one distinguishes between two construction principles, a *massless* scaling limit and a *massive* scaling limit. In the first case all the actions/measures on a trajectory emanating from the fixed point describe a scale invariant system, in the second case this is true only for the action/measure at the fixed point. In either case the unstable manifold of the given fixed point has to be at least one-dimensional. Here we shall exclusively be interested in the second construction principle. Given a coarse graining operation and a fixed point of it with a nontrivial unstable manifold a scaling limit is then constructed by ‘backtracing’ a renormalized trajectory emanating from the fixed point. The number of parameters needed to specify a point on the unstable manifold gives the number of possible scaling limits – not all of which must be physically distinct, however.

In this context it should be emphasized that the number of relevant directions in a chosen basis is *not* directly related to the predictive power of the theory. A number of authors have argued in the effective field theory framework that even theories with an infinite number of relevant parameters can be predictive [126, 16, 32]. This applies all the more if the theory under consideration is based on a fixed point, and thus not merely effective. One reason lies in the fact that the number of independent relevant directions connected to the fixed point might not be known. Hidden dependencies would then allow for a (genuine or effective) reduction of couplings [236, 160, 174, 11, 16]. For quantum gravity the situation is further complicated by the fact that *generic* physical quantities are likely to be related only nonlocally and nonlinearly to the metric. What matters for the predictive power is not the total number of relevant parameters but how the observables depend on them. To illustrate the point imagine a (hypothetical) case where  $n^2$  observables are injective functions of  $n$  relevant couplings each; then  $n$  measurements will determine the couplings, leaving  $n^2 - n$  predictions. This gives plenty of predictions, for any  $n$ , and it remains true in the limit  $n \rightarrow \infty$ , despite the fact that one then has infinitely many relevant couplings.

Infinitely many essential couplings naturally arise when a perturbative treatment of Quantum Gravidynamics is based on a  $1/p^2$  type propagator. As first advocated by Gomis and Weinberg [94] the use of a  $1/p^2$  type graviton propagator in combination with higher derivative terms avoids the problems with unitarity that occur in other treatments of higher derivative theories. Consistency requires that quadratic counterterms (those which contribute to the propagator) can be absorbed by field redefinitions. This can be seen to be the case [10] either in the absence of a cosmological constant term or when the background spacetime admits a metric with constant curvature. The price to pay for the  $1/p^2$  type propagator is that all nonquadratic counterterms have to be included in the bare action, so that independence of the UV cutoff can only be achieved with infinitely many essential couplings, but it can be [94]. In order to distinguish this from the familiar notion of perturbative renormalizability with finitely many couplings we shall call such theories (perturbatively) *weakly renormalizable*. Translated into Wilsonian terminology the above results then show the existence of a “weakly renormalizable” but “propagator unitary” Quantum Gravidynamics based on a **perturbative Gaussian fixed point**.

The beta functions for this infinite set of couplings are presently unknown. If they were known, expectations are that at least a subset of the couplings would blow up at some finite momentum scale  $\mu = \mu_{\text{term}}$  and would be unphysical for  $\mu > \mu_{\text{term}}$ . In this case the computed results for physical quantities (“reaction rates”) are likely to blow up likewise at some (high) energy scale  $\mu = \mu_{\text{term}}$ .

This illustrates Weinberg’s concept of asymptotic safety. To quote from [227]: “A theory is said to be **asymptotically safe** if the essential coupling parameters approach a fixed point as the momentum scale of their renormalization point goes to infinity”. Here ‘the’ essential couplings  $g_i$  are those which are useful for the absorption of cutoff dependencies, i.e. not irrelevant ones. The momentum scale is the above  $\mu$ , so that the condition amounts to having nonterminating

trajectories for the  $g_i$ 's with a finite limit:

$$\sup_{\mu_0 \leq \mu \leq \infty} g_i(\mu) < \infty, \quad \lim_{\mu \rightarrow \infty} g_i(\mu) = g_i^* < \infty. \quad (1.3)$$

In other words in an asymptotically safe theory the above blow-up in the couplings and hence in physical observables does not occur. We suggest to call couplings satisfying Equation (1.3) *asymptotically safe*. As a specification one should add [227]: “Of course the question whether or not an infinity in coupling constants betokens a singularity in reaction rates depends on how the coupling constants are parameterized. We could always adopt a perverse definition (e.g.  $\tilde{g}(\mu) = (g_* - g(\mu))^{-1}$ ) such that reaction rates are finite even at an infinity of the coupling parameters. This problem can be avoided if we define the coupling constants as coefficients in a power series expansion of the reaction rates themselves around some physical renormalization point”.

A similar remark applies to the signs of coupling constants. When defined through physical quantities certain couplings or coupling combinations will be constrained to be positive. For example in a (nongravitational) effective field theory this constrains the couplings of a set of leading power counting irrelevant operators to be positive [2]. In an asymptotically safe theory similar constraints are expected to arise and are crucial for its physics viability.

Note that whenever the criterion for asymptotic safety is met, *all* the relevant couplings lie in the unstable manifold of the fixed point (which is called the “UV critical surface” in [227], Page 802, a term now usually reserved for the surface of infinite correlation length). The regularity property described earlier is then satisfied, and the space of actions decomposes in the vicinity of the fixed point into a stable and an unstable manifold.

Comparing the two perturbative treatments of Quantum Gravidynamics described earlier, one sees that they have complementary advantages and disadvantages: Higher derivative theories based on a  $1/p^4$  propagator are strictly renormalizable with couplings that are presumed to be asymptotically safe; however unphysical propagating modes are present. Defining higher derivative gravity perturbatively with respect to a  $1/p^2$  propagator has the advantage that all propagating modes are physical, but infinitely many essential couplings are needed, a subset of which is presumed to be not asymptotically safe. From a technical viewpoint the challenge of Quantum Gravidynamics lies therefore not so much in achieving renormalizability but to **reconcile asymptotically safe couplings with the absence of unphysical propagating modes**.

The solution of this ‘technical’ problem is likely also to give rise to enhanced predictability properties, which should be vital to make the theory phenomenologically interesting. Adopting the second of the above perturbative constructions one sees that situation is similar to, for example, perturbative QED. So, apart from esthetic reasons, why not be content with physically motivated couplings that display a ‘Landau’ pole, and hence with an effective field theory description? Predictability in principle need not a be problem. The previous remarks about the predictability of theories with infinitely many essential couplings apply here. Even in Quantum Gravidynamics based on the perturbative Gaussian fixed point, some lowest order corrections are unambiguously defined (independent of the scale  $\mu_{\text{term}}$ ), as stressed by Donoghue (see [32] and references therein). In our view [82], as mentioned earlier, the main rationale for trying to go beyond Quantum Gravidynamics based on the perturbative Gaussian fixed point is not the infinite number of essential couplings, but the fact that the size of the corrections is invariably governed by power-counting dimensions. As a consequence, in the energy range where the computations are reliable the corrections are way too small to be phenomenologically interesting. Conversely, if there is a physics of quantum gravity, which is experimentally accessible and adequately described by some Quantum Gravidynamics, the above two features need to be reconciled – perturbatively or nonperturbatively.

Assuming that this can be achieved certain qualitative features such a gravitational functional integral must have can be inferred without actually evaluating it. One is the presence of anti-screening configurations, the other is a dimensional reduction phenomenon in the ultraviolet.

In non-Abelian gauge theories the **anti-screening** phenomenon can be viewed as the physics mechanism underlying their benign high energy behavior (as opposed to Abelian gauge theories, say); see e.g. [175] for an intuitive discussion. It is important not to identify “anti-screening” with its most widely known manifestation, the sign of the dominant contribution to the one-loop beta function. In an exact continuum formulation of a pure Yang–Mills theory, say, the correlation functions do not even depend on the gauge coupling. Nevertheless they indirectly do know about “asymptotic freedom” through their characteristic high energy behavior. In the functional integral measure this comes about through the dominance of certain configurations/histories which one might also call “anti-screening”.

By analogy one would expect that in a gravitational functional integral which allows for a continuum limit, a similar mechanism is responsible for its benign ultraviolet behavior (as opposed to the one expected by power counting considerations with respect to a  $1/p^2$  propagator, say). Some insight into the nature of this mechanism can be gained from a Hamiltonian formulation of the functional integral (authors, unpublished) but a concise characterization of the “anti-screening” geometries/histories, ideally in a discretized setting, remains to be found. By definition the dominance of these configurations/histories would be responsible for the benign ultraviolet properties of the discretized functional integral based on a non-Gaussian fixed point. Conversely understanding the nature of these antiscreening geometries/histories might help to design good discretizations. A discretization of the gravitational functional integral which allows for a continuum limit might also turn out to exclude or dynamically disfavor configurations that are taken into account in other, off-hand equally plausible, discretizations. Compared to such a naive discretization it will look as if a constraint on the allowed configurations/histories has been imposed. For want of a better term we call this an “anti-screening constraint”. A useful analogy is the inclusion of a causality constraint in the definition of the (formal Euclidean) functional integral originally proposed by Teitelboim [212, 213], and recently put to good use in the framework of dynamical triangulations [5]. Just as the inclusion of a good causality constraint is justified retroactively, so would be the inclusion of a suitable “antiscreening” constraint.

A second qualitative property of a gravitational functional integral where the continuum limit is based on a non-Gaussian fixed point is a **dimensional reduction of the residual interactions in the UV**. There are several arguments for this phenomenon which will be described in Section 2.4. Perhaps the simplest one is based on the large anomalous dimensions at a non-Gaussian fixed point and runs as follows: (We present here a formulation independent variant [157] of the argument first used in [133].) Suppose that the unknown microscopic action is local and reparameterization invariant. The only term containing second derivatives then is the familiar Einstein–Hilbert term  $\int dx \sqrt{g} R(g)$  of mass dimension  $2 - d$  in  $d$  dimensions, if the metric is taken dimensionless. As explained before the dimensionful running prefactor  $Z_N$  multiplying it plays a double role, once as a wave function renormalization constant and once as a quasi-essential coupling  $g_N(\mu)$ . Both aspects are related as outlined before; in particular

$$Z_N(\mu) = \frac{\mu^{d-2}}{g_N(\mu)}. \quad (1.4)$$

Here  $g_N$  is a dimensionless coupling which is treated as “quasi-essential” and whose running may also depend on all the other couplings (gravitational and non-gravitational) made dimensionless by taking out a suitable power of  $\mu$ . The short distance behavior of the propagator will now be governed by the “anomalous dimension”  $\eta_N = -\mu \partial_\mu \ln Z_N(\mu)$ , by the usual field theoretical arguments, say, via the Callan–Symanzik equation for the effective action. On the other hand the flow equation for  $g_N$  can be expressed in terms of  $\eta_N$  as

$$\mu \partial_\mu g_N = [d - 2 + \eta_N(g_N, \text{other})] g_N, \quad (1.5)$$

where we schematically indicated the dependence on the other dimensionless couplings. *If* this

flow equation now has a nontrivial fixed point  $\infty > g_N^* > 0$ , the only way how the right-hand-side can vanish is for  $\eta_N(g_N^*, \text{other}) = 2 - d$ , irrespective of the detailed behavior of the other couplings as long as no blow-up occurs. This is a huge anomalous dimension. For a graviton “test propagator” (see below) the key property of  $\eta_N$  is that it gives rise to a high momentum behavior of the form  $(p^2)^{-1+\eta_N/2}$  modulo logarithms, or a short distance behavior of the form  $(\sqrt{x^2})^{2-d-\eta_N}$  modulo logarithms. Keeping only the leading part the vanishing power at  $\eta_N = 2 - d$  translates into a logarithmic behavior,  $\ln x^2$ , formally the same as for massless Klein–Gordon fields in a two-dimensional field theory. We shall comment on potential pitfalls of such an argument below.

In accordance with this argument a  $1/p^4$  type propagator goes hand in hand with a non-Gaussian fixed point for  $g_N$  in two other computational settings: in strictly renormalizable higher derivative theories (see Section 2.3.2 and in the  $1/N$  expansion [216, 217, 203]. In the latter case a nontrivial fixed point goes hand in hand with a graviton propagator whose high momentum behavior is of the form  $1/(p^4 \ln p^2)$ , in four dimensions, and formally  $1/p^d$  in  $d$  dimensions.

The fact that a large anomalous dimension occurs at a non-Gaussian fixed point was first observed in the context of the  $2 + \epsilon$  expansion [116, 117] and then noticed in computations based on truncated flow equations [133]. The above variant of the argument [157] shows that no specific computational information enters. It highlights what is special about the Einstein–Hilbert term (within the class of local gravitational actions): it is the kinetic (second derivative) term itself which carries a dimensionful coupling. Of course one could assign to the metric a mass dimension 2, in which case Newton’s constant would be dimensionless. However one readily checks that then the wave function renormalization constant of a standard matter kinetic term acquires a mass dimension  $d - 2$  for bosons and  $d - 1$  for fermions, respectively. Assuming that the dimensionless parameter associated with them remains nonzero as  $\mu \rightarrow \infty$ , one can repeat the above argument and finds now that *all* matter propagators have a  $1/p^d$  high momentum behavior, or a  $\ln x^2$  short distance behavior. It is this universality which justifies to attribute the modification in the short distance behavior of the fields to a modification of the underlying (random) geometry. This may be viewed as a specific variant of the old expectation that gravity acts as a short distance regulator.

Let us stress that while the anomalous dimension always governs the UV behavior in the vicinity of a (UV) fixed point, it is in general *not* related to the geometry of field propagation (see [125] for a discussion in QCD). What is special about gravity is ultimately that the propagating field itself determines distances. In the context of the above argument this is used in the reshuffling of the soft UV behavior to matter propagators. The propagators used here should be viewed as “test propagators”, not as physical ones. One transplants the information in  $\eta_N$  derived from the gravitational functional integral into a conventional propagator on a (flat or curved) background spacetime. The reduced dimension two should be viewed as an “interaction dimension” specifying roughly the (normalized) number of independent degrees of freedom a randomly picked one interacts with.

The same conclusion ( $1/p^d$  propagators or interaction dimension 2) can be reached in a number of other ways as well, which are described in Section 2.4. A more detailed understanding of the microstructure of the random geometries occurring in an asymptotically safe functional integral remains to be found (see however [135, 134]).

Accepting this dimensional reduction as a working hypothesis it is natural to ask whether there exists a two-dimensional field theory which provides an quantitatively accurate (‘effective’) description of this extreme UV regime. Indeed, one can identify a number of characteristics such a field theory should have, using only the main ideas of the scenario (see the end of Section 2.4). The asymptotic safety of such a field theory would then strongly support the corresponding property of the full theory and the self-consistency of the scenario. In summary, we have argued that the qualitative properties of the gravitational functional integral in the extreme ultraviolet follow directly from the previously highlighted principles: the existence of a nontrivial UV fixed point, asymptotic safety of the couplings, and antiscreening. Moreover these UV properties can be probed



for self-consistency.

## 1.2 Evidence for asymptotic safety

Presently the evidence for asymptotic safety in quantum gravity comes from the following very different computational settings: the  $2 + \epsilon$  expansion, perturbation theory of higher derivative theories, a large  $N$  expansion in the number of matter fields, the study of symmetry truncations, and that of truncated functional flow equations. Arguably none of the pieces of evidence is individually compelling but taken together they make a strong case for asymptotic safety.

The results from the  $2 + \epsilon$  expansion were part of Weinberg’s original motivation to propose the scenario. Since gravity in two and three dimensions is non-dynamical, however, the lessons for a genuine quantum gravitational dynamics are somewhat limited. Higher derivative theories were known to be strictly renormalizable with a finite number of couplings, at the expense of having unphysical propagating modes (see [207, 206, 83, 19, 59]). In hindsight one can identify a non-Gaussian fixed point for Newton’s constant already in this setting (see [54] and Section 2.3). The occurrence of this non-Gaussian fixed point is closely related to the  $1/p^4$ -type propagator that is used. The same happens when (Einstein or a higher derivative) gravity is coupled to a large number  $N$  of matter fields and a  $1/N$  expansion is performed. A nontrivial fixed point is found that goes hand in hand with a  $1/p^4$ -type propagator (modulo logs), which here arises from a resummation of matter self-energy bubbles, however.

As emphasized before the challenge of Quantum Gravidynamics is not so much to achieve (perturbative or nonperturbative) renormalizability but to reconcile asymptotically safe couplings with the absence of unphysical propagating modes. Two recent developments provide complementary evidence that this might indeed be feasible. Both of these developments take into account the dynamics of infinitely many physical degrees of freedom of the four-dimensional gravitational field. In order to be computationally feasible the ‘coarse graining’ has to be constrained somehow. To do this the following two strategies have been pursued (which we label Strategies (c) and (d) according to the discussion below):

(c) The metric fluctuations are constrained by a symmetry requirement, but the full (infinite-dimensional) renormalization group dynamics is considered. We shall refer to this as the strategy via *symmetry reductions*.

(d) All metric fluctuations are taken into account but the renormalization group dynamics is projected onto a low-dimensional submanifold. Since this is done using truncations of functional renormalization group equations, we shall refer to this as the strategy via *truncated functional flow equations*.

Both strategies (truncation in the fluctuations but unconstrained flow and unconstrained quantum fluctuations but constrained flow) are complementary. Tentatively both results are related by the dimensional reduction phenomenon described before (see Section 2.4). The techniques used are centered around the background effective action, but are otherwise fairly different. For the reader’s convenience we included summaries of the relevant aspects in Appendices A and B. The main results obtained from Strategies (c) and (d) are reviewed in Sections 3 and 4, respectively.

For the remainder of this section we now first survey the pieces of evidence from all the computational settings (a–d):

**(a) Evidence from  $2 + \epsilon$  expansions:** In the non-gravitational examples of perturbatively non-renormalizable field theories with a non-Gaussian fixed point the non-Gaussian fixed point can be viewed as a ‘remnant’ of an asymptotically free fixed point in a lower-dimensional version of the theory. It is thus natural to ask how gravity behaves in this respect. In  $d = 2$  space-time dimensions Newton’s constant  $g_N$  is dimensionless, and formally the theory with the bare action  $g_N^{-1} \int d^2x \sqrt{g} R(g)$  is power counting renormalizable in perturbation theory. However, as

the Einstein–Hilbert term is purely topological in two dimensions, the inclusion of local dynamical degrees of freedom requires, at the very least, starting from  $2 + \epsilon$  dimensions and then studying the behavior near  $\epsilon \rightarrow 0^+$ . The resulting “ $\epsilon$ -expansion” amounts to a double expansion in the number of ‘graviton’ loops and in the dimensionality parameter  $\epsilon$ . Typically dimensional regularization is used, in which case the UV divergencies give rise to the usual poles in  $1/\epsilon$ . Specific for gravity are however two types of complications. The first one is due to the fact that  $\int d^{2+\epsilon}x \sqrt{g}R(g)$  is topological at  $\epsilon = 0$ , which gives rise to additional “kinematical” poles of order  $1/\epsilon$  in the graviton propagator. The goal of the renormalization process is to remove both the ultraviolet and the kinematical poles in physical quantities. The second problem is that in pure gravity Newton’s constant is an inessential parameter, i.e. it can be changed at will by a field redefinition. Newton’s constant  $g_N$  can be promoted to a coupling proper by comparing its flow with that of the coefficient of some reference operator, which is fixed to be constant.

For the reference operator various choices have been adopted (we follow the discussion in Kawai et al. [118, 116, 117, 3] with the conventions of [117]):

- (i) a cosmological constant term  $\int d^{2+\epsilon}x \sqrt{g}$ ,
- (ii) monomials from matter fields which are quantum mechanically non-scale invariant in  $d = 2$ ,
- (iii) monomials from matter fields which are quantum mechanically scale invariant in  $d = 2$ ,
- (iv) the conformal mode of the metric itself in a background field expansion.

All choices lead to a flow equation of the form

$$\mu \frac{d}{d\mu} g_N = \epsilon g_N - \gamma g_N^2, \quad (1.6)$$

but the coefficient  $\gamma$  depends on the choice of the reference operator [118]. For all  $\gamma > 0$  there is a nontrivial fixed point  $g_N^* = \epsilon/\gamma > 0$  with a one-dimensional unstable manifold. In other words  $g_N$  is an asymptotically safe coupling in  $2 + \epsilon$  dimensions, and the above rule of thumb suggests that this a remnant of a nontrivial fixed point in  $d = 4$  with respect to which  $g_N$  is asymptotically safe (see Section 1.4 for the renormalization group terminology).

Technically the non-universality of  $\gamma$  arises from the before-mentioned kinematical poles. In the early papers [86, 53, 227] the Choice **i** was adopted giving  $\gamma = 19/(24\pi)$ , or  $\gamma = (19 - c)/(24\pi)$  if free matter of central charge  $c$  is minimally coupled. A typical choice for Choice **ii** is a mass term of a Dirac fermion, a typical choice for Choice **iii** is the coupling of a four-fermion (Thirring) interaction. Then  $\gamma$  comes out as  $\gamma = (19 + 6\Delta_0 - c)/(24\pi)$ , where  $\Delta_0 = 1/2, 1$ , respectively. Here  $\Delta_0$  is the scaling dimension of the reference operator, and again free matter of central charge  $c$  has been minimally coupled. It has been argued in [118] that the loop expansion in this context should be viewed as double expansion in powers of  $\epsilon$  and  $1/c$ , and that reference operators with  $\Delta_0 = 1$  are optimal. The Choice **iv** has been pursued systematically in a series of papers by Kawai et al. [116, 117, 3]. It is based on a parameterization of the metric in terms of a background metric  $\bar{g}_{\mu\nu}$ , the conformal factor  $e^\sigma$ , and a part  $f_{\mu\nu}$  which is traceless,  $\bar{g}^{\mu\nu} f_{\mu\nu} = 0$ . Specifically  $g_{\mu\nu} = \bar{g}_{\mu\rho} (e^f)^\rho{}_\nu e^\sigma$  is inserted into the Einstein–Hilbert action; propagators are defined (after gauge fixing) by the terms quadratic in  $\sigma$  and  $f_{\mu\nu}$ , and vertices correspond to the higher order terms. This procedure turns out to have a number of advantages. First the conformal mode  $\sigma$  is renormalized differently from the  $f_{\mu\nu}$  modes and can be viewed as defining a reference operator in itself; in particular the coefficient  $\gamma$  comes out as  $\gamma = (25 - c)/(24\pi)$ . Second, and related to the first point, the system has a well-defined  $\epsilon$ -expansion (absence of poles) to all loop orders. Finally this setting allows one to make contact to the exact (KPZ [122]) solution of two-dimensional quantum gravity

in the limit  $\epsilon \rightarrow 0$ .

**(b) Evidence from perturbation theory and large  $N$ :** Modifications of the Einstein–Hilbert action where fourth derivative terms are included are known to be perturbatively renormalizable [206, 83, 19, 59]. A convenient parameterization is

$$S_4 = \int dx \sqrt{g} \left[ \Lambda - \frac{1}{c_d G_N} R + \frac{1}{2s} C^2 - \frac{\omega}{3s} R^2 + \frac{\theta}{s} E \right]. \quad (1.7)$$

Here  $c_d$  is a constant such that  $c_4 = 16\pi$ ,  $C^2$  is the square of the Weyl tensor, and  $E$  is the integrand of the Gauss–Bonnet term. In  $d = 4$  the latter is negligible, unless dimensional regularization is used. The sign of the crucial  $C^2$  coupling  $s > 0$  is fixed by the requirement that the Euclidean functional integral is damping. The one-loop beta functions for the (non-negative) couplings,  $s$ ,  $\omega$ ,  $\theta$ , are known and on the basis of them these couplings are expected to be asymptotically safe. In particular  $s$  is asymptotically free,  $\lim_{\mu \rightarrow 0} s(\mu) = 0$ . The remaining couplings  $\Lambda$  and  $c_d G_N$  are made dimensionless via  $c_d G_N = \mu^{-2} g_N$ ,  $\Lambda = \mu^4 2\lambda/g_N$ , where  $\mu$  is the renormalization scale. At  $s = 0$  these flow equations are compatible with the existence of a non-trivial fixed point for Newton’s constant,  $g_N^* \neq 0$ . The value of  $g_N^*$  is highly nonuniversal but it cannot naturally be made to vanish, i.e. the nontrivial and the trivial fixed point,  $g_N^* = 0$ , do not merge. The rationale for identifying a nontrivial fixed point by perturbative means is explained in Section 2.2. The benign renormalizability properties seen in this framework are mostly due to the  $1/p^4$  type propagator, at the expense of unphysical propagating modes.

The action (1.7) can be supplemented by a matter action, containing a large number,  $O(N)$ , of free matter fields. One can then keep the product  $N \cdot c_d G_N$  fixed, retain the usual normalization of the matter kinetic terms, and expand in powers of  $1/N$ . Renormalizability of the resulting ‘large  $N$  expansion’ then amounts to being able to remove the UV cutoff order by order in the formal series in  $1/N$ . This type of studies was initiated by Tomboulis where the gravity action was taken either the pure Ricci scalar [216], Ricci plus cosmological term [203], or a higher derivative action [217], with free fermionic matter in all cases. More recently the technique was reconsidered [169] with Equation (1.7) as the gravity action and free matter consisting of  $Nn_S$  scalar fields,  $Nn_D$  Dirac fields, and  $Nn_M$  Maxwell fields.

Starting from the Einstein–Hilbert action the high energy behavior of the usual  $1/p^2$ -type propagator gets modified. To leading order in  $1/N$  the modified propagator can be viewed as the graviton propagator with an infinite number of fermionic self-energy bubbles inserted and resummed. The resummation changes the high momentum behavior from  $1/p^2$  to  $1/(p^4 \ln p^2)$ , in four dimensions. In  $2 < d < 4$  dimensions the resulting  $1/N$  expansion is believed to be renormalizable in the sense that the UV cutoff  $\Lambda$  can strictly be removed order by order in  $1/N$  without additional (counter) terms in the Lagrangian. In  $d = 4$  the same is presumed to hold provided an extra  $C^2$  term is included in the bare Lagrangian, as in Equation (1.7). After removal of the cutoff the beta functions of the dimensionless couplings can be analyzed in the usual way and already their leading  $1/N$  term will decide about the flow pattern.

The qualitative result (due to Tomboulis [216] and Smolin [203]) is that there exists a nontrivial fixed point for the dimensionless couplings  $g_N$ ,  $\lambda$ , and  $s$ . Its unstable manifold is three dimensional, i.e. all couplings are asymptotically safe. Repeating the computation in  $2 + \epsilon$  dimensions the fixed point still exists and (taking into account the different UV regularization) corresponds to the large  $c$  (central charge) limit of the fixed point found the  $2 + \epsilon$  expansion.

These results have recently been confirmed and extended by Percacci [169] using the heat kernel expansion. In the presence of  $Nn_S$  scalar fields,  $Nn_D$  Dirac fields, and  $Nn_M$  Maxwell fields, the



flow equations for  $g_N$ ,  $\lambda$  and  $s$  come out to leading order in  $1/N$  as

$$\begin{aligned}\mu \frac{d}{d\mu} g_N &= 2g_N + \frac{1}{(4\pi)^2} \frac{1}{6} (n_S - 2n_D - 4n_M) g_N^2, \\ \mu \frac{d}{d\mu} \lambda &= -2\lambda + \frac{1}{(4\pi)^2} \left[ \frac{1}{6} (n_S - 2n_D - 4n_M) \lambda g_N - \frac{1}{4} (n_S - 4n_D + 2n_M) g_N \right], \\ \mu \frac{d}{d\mu} s &= -\frac{1}{(4\pi)^2} \frac{1}{280} (6n_S + 25n_D + 72n_M) s^2.\end{aligned}\quad (1.8)$$

One sees that the  $C^2$  coupling is always asymptotically free, and that Newton's constant has a nontrivial fixed point,  $g_N/(4\pi)^2 = 12/(-n_S + 2n_D + 4n_M)$ , which is positive if the number of matter fields is not too large.

As a caveat one should add that the  $1/p^4$ -type propagators occurring both in the perturbative and in the large  $N$  framework are bound to have an unphysical pole at some intermediate momentum scale. This pole corresponds to unphysical propagating modes and it is the price to pay for (strict) perturbative renormalizability combined with asymptotically safe couplings. From this point of view, the main challenge of Quantum Gravidynamics lies in reconciling asymptotically safe couplings with the absence of unphysical propagating modes. Precisely this can be achieved in the context of the  $2+2$  reduction.

**(c) Evidence from symmetry reductions:** Here one considers the usual gravitational functional integral but restricts it from “4-geometries modulo diffeomorphisms” to “4-geometries constant along a  $2+2$  foliation modulo diffeomorphisms”. This means that instead of the familiar  $3+1$  foliation of geometries one considers a foliation in terms of two-dimensional hypersurfaces  $\Sigma$  and performs the functional integral only over configurations that are constant as one moves along the stack of two-surfaces. Technically this constancy condition is formulated in terms of two commuting vector fields  $K_a = K_a^\alpha \partial_\alpha$ ,  $a = 1, 2$ , that are Killing vectors of the class of geometries  $g$  considered,  $\mathcal{L}_{K_a} g_{\alpha\beta} = 0$ . For definiteness we consider here only the case where both Killing vectors are spacelike. From this pair of Killing vector fields one can form the symmetric  $2 \times 2$  matrix  $M_{ab} := g_{\alpha\beta} K_a^\alpha K_b^\beta$ . Then  $\gamma_{\alpha\beta} := g_{\alpha\beta} - M^{ab} K_{a\alpha} K_{b\beta}$  (with  $M^{ab}$  the components of  $M^{-1}$  and  $K_{a\alpha} := g_{\alpha\beta} K_a^\beta$ ) defines a metric on the orbit space  $\Sigma$  which obeys  $\mathcal{L}_{K_a} \gamma_{\alpha\beta} = 0$  and  $K_a^\alpha \gamma_{\alpha\beta} = 0$ . The functional integral is eventually performed over metrics of the form

$$g_{\alpha\beta} = \gamma_{\alpha\beta} + M^{ab} K_{a\alpha} K_{b\beta}, \quad (1.9)$$

where the 10 components of a metric tensor are parameterized by the  $3+3$  independent functions in  $\gamma_{\alpha\beta}$  and  $M_{ab}$ . Each of these functions is constant along the stack of two-surfaces but may be arbitrarily rough within a two-surface.

In the context of the asymptotic safety scenario the restriction of the functional integral to metrics of the form (1.9) is a very fruitful one:

- The restricted functional integral inherits the perturbative non-renormalizability (with finitely many relevant couplings) from the full theory.
- It takes into account the crucial ‘spin-2’ aspect, that is, linear and nonlinear gravitational waves with two independent polarizations per spacetime point are included.
- It goes beyond the Eikonal approximation [209, 71] whose dynamics can be understood via a related  $2+2$  decomposition [113, 73].
- Based on heuristic arguments the dynamics of full Quantum Gravidynamics is expected to be effectively two-dimensional in the extreme ultraviolet with qualitative properties resembling that of the  $2+2$  truncation. The renormalization of the  $2+2$  truncation can thus serve as a prototype study and its asymptotic safety probes the self-consistency of the scenario.

- For the restricted functional integral the full infinite-dimensional renormalization group dynamics can be studied; it reveals both a Gaussian and a non-Gaussian fixed point, where the properties of the latter are compatible with the existence of a non-perturbative continuum limit.

Two additional bonus features are: In this sector the explicit construction of Dirac observables is feasible (classically and presumably also in the quantum theory). Finally a large class of matter couplings is easily incorporated.

As mentioned the effective dynamics looks two-dimensional. Concretely the classical action describing the dynamics of the 2-Killing vector subsector is that of a non-compact symmetric space sigma-model non-minimally coupled to 2D gravity via the “area radius”  $\rho := \sqrt{\det(M_{ab})_{1 \leq a, b \leq 2}}$ , of the two Killing vectors. To avoid a possible confusion let us stress, however, that the system is very different from most other models of quantum gravity (mini-superspace, 2D quantum gravity or dilaton gravity, Liouville theory, topological theories) in that it has infinitely many local and self-interacting dynamical degrees of freedom. Moreover these are literally (an infinite subset of) the degrees of freedom of the four-dimensional gravitational field, not just analogues thereof. The corresponding classical solutions (for both signatures of the Killing vectors) have been widely studied in the general relativity literature, c.f. [98, 26, 121]. We refer to [45, 46, 56] for details on the reduction procedure and [197] for a canonical formulation.

Technically the renormalization is done by borrowing covariant background field techniques from Riemannian sigma-models (see [84, 110, 201, 57, 220, 162]). In the particular application here the sigma-model perturbation theory is partially nonperturbative from the viewpoint of a graviton loop expansion as not all of the metric degrees of freedom are Taylor expanded in the bare action (see Section 3.2). This together with the field reparameterization invariance blurs the distinction between a perturbative and a non-perturbative treatment of the gravitational modes. The renormalization can be done to all orders of sigma-model perturbation theory, which is ‘non-really-perturbative’ for the gravitational modes. It turns out that strict cutoff independence can be achieved only by allowing for infinitely many essential couplings. They are conveniently combined into a generating functional  $h$ , which is a positive function of one real variable. Schematically the renormalized action takes the form [154]

$$S[g] = S_{\text{EH}} \left[ \frac{h(\rho)}{\rho} g \right] + \text{other second derivative terms.} \quad (1.10)$$

Here  $g$  is a metric of the form (1.9),  $S_{\text{EH}}[g]$  is the Einstein–Hilbert action evaluated on it, and  $h(\rho)$  is the generating coupling function evaluated on the renormalized area radius field  $\rho$ . Higher derivative terms are not needed in this subsector for the absorption of counter terms; the “other second derivative terms” needed are known explicitly.

This “coupling functional” is scale dependent and is subject to a flow equation of the form

$$\mu \frac{d}{d\mu} h = \beta_h(h), \quad (1.11)$$

where  $\mu$  is the renormalization scale and  $\mu \mapsto h(\cdot, \mu)$  is the ‘running’ generating functional. To preclude a misunderstanding let us stress that the function  $h(\cdot, \mu)$  changes with  $\mu$ , irrespective of the name of the argument, not just its value on  $\rho$ , say. Interestingly a closed formula for the beta function (or functional) in Equation (1.11) can be found [154, 155]. The resulting flow equation is a nonlinear partial integro-differential equation and difficult to analyze. The fixed points however are easily found. Apart from the degenerate ‘Gaussian’ one,  $1/h \equiv 0$ , there is a nontrivial fixed point  $h^{\text{beta}}(\cdot)$ . For the Gaussian fixed point a linearized stability analysis is empty, the structure of the quadratic perturbation equation suggests that it has both attractive and repulsive directions in the space of functions  $h$ . For the non-Gaussian fixed point  $h^{\text{beta}}(\cdot)$  a linearized

stability analysis is non-empty and leads to a system of linear integro-differential equations. It can be shown [155] that *all* linearized perturbations decay for  $\mu \rightarrow \infty$ , which is precisely what Weinberg's criterion for asymptotic safety asks for. Moreover the basic propagator used is free from unphysical poles. Applying the criterion described in Section 1.3 this strongly suggests that a continuum limit exist for the  $2 + 2$  reduced Quantum Gravidynamics beyond approximations (like the sigma-model perturbation theory/partially nonperturbative graviton expansion used to compute Equation (1.11)). See [158] for a proposed 'exact' bootstrap construction, whose relation to a  $2 + 2$  truncated functional integral however remains to be understood.

In summary, in the context of the  $2 + 2$  reduction an asymptotically safe coupling flow can be reconciled with the absence of unphysical propagating modes. In contrast to the technique on which Evidence (d) below is based the existence of an infinite cutoff limit here can be *shown* and does not have to be stipulated as a hypothesis subsequently probed for self-consistency. Since the properties of the  $2 + 2$  truncation qualitatively are the ones one would expect from an 'effective' field theory describing the extreme UV aspects of Quantum Gravidynamics (see the end of Section 2.4), its asymptotic safety is a strong argument for the self-consistency of the scenario.

**(d) Evidence from truncated flows of the effective average action:** The *effective average action*  $\Gamma_{\Lambda,k}$  is a generating functional generalizing the usual effective action, to which it reduces for  $k = 0$ . Here  $\Gamma_{\Lambda,k}$  depends on the UV cutoff  $\Lambda$  and an additional scale  $k$ , indicating that in the defining functional integral roughly the field modes with momenta  $p$  in the range  $k \leq p \leq \Lambda$  have been integrated out. Correspondingly  $\Gamma_{\Lambda,\Lambda}$  gives back the bare action and  $\Gamma_{\Lambda,0} = \Gamma_{\Lambda}$  is the usual quantum effective action, in the presence of the UV cutoff  $\Lambda$ . The modes in the momentum range  $k \leq p \leq \Lambda$  are omitted or suppressed by a mode cutoff 'action'  $C_{\Lambda,k}$ , and one can think of  $\Gamma_{\Lambda,k}$  as being the conventional effective action  $\Gamma_{\Lambda}$  but computed with a bare action that differs from the original one by the addition of  $C_{\Lambda,k}$ ; specifically

$$\Gamma_{\Lambda,k} = -C_{\Lambda,k} + \Gamma_{\Lambda} \Big|_{S \mapsto S + C_{\Lambda,k}}. \quad (1.12)$$

A summary of the key properties of the effective average action (1.12) can be found in Appendix C. Here we highlight that from the regularized functional integral defining  $\Gamma_{\Lambda,k}$  an ('exact') functional renormalization group equation (FRGE) can be derived. Schematically it has the form  $k \frac{d}{dk} \Gamma_{\Lambda,k} = \text{rhs}$ , where the "right-hand-side" involves the Hessian of  $\Gamma_{\Lambda,k}$  with respect to the dynamical fields. The FRGE itself (that is, its right-hand-side) carries no *explicit* dependence on the UV cutoff, or one which can trivially be removed. However the removal of the UV regulator  $\Lambda$  implicit in the definition of  $\Gamma_{\Lambda,k}$  is nontrivial and is related to the traditional UV renormalization problem (see Section 2.2). Whenever massless degrees of freedom are involved, also the existence of the  $k \rightarrow 0$  limit of  $\Gamma_{\Lambda,k}$  is nontrivial and requires identification of the proper infrared degrees of freedom. In the present context we take this for granted and focus on the UV aspects.

The effective average action has been generalized to gravity [179] and we shall describe it and its properties in more detail in Sections 4.1 and 4.2. As before the metric is taken as the dynamical variable but the bare action  $\Gamma_{\Lambda,\Lambda}$  is not specified from the outset. In fact, conceptually it is largely determined by the requirement that a continuum limit exists (see the criterion in Section 2.2).  $\Gamma_{\Lambda,\Lambda}$  can be expected to have a well-defined derivative expansion with the leading terms roughly of the form (1.7). Also the gravitational effective average action  $\Gamma_{\Lambda,k}$  obeys an 'exact' FRGE, which is a new computational tool in quantum gravity not limited to perturbation theory. In practice  $\Gamma_{\Lambda,k}$  is replaced in this equation with a  $\Lambda$  independent functional interpreted as  $\Gamma_{\infty,k}$ . The assumption that the 'continuum limit'  $\Gamma_{\infty,k}$  for the gravitational effective average action exists is of course what is at stake here. The strategy in the FRGE approach is to show that this assumption, although without a-priori justification, is consistent with the solutions of the flow equation  $k \frac{d}{dk} \Gamma_{\infty,k} = \text{rhs}$  (where right-hand-side now also refers to the Hessian of  $\Gamma_{\infty,k}$ ). The structure of the solutions

$\Gamma_k$  of this cut-off independent FRGE should be such that they can plausibly be identified with  $\Gamma_{\infty,k}$ . Presupposing the ‘infrared safety’ in the above sense, a necessary condition for this is that the limits  $\lim_{k \rightarrow \infty} \Gamma_k$  and  $\lim_{k \rightarrow 0} \Gamma_k$  exist. Since  $k \leq \Lambda$  the first limit probes whether  $\Lambda$  can be made large; the second condition is needed to have all modes integrated out. In other words one asks for *global existence* of the  $\Gamma_k$  flow obtained by solving the cut-off independent FRGE. Being a functional differential equation the cutoff independent FRGE requires an initial condition, i.e. the specification of a functional  $\Gamma_{\text{initial}}$  which coincides with  $\Gamma_k$  at some scale  $k = k_{\text{initial}}$ . The point is that only for very special ‘fine tuned’ initial functionals  $\Gamma_{\text{initial}}$  will the associated solution of the cutoff independent FRGE exist globally [157]. The existence of the  $k \rightarrow \infty$  limit in this sense can be viewed as the counterpart of the UV renormalization problem, namely the determination of the unstable manifold associated with the fixed point  $\lim_{k \rightarrow \infty} \Gamma_k$ . We refer to Section 2.2 for a more detailed discussion of this issue.

In practice of course a nonlinear functional differential equation is very difficult to solve. To make the FRGE computationally useful the space of functionals is truncated typically to a finite-dimensional one of the form

$$\Gamma_k[\cdot] = \sum_{i=0}^N g_i(k) k^{d_i} I_i[\cdot], \quad (1.13)$$

where the  $I_i$  are given ‘well-chosen’ – local or nonlocal – functionals of the fields (among them the expectation value of the metric  $\langle g_{\alpha\beta} \rangle$  in the case of gravity) and the  $g_i(k)$  are numerical parameters that carry the scale dependence. For  $I_i$ ’s obeying a non-redundancy condition, the  $g_i$  play the role of essential couplings which have been normalized to have vanishing mass dimension by taking out a power  $k^{d_i}$ . The original FRGE then can be converted into a system of nonlinear ordinary differential equations for these couplings. In the case of gravity the following ansatz has been made [133, 131]:

$$I_0[g] = \int dx \sqrt{g}, \quad I_1[g] = \int dx \sqrt{g} R(g), \quad I_2[g] = \int dx \sqrt{g} R(g)^2, \quad (1.14)$$

where  $g = (g_{\alpha\beta})_{1 \leq \alpha, \beta \leq 4}$  is the metric and  $R(g)$  is the associated curvature scalar. The flow pattern  $k \mapsto (g_0(k), g_1(k), g_2(k))$  displays a number of remarkable properties. Most importantly a non-Gaussian fixed point exists (first found in [204] based on [179] and extensively corroborated in [205, 133, 131, 136, 39]). Within the truncation (1.14) a *three-dimensional* subset of initial data is attracted to the fixed point under the reversed flow

$$\lim_{k \rightarrow \infty} (g_0(k), g_1(k), g_2(k)) = (g_0^*, g_1^*, g_2^*), \quad (1.15)$$

where the fixed point couplings  $g_i^*$ ,  $i = 0, 1, 2$ , are finite and positive and no blow-up occurs in the flow for large  $k$ . Moreover unphysical propagating modes appear to be absent. Again this adheres precisely to the asymptotic safety criterion. Some of the trajectories with initial data in the unstable manifold cannot quite be extended to  $k \rightarrow 0$ , due to (infrared) singularities. This problem is familiar from nongravitational theories and is presumably an artifact of the truncation. In the vicinity of the fixed point, on the other hand, all trajectories show remarkable robustness properties against modifications of the mode cutoff scheme (see Section 4.3) which provide good reasons to believe that the structural aspects of the above results are not an artifact of the truncation used. The upshot is that there is a clear signal for asymptotic safety in the subsector (1.13), obtained via truncated functional renormalization flow equations.

The impact of matter has been studied by Percacci et al. [72, 171, 170]. Minimally coupling free fields (bosons, fermions, or Abelian gauge fields) one finds that the non-Gaussian fixed point is robust, but the positivity of the fixed point couplings  $g_0^* > 0$ ,  $g_1^* > 0$  puts certain constraints on the allowed number of copies. When a self-interacting scalar  $\chi$  is coupled non-minimally via

$\sqrt{g}[(\kappa_0 + \kappa_2\chi^2 + \kappa_4\chi^4 + \dots)R(g) + \lambda_0 + \lambda_2\chi^2 + \lambda_4\chi^4 + \dots + \partial\chi\partial\chi]$ , one finds a fixed point  $\kappa_0^* > 0$ ,  $\lambda_0^* > 0$  (whose values are with matched normalizations the same as  $g_1^*, g_0^*$  in the pure gravity computation) while all self-couplings vanish,  $\kappa_2^* = \kappa_4^* = \dots = 0$ ,  $\lambda_2^* = \lambda_4^* = \dots = 0$ . In the vicinity of the fixed point a linearized stability analysis can be performed; the admixture with  $\lambda_0$  and  $\kappa_0$  then lifts the marginality of  $\lambda_4$ , which becomes marginally irrelevant [171, 170]. The running of  $\kappa_0$  and  $\lambda_0$  is qualitatively unchanged as compared to pure gravity, indicating that the asymptotic safety property is robust also with respect to the inclusion of self-interacting scalars.

Both Strategies (c) and (d) involve truncations and one may ask to what extent the results are significant for the (intractable) full renormalization group dynamics. In our view they are significant. This is because even for the truncated problems there is no a-priori reason for the asymptotic safety property. In the Strategy (c) one would in the coupling space considered naively expect a zero-dimensional unstable manifold rather than the co-dimension zero one that is actually found! In Case (d) the ansatz (1.13, 1.14) implicitly replaces the full gravitational dynamics by one whose functional renormalization flow is confined to the subspace (1.13, 1.14) (similar to what happens in a hierarchical approximation). However there is again no a-priori reason why this approximate dynamics should have a non-Gaussian fixed point with positive fixed point couplings and with an unstable manifold of co-dimension zero. Both findings are genuinely surprising.

Nevertheless even surprises should have explanations in hindsight. For the asymptotic safety property of the truncated Quantum Gravidynamics in Strategies (c) and (d) the most natural explanation seems to be that it reflects the asymptotic safety of the *full* dynamics with respect to a nontrivial fixed point.

Tentatively both results are related by the dimensional reduction of the residual interactions in the ultraviolet. Alternatively one could try to merge both strategies as follows. One could take the background metrics in the background effective action generic and only impose the 2-Killing vector condition on the integration variables in the functional integral. Computationally this is much more difficult; however it would allow one to compare the lifted 4D flow with the one obtained from the truncated flows of the effective average action, presumably in truncations far more general than the ones used so far. A better way to relate both strategies would be by trying to construct a two-dimensional UV field theory with the characteristics to be described at the end of Section 2.4 and show its asymptotic safety.

### 1.3 Some working definitions

Here we attempt to give working definitions for some of the key terms used before.

**Quantum Gravidynamics:** We shall use the term “Quantum Gravidynamics” to highlight a number of points in the present circle of ideas. First, that one aims at relating the micro- and the macro-physics of the gravitational field through a renormalization flow defined conceptually in terms of a functional integral. In contrast to “Quantum General Relativity” the microscopic action is allowed to be very different from the Einstein–Hilbert action or a discretization thereof. Plausibly it should be still quasilocal, i.e. have a well-defined derivative expansion, and based on perturbatively renormalizable higher derivative theories one would expect it to contain at least quartic derivative terms. This means that also the number of physical propagating degrees of freedom (with respect to a background) may be different from the number entailed by the Einstein–Hilbert action. The second motivation for the term comes from the analogy with Quantum Chromodynamics. Indeed, the premise is that the self-interaction for the quantized gravitational field is predominantly “anti-screening” in the ultraviolet in a similar sense as in Quantum Chromodynamics, where it is responsible for the characteristic high energy behavior of physical quantities. As in Quantum Chromodynamics the proper identification of the antagonistic degrees of freedom (screening versus anti-screening) may well depend on the choice of field variables.

As with “Quantum General Relativity” we take the term “Gravidynamics” in a broad sense, allowing for any set of field variables (e.g. vielbein and spin connection, Ashtekar’s variables, Plebanski and BF type formulations, teleparallel, etc.) that can be used to recast general relativity (see e.g. the review [167]). For example the coupling of fermions might be a good reason to use a vielbein formulation. If the metric is taken as dynamical variable in four dimensions we shall also use the term “Quantum Einstein Gravity” as in [154, 133, 131]. It is of course *not* assumed from the outset that the quantum gravidynamics based on the various set of field variables are necessarily equivalent.

**Gaussian fixed point:** A fixed point is called *Gaussian* if there exists a choice of field variables for which the fixed point action is quadratic in the fields and the functional measure is Gaussian. This includes the local case but also allows for nonlocal quadratic actions. The drawback of this definition is that the proper choice of field variables in which the measure reveals its Gaussian nature may be hard to find. (For example in the correlation functions of the spin field in the two-dimensional Ising model the underlying free fermionic theory is not visible.)

A non-Gaussian fixed point is simply one where no choice of fields can be found in which the measure becomes Gaussian. Unfortunately this, too, is not a very operational criterion.

**Unstable manifold:** The unstable manifold of a fixed point with respect to a coarse graining operation is the set of all points that can be reached along flow lines emanating from the fixed point, the so-called *renormalized trajectories*. Points on such a flow line correspond to *perfect actions*. The *stable* manifold is the set of points attracted to the fixed point in the direction of coarse graining.

**Strict (weak) renormalizability:** We call a field theory strictly (weakly) renormalizable with respect to a fixed point and a coarse graining operation if the dimension of its unstable manifold is finite (infinite). It is implied that if a field theory has this property with respect to one coarse graining operation it will have it with respect to many others (“universality”). Strict or weak renormalizability is believed to be a sufficient condition for the existence of a genuine continuum limit for observables.

**Relevant coupling:** Given an expansion “sum over couplings times interaction monomials”, a coarse graining operation, and a fixed point of it, a coupling is called *relevant (irrelevant)* if it is driven away from (towards) the value the corresponding coordinate has at the fixed point, under a sufficient number of coarse graining steps. Note that this distinction makes sense even for trajectories not connected to the fixed point (because they terminate). It is however an explicitly ‘coordinate dependent’ notion. The same terms are used for the interaction monomials associated with the couplings. The dimension of the unstable manifold equals the maximal number of independent relevant interaction monomials ‘connected’ to the fixed point. All points on the unstable manifold are thus parameterized by relevant couplings but not vice versa.

Couplings which are relevant or irrelevant in a linearized analysis are called linearly relevant or linearly irrelevant, respectively. A coupling which is neither linearly relevant nor linearly irrelevant is called (linearly) marginal.

**Continuum limit:** By a genuine continuum limit we mean here a limit in which physical quantities become:



- (C1) strictly independent of the UV cutoff,
- (C2) independent of the choice of the coarse graining operation (within a certain class), and
- (C3) invariant under point transformations of the fields.

Usually one stipulates Properties (C1) and (C2) for the functional measure after which Property (C3) should be a provable property of physical quantities like the S-matrix. The requirement of having also Properties (C1) and (C2) only for observables is somewhat weaker and in the spirit of the asymptotic safety scenario.

Typically the Properties (C1, C2, C3) cannot be rigorously established, but there are useful criteria which render the existence of a genuine continuum limit plausible in different computational frameworks. In Sections 2.1 and 2.2 we discuss in some detail such criteria for the perturbative and for the FRGE approach, respectively. For convenience we summarize the main points here.

In *renormalized perturbation* theory the criterion involves two parts:

**(PTC1)** Existence of a formal continuum limit. This means, the removal of the UV cutoff is possible and the renormalized physical quantities are independent of the scheme and of the choice of interpolating fields – all in the sense of formal power series in the loop counting parameter. The perturbative beta functions always have a trivial (Gaussian) fixed-point but may also have a nontrivial (non-Gaussian) fixed point.

The second part of the criterion is:

**(PTC2)** The dimension of the unstable manifold of the (Gaussian or non-Gaussian) fixed point as computed from the perturbative beta functions equals the number of independent essential couplings.

For example  $\phi_4^4$  and QED meet Criterion (PTC1) but not (PTC2) while QCD satisfies both Criterion (PTC1) and (PTC2). In the framework of the *functional renormalization group equations* (FRGE) similar criteria for the existence of a genuine continuum limit can be formulated. Specifically for the FRGE of the effective average action one has:

**(FRGC1)** The solution of the FRG equation admits (for fine tuned initial data  $\Gamma_{\text{initial}}$  at some  $k = k_{\text{initial}}$ ) a global solution  $\Gamma_k$ , i.e. one that can be extended both to  $k \rightarrow \infty$  and to  $k \rightarrow 0$  (where the latter limit is not part of the UV problem in itself).

**(FRGC2)** The functional derivatives of  $\lim_{k \rightarrow 0} \Gamma_k$  (vertex functions) meet certain requirements which ensure stability/positivity/unitarity.

In Criterion (FRGC1) the existence of the  $k \rightarrow 0$  limit in theories with massless degrees of freedom is nontrivial and the problem of gaining computational control over the infrared physics should be separated from the UV aspects of the continuum limit as much as possible. However the  $k \rightarrow 0$  limit is essential to probe stability/positivity/unitarity. For example, to obtain a (massive) Euclidean quantum field theory the Schwinger functions constructed from the vertex functions have to obey nonlinear relations which ensure that the Hilbert space reconstructed via the Osterwalder–Schrader procedure has a positive definite inner product.

**Perturbative (weak) renormalizability:** We call a theory perturbatively (weakly) renormalizable if Criterion (PTC1) can be achieved with finitely (infinitely) many essential couplings. A theory were neither can be achieved is called perturbatively non-renormalizable. Perturbative (weak) renormalizability is neither necessary nor sufficient for (weak or strict) renormalizability in the above nonperturbative sense. It is only in combination with Criterion (PTC2) that perturbative results are indicative for the existence of a genuine continuum limit.

**Asymptotically free coupling:** A non-constant coupling in the unstable manifold of a Gaussian fixed point.

The “non-constant” proviso is needed to exclude cases like a trivial  $\phi_4^4$  coupling. In a nonperturbative lattice construction of  $\phi_4^4$  theory only a Gaussian fixed point with a one-dimensional unstable manifold (parameterized by the renormalized mass) is thought to exist, along which the renormalized  $\phi_4^4$  coupling is constant and identically zero. The Gaussian nature of the fixed-point, on the other hand, is not crucial and we define:

**Asymptotically safe coupling:** A non-constant coupling in the unstable manifold of a fixed point.

**Asymptotically safe functional measure:** The functional measure of a statistical field theory is said to be asymptotically safe if it is perturbatively weakly renormalizable or nonrenormalizable, but possesses a fixed point with respect to which it is strictly renormalizable. Subject to the regularity assumption that the space of actions can in the vicinity of the fixed point be decomposed into a stable and an unstable manifold, this is equivalent to the following requirement: All relevant couplings are asymptotically safe and there is only a finite number of them. Note that unitarity or other desirable properties that would manifest itself on the level of observables are not part of this definition.

In a non-gravitational context the functional measure of the 3D Gross–Neveu model is presently the best candidate to be asymptotically safe in the above sense (see [101, 60, 198, 105] and references therein). Also 5D Yang–Mills theories (see [93, 148] and references therein) are believed to provide examples. In a gravitational context, however, there are good reasons to modify this definition.

First the choice of couplings has to be physically motivated, which requires to make contact to observables. In the above nongravitational examples with a single coupling the ‘meaning’ of the coupling is obvious; in particular it is clear that it must be finite and positive at the non-Gaussian fixed point. In general however one does not know whether ill behaved couplings are perverse redefinitions of better behaved ones. To avoid this problem the couplings should be defined as coefficients in a power series expansion of the observables themselves (Weinberg’s “reaction rates”; see the discussion in Section 1.1). Of course painfully little is known about (generic) quantum gravity observables, but as a matter of principle this is how couplings should be defined. In particular this will pin down the physically adequate notion of positivity or unitarity.

Second, there may be good reasons to work initially with infinitely many essential or potentially relevant couplings. Recall that the number of essential couplings entering the initial construction of the functional measure is not necessarily equal to the number eventually indispensable. In a secondary step a reduction of couplings might be feasible. That is, relations among the couplings might exist which are compatible with the renormalization flow. If these relations are sufficiently complicated, it might be better to impose them retroactively than to try to switch to a more adapted basis of interaction monomials from the beginning.

Specifically in the context of quantum gravity microscopic actions with infinitely many essential couplings occur naturally in several ways. First, when starting from the Gomis–Weinberg picture [94] of perturbative quantum gravity (which is implemented in a non-graviton expansion in Section 3 for the 2 + 2 reduction). Second, when power counting considerations are taken as a guideline one can use Newton’s constant (frozen in Planck units) to build dimensionless scalars (dilaton, conformal factor) and change the conformal frame arbitrarily. The way how these dimensionless scalars enter the (bare versus renormalized) action is not constrained by power counting considerations. This opens the door to an infinite number of



essential couplings. The effective action for the conformal factor [149] and the dilaton field in the  $2 + 2$  reduction [154] provide examples of this phenomenon.

Third, the dimension of the unstable manifold is of secondary importance in this context. Recall that the dimension of the unstable manifold is the *maximal* number of independent relevant interaction monomials ‘connected’ to the fixed point. This maximal number may be very difficult to determine in Quantum Gravidynamics. It would require identification of *all* renormalized trajectories emanating from the fixed point – which may be more than what is needed physicswise: The successful construction of a subset of renormalized trajectories for physically motivated couplings may already be enough to obtain predictions/explanations for some observables. Moreover, what matters is not the total number of relevant couplings but the way how observables depend on them. Since generic observables (in the sense used in Section 1.1) are likely to be nonlinearly and nonlocally related to the metric or to the usual basis of interaction monomials (scalars built from polynomials in the curvature tensors, for instance) the condition that the theory should allow for predictions in terms of observables is only indirectly related to the total number of relevant couplings.

In summary, the interplay between the microscopic action, its parameterization through essential or relevant couplings, and observables is considerably more subtle than in the presumed non-gravitational examples of asymptotically safe theories with a single coupling. The existence of an asymptotically safe functional measure in the above sense seems to be neither necessary nor sufficient for a physically viable theory of Quantum Gravidynamics. This leads to our final working definition.

**Asymptotically safe Quantum Gravidynamics:** A quantum theory of gravity based on a notion of an asymptotically safe functional integral measure which incorporates the interplay between couplings and observables described above. In brief:

- The choice of couplings has to be based on observables; this will pin down the physically relevant notion of positivity/unitarity.
- The number of essential or relevant couplings is not a-priori finite.
- What matters is not so much the dimension of the unstable manifold than how observables depend on the relevant couplings.

## 1.4 Relation to other approaches

For orientation we offer here some sketchy remarks on how Quantum Gravidynamics relates to some other approaches to Quantum Gravity, notably the Dynamical Triangulations approach, Loop Quantum Gravity, and String Theory. These remarks are of course not intended to provide a comprehensive discussion of the relative merits but merely to highlight points of contact and stark differences to Quantum Gravidynamics.

### 1.4.1 Dynamical triangulations

The framework closest in spirit to the present one are discretized approaches to the gravitational functional integral, where a continuum limit in the statistical field theory sense is aimed at. See [138] for a general review and [8] for the dynamical triangulations approach. Arguably the most promising variant of the latter is the causal dynamical triangulations approach by Ambjørn, Jurkiewicz, and Loll [5]. In this setting the formal four dimensional quantum gravity functional integral is replaced by a sum over discrete geometries,  $Z = \sum_T m(T) e^{iS(T)}$ . The geometries  $T$  are piecewise Minkowskian and selected such that they admit a Wick rotation to piecewise Euclidean geometries. The edge lengths in the spatial and the temporal directions are  $\ell_{\text{space}} = a^2$

and  $\ell_{\text{time}} = -\alpha a^2$ , where  $a$  sets the discretization scale and  $\alpha > 0$  is an adjustable parameter. The flip  $\alpha \mapsto -\alpha$  defines a Wick rotation under which the weights in the partition function become real:  $e^{iS(T)} \mapsto e^{-S_{\text{eucl}}(T)}$ . For  $\alpha = -1$  the usual expressions for the action used [8] in equilateral Euclidean dynamical triangulations are recovered, but the sum is only over those Euclidean triangulations  $T$  which lie in the image of the above Wick rotation. The weight factor  $m(T)$  is the inverse of the order of the automorphism group of the triangulation, i.e. 1 for almost all of them. With these specifications the goal is to construct a continuum limit by sending  $a \rightarrow 0$  and the number  $N$  of simplices to infinity, while adjusting the two bare parameters (corresponding to Newton's constant and a cosmological constant) in  $S_{\text{eucl}}(T)$  as well as the overall scale of  $Z$ . Very likely, in order for such a continuum limit to exist and to be insensitive against modifications of the discretized setting, a renormalization group fixed point in the coupling flow is needed. Assuming that the system indeed has a fixed point, this fixed point would by construction have a nontrivial unstable manifold, and ideally both couplings would be asymptotically safe, thereby realizing the strong asymptotic safety scenario (using the terminology of Section 1.3). Consistent with this picture and the previously described dimensional reduction phenomenon for asymptotically safe functional measures (see also Section 2.4), the microscopic geometries appear to be effectively two-dimensional [6, 7].

Despite these similarities there are (for the time being) also important differences. First the discretized action depends on two parameters only and it is hoped that a renormalized trajectory can be found by tuning only these two parameters. Since in dynamical triangulations there is no naive (classical) continuum limit, one cannot directly compare the discretized action used with a microscopic action in the previous sense. Conceptually one can assign a microscopic action to the two parametric measure defined by the causal dynamical triangulations by requiring that combined with the regularized kinematical continuum measure (see Section 2.3.3) it reproduces the same correlation functions in the continuum limit. The microscopic action defined that way would presumably be different from the Einstein–Hilbert action, but it would still contain only two tunable parameters. In other words the hope is that the particular non-naive discretization procedure gets all but two coordinates of the unstable manifold automatically right. A second difference concerns the role of averages of the metric. The transfer matrix used in [8] is presumed to have a unique ground state for both finite and infinite triangulations. Expectation values in a reconstructed Hilbert space will refer to this ground state and hence be unique for a given operator. A microscopic metric operator does not exist in a dynamical triangulations approach but if one were to define coarse grained variants, their expectation value would have to be unique. In contrast the field theoretical formulations based on a background effective action allow for a large class of averaged metrics.

### 1.4.2 Loop quantum gravity

The term loop quantum gravity is by now used for a number of interrelated formulations (see [199] for a guide). For definiteness we confine our comparative remarks to the original canonical formulation using loop (holonomy) variables.

Here a reformulation of general relativity in terms of Ashtekar variables  $(A, E)$  is taken as a starting point, where schematically  $A$  and  $E$  are defined on a three-dimensional time slice and are conjugate to each other,  $\{A, E\} = \delta$ , with respect to the canonical symplectic structure (see [199, 15]). From  $A$  one can form holonomies (line integrals along loops) and from  $E$  one can form fluxes (integrals over two-dimensional hypersurfaces) without using more than the manifold structure. The Poisson bracket  $\{A, E\} = \delta$  is converted into a Poisson algebra for the holonomy and the flux variables. Two basic assumptions then govern the transition to the quantum theory: First the Poisson bracket  $\{A, E\} = \delta$  is replaced by a commutator  $[A, E] = i\hbar\delta$  and is subsequently converted into an algebraic structure among the holonomy and flux variables. Second, representations of

this algebra are sought on a state space built from multiple products of holonomies associated with a graph (spin network states). The inner product on this space is sensitive only to the coincidence or non-coincidence of the graphs labeling the states (not to their embedding into the three-manifold). Based on a Gelfand triple associated with this kinematical state space one then aims at the incorporation of dynamics via a (weak) solution of the Hamiltonian constraint of general relativity (or a ‘squared’ variant thereof). To this end one has to transplant the constraint into holonomy and flux variables so that it can act on the above state space. This step is technically difficult and the results obtained do not allow one to address the off-shell closure of the constraint algebra, an essential requirement emphasized in [151].

As far as comparison with Quantum Gravidynamics is concerned, important differences occur both on a kinematical and on a dynamical level, even if a variant of Gravidynamics formulated in terms of the Ashtekar variables  $(A, E)$  was used [156]. Step one in the above quantization procedure keeps the right-hand-side of the commutator  $[A, E] = i\hbar\delta$  free from dynamical information. In any field theoretical framework, on the other hand, one would expect the right-hand-side to be modified: minimally (if  $A$  and  $E$  are multiplicatively renormalized) by multiplication with a (divergent) wave function renormalization constant, or (if  $A$  and  $E$  are nonlinearly renormalized) by having  $\delta$  replaced with a more general, possibly field dependent, distribution. Stipulation of unmodified canonical commutation relations might put severe constraints on the allowed interactions, as it does in quantum field theories with a sufficiently soft ultraviolet behavior. (We have in mind here “triviality” results, where e.g. for scalar quantum field theories in dimensions  $d \geq 4$  a finite wave function renormalization constant goes hand in hand with the absence of interaction [24, 77]). A second marked difference to Quantum Gravidynamics is that in Loop Quantum Gravity there appears to be no room for the distinction between fine grained (‘rough’) and coarse grained (‘smooth’) geometries. The inner product used in the second of the above steps sees only whether the graphs of two spin network states coincide or not, but is insensitive to the ‘roughness’ of the geometry encoded initially in the  $(A, E)$  pair. This information appears to be lost [151]. In a field theory the geometries would be sampled according to some underlying measure and the typical configurations are very rough (non-differentiable). As long as the above ‘holonomy inner product’ on such sampled geometries is well defined and depends only on the coincidence or non-coincidence of the graphs the information about the measure according to which the sampling is done appears to be lost. Every measure will look the same. This property seems to match the existence of a preferred diffeomorphism invariant measure [14] (on a space generated by the holonomies) which is uniquely determined by some natural requirements. The typical  $A$  configurations are also of distributional type [14, 140]. This uniqueness translates into the uniqueness of the associated representation of the holonomy-flux algebra (which rephrases the content of the original  $[A, E] = i\hbar\delta$  algebra). In a field theory based on the  $(A, E)$  variables, on the other hand, there would be a cone of regularized measures which incorporate dynamical information and on which the renormalization group acts.

Another difference concerns the interplay between the dynamics and the canonical commutation relations. In a field theory the moral from Haag’s theorem is that “the choice of the representation of the canonical commutation relations is a dynamical problem” [99]. Further the inability to pick the ‘good’ representation beforehand is one way to look at the origin of the divergencies in a canonically quantized relativistic field theory. (To a certain extent the implications of Haag’s theorem can be avoided by considering scattering states and spatially cutoff interactions; in a quantum gravity context, however, it is unclear what this would amount to.) In contrast, in the above holonomy setting a preferred representation of the holonomy-flux algebra is uniquely determined by a set of natural requirements which do not refer to the dynamics. The dynamics formulated in terms of the Hamiltonian constraint thus must be automatically well-defined on the above kinematical arena (see [151, 173] for a discussion of the ambiguities in such constructions). In a field theoretical framework, on the other hand, the constraints would be defined as composite

operators in a way that explicitly requires dynamical information (fed in through the renormalized action). So the constraints and the space on which they act are dynamically correlated. In loop quantum gravity, in contrast, both aspects are decoupled.

Finally, the microscopic action for asymptotically safe Quantum Gravidynamics is very likely different from the Einstein–Hilbert action and thus not of second order. This changes the perspective on a canonical formulation considerably.

### 1.4.3 String theory

String theory provides a possible context for the unification of known and unknown forces including quantum gravity [97, 177, 112]. As far as quantum gravity is concerned the point of departure is the presupposition that the renormalization problem for the quantized gravitational field is both insoluble and irrelevant. Presently a clearly defined dynamical principle that could serve as a substitute seems to be available only for so-called perturbative first quantized string theory, to which we therefore confine the following comparative comments.

In this setting certain two-dimensional (supersymmetric) conformal field theories are believed to capture (some of) the ‘ultimate degrees of freedom of Nature’. The attribute ‘perturbative’ mostly refers to the fact that a functional integral over the two-surfaces on which the theories are defined is meant to be performed, too, but in a genus expansion this gives rise to a divergent and not Borel summable series. (In a non-perturbative formulation aimed at degrees of freedom corresponding to other extended objects are meant to occur and to cure this problem.) For the relation to gravity it is mainly the bosonic part of the conformal field theories which is relevant, so we take the 2D fermions to be implicitly present in the following without displaying them.

A loose relation to a gravitational functional integral then can be set up as follows. Schematically, the so-called low energy effective action  $S_{\text{eff}}[g]$  arises by functional integration over the fields of a Riemannian sigma-model  $X : 2\text{D surface} \rightarrow 10\text{D target space}$  with metric  $g_{\alpha\beta}$ . Thus  $\int dX e^{-S_g[X]} = \exp\{-S_{\text{eff}}[g]\}$  is a functional integral that depends parametrically on the metric  $g_{\alpha\beta}$  which is viewed as a set of generalized couplings. The functional integral one (morally speaking) would like to make sense of in the present context is however

$$\int Dg \int DX e^{-S_g[X]} = \int Dg e^{-S_{\text{eff}}[g]}. \quad (1.16)$$

From the present point of view the ‘low energy’ effective action thus plays the role of a ‘Planck-scale’ or microscopic action, and the sigma-model approach makes a specific proposal for this action. (Alternatively one can reconstruct  $S_{\text{eff}}[g]$  by studying string scattering amplitudes in a flat target space  $g_{\alpha\beta} = \eta_{\alpha\beta}$ .) Performing the additional functional integral over the metrics loosely speaking corresponds to some string field theory. However as far as the functional integral on the right-hand-side is concerned the string field only serves as an ‘auxiliary field’ needed to come up with the proposal for the action (whatever name one gives it) used to ‘weigh’ the metrics. In contrast, in the string paradigm, it is primarily the degrees of freedom  $X$  of the Riemannian sigma-model one should quantize (i.e. perform a functional integral over), not the degrees freedom in the gravitational field represented by the (equivalence classes of the) metric field (or any other set of fields classically equivalent to it). In this sense the degrees of freedom gravitational field are not taken seriously in the quantum regime, as the fundamental degrees of freedom (extended objects) are supposed to be known. The point we are trying to make may become clearer if one considers superstring theory, where the effective action  $S_{\text{eff}}$  also includes Yang–Mills fields. After compactification to four dimensions one still has to perform a functional integral over the 4D Yang–Mills field, in order to make contact to fully-fledged QCD.

Since arguments presented after Equation (1.5) suggest a kind of ‘dimensional reduction’ to  $d = 2$ , one might be tempted to see this as a vindication of string theory from the present viewpoint. However string theory’s very departure was the presupposition that no fixed point exists

for the gravitational functional integral. Moreover in string theory the sigma-model fields relate the worldsheet to a (4 + 6-dimensional) target manifold with a prescribed metric (or pairs thereof related by T-duality). The at least perturbatively known dynamics of the sigma-model fields does not appear to simulate the functional integral over metrics (see Equation (1.16)). The additional functional integral over Euclidean worldsheet geometries is problematic in itself and leaves unanswered the question how and why it successfully captures or replaces the ultraviolet aspects of the original functional integral, other than by definition. In the context of the asymptotic safety scenario, on the other hand, the presumed reduction to effectively two-dimensional propagating degrees of freedom is a consequence of the renormalization group dynamics, which in this case acts like a ‘holographic map’. This holographic map is of course not explicitly known, nor is it off-hand likely that it can be described by some effective string theory. A more immediate difference is that Quantum Gravidynamics does not require the introduction of hitherto unseen degrees of freedom.

## 1.5 Discussion of possible objections

Here we discuss some of the possible objections to a physically viable asymptotically safe theory of quantum gravidynamics.

**Q1** Since the microscopic action is likely to contain higher derivative terms, don’t the problems with non-unitarity notorious in higher derivative gravity theories reappear?

**A1** In brief, the unitarity issue has not much been investigated so far, but the presumed answer is No.

First, the problems with perturbatively strictly renormalizable higher derivative theories stem mostly from the  $1/p^4$ -type propagator used. The alternative perturbative framework already mentioned, namely to use a  $1/p^2$ -type propagator at the expense of infinitely many essential (potentially ‘unsafe’) couplings, avoids this problem [94, 10]. The example of the  $2 + 2$  reduction shows that the reconciliation of safe couplings with the absence of unphysical propagating modes can be achieved in principle. Also the superrenormalizable gravity theories with unitary propagators proposed in [218] are intriguing in this respect.

Second, when the background effective action is used as the central object to define the quantum theory, the ‘background’ is *not* a solution of the classical field equations. Rather it is adjusted self-consistently by a condition involving the full quantum effective action (see Appendix B). If the background effective action is computed nonperturbatively (by whatever technique) the intrinsic notion of unitarity will not be related to the ‘propagator unitarity’ around a solution of the classical field equations in any simple way.

One aspect of this intrinsic positivity is the convexity of the background effective action. In the flow equation for the effective average action one can see, for example, that the wrong-sign of the propagator is not an issue: If  $\Gamma_k$  is of the  $R + R^2$  type, the running inverse propagator  $\Gamma_k^{(2)}$  when expanded around flat space has ghosts similar to those in perturbation theory. For the FRG flow, however, this is irrelevant since in the derivation of the beta functions no background needs to be specified explicitly. All one needs is that the RG trajectories are well defined down to  $k = 0$ . This requires that  $\Gamma_k^{(2)} + \mathcal{R}_k$  is a positive operator for all  $k$ . In the untruncated functional flow this is believed to be the case. A rather encouraging first result in this direction comes from the  $R^2$  truncation [131].

More generally, the reservations towards higher derivative theories came from a loop expansion near the perturbative Gaussian fixed point. In contrast in Quantum Gravidynamics one aims at constructing the continuum limit nonperturbatively at a different fixed point. In the conventional setting one quantizes  $R + R^2$  as the bare action, while in Quantum Gravidynamics the bare action, defined by backtracing the renormalized trajectory to the non-Gaussian

fixed point, may in principle contain all sorts of curvature invariants whose impact on the positivity and causality of the theory is not even known in perturbation theory.

In the previous discussion we implicitly assumed that generic physical quantities are related in a rather simple way to the interaction monomials entering the microscopic action. For Dirac observables however this is clearly not the case. Assuming that the physically correct notion of unitarity concerns such observables it is clear that the final word on unitarity issues can only be spoken once actual observables are understood.

**Q2** Doesn't the very notion of renormalizability presuppose a length or momentum scale? Since the latter is absent in a background independent formulation, the renormalizability issue is really an artifact of the perturbative expansion around a background.

**A2** No. Background independence is a subtle property of classical general relativity (see e.g. [199] for a discussion) for which it is unclear whether or not it has a compelling quantum counterpart. As far as the renormalization problem is concerned it is part of the physics premise of a functional integral type approach that there is a description independent and physically relevant distinction between coarse grained and fine grained geometries. On a classical level this amounts to the distinction, for example, between a perfect fluid solution of the field equations and one generated by its  $10^{30}$  or so molecular constituents. A sufficiently large set of Dirac observables would be able to discriminate two such spacetimes. Whenever we shall refer later on to "coarse grained" versus "fine grained" geometries we have a similar picture in mind for the ensembles of off-shell geometries entering a functional integral.

Once such a physics premise is made, the renormalization in the Kadanoff–Wilson sense is clearly relevant for the computation of observable quantities and does not just amount to a reshuffling of artifacts. Renormalization in this sense is, for example, very likely not related to the regularization ambiguities [151, 173] appearing in loop quantum gravity. A minimal requirement for such an interpretation of the regularization ambiguities would be that reasonable coarse graining operations exist which have a preferred discretization of the Einstein–Hilbert action as its fixed point. This preferred discretization would have to be such that the observables weakly commuting with the associated Hamiltonian constraint reproduce those of loop quantum gravity.

For clarity's sake let us add that the geometries entering a functional integral are expected to be very rough on the cutoff scale (or of a distributional type without a cutoff) but superimposed to this 'short wavelength zigzag' should be 'long wavelength' modulations (defined in terms of dimensionless ratios) to which different observables are sensitive in different degrees. In general it will be impractical to base the distinction between ensembles of fine grained and coarse grained geometries directly on observables. In the background field formalism the distinction is made with respect to an initially prescribed but generic background geometry which after the functional integral is performed (entirely or in a certain mode range) gets related to the expectation value of the quantum metric by a consistency condition involving the full quantum dynamics.

**Q3** Doesn't such a non-perturbative renormalizability scenario require a hidden enhanced symmetry?

**A3** Improved renormalizability properties around a given fixed point are indeed typically rooted in symmetries. A good example is QCD in a lightfront formulation where gauge invariance is an 'emergent phenomenon' occurring only after an infinite reduction of couplings [174]. In the case of Quantum Gravidynamics, the symmetry in question would be one that becomes visible only around the non-Gaussian fixed point. If it exists, its identification would constitute



a breakthrough. From the Kadanoff–Wilson view of renormalization it is however the fixed point which is fundamental – the enhanced symmetry properties are a consequence (see the notion of generalized symmetries in [236, 160]).

**Q4** Shouldn't the proposed anti-screening be seen in perturbation theory?

**A4** Maybe, maybe not. Presently no good criterion for antiscreening in this context is known. For the reasons explained in Section 1.1 it should not merely be identified with the sign of the dominant contribution to some beta function. The answer to the above question will thus depend somewhat on the identification of the proper degrees of freedom and the quantity considered.

As an example one can look at quantum gravity corrections to the Newton potential, which have been considered in some detail. The result is always of the form

$$V(r) = -\frac{Gm_1m_2}{r} \left[ 1 + 3\frac{G(m_1+m_2)}{r} + \zeta\frac{\hbar G}{r^2} \right].$$

Interpreted as a modification of Newton's constant  $G(r) = G(1 + \zeta\hbar G/r^2)$ , one sees that  $\zeta > 0$  roughly corresponds to screening and  $\zeta < 0$  to anti-screening behavior. The value of  $\zeta$  is unambiguously defined in 1-loop perturbation theory and is a genuine prediction of quantum gravity viewed as an effective field theory (as stressed by Donoghue). However  $\zeta$  will depend on the precise definition of the nonrelativistic potential and there are various options for it.

One is via the  $2 \rightarrow 2$  scattering amplitude. The coefficient  $\zeta_{\text{scatt}}$  was computed initially by Donoghue and later by Khriplovich and Kirilin; the result considered definite in [32] is  $\zeta_{\text{scatt}} = \frac{41}{10\pi}$ . It decomposes into a negative vertex and triangle contributions  $\zeta_v = -\frac{105}{3\pi}$ , and a just slightly larger positive remainder  $\zeta_{\text{scatt}} - \zeta_v = \frac{117.3}{3\pi}$  coming from box, seagull, and vacuum polarization diagrams.

Another option is to consider corrections to the Schwarzschild metric. Different sets of diagrams have been used for the definition [119, 33] and affect the parameterization (in-)dependence and other properties of the corrections. Both choices advocated lead to  $\zeta_{\text{metric}} < 0$ , which amounts to antiscreening.

Let us also mention alternative definitions of an effective Newton potential via Wilson lines in Regge calculus [100] or by resummation of scalar matter loops [226]. The latter gives rise to an “antiscreening” Yukawa type correction of the form  $V(r) = -\frac{G}{4\pi r}(1 - e^{-r/\sqrt{\zeta G}})$ , with  $\zeta > 0$ . Via  $V(r) = \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}} G(k)/k^2$  it can be interpreted as a running Newton constant  $G(k) = G/(1 + \zeta k^2)$ .

**Q5** There are several thought experiments suggesting a fundamental limit to giving an operational meaning to spacetime resolutions, for example via generalized uncertainty relations of the form (see [85, 163, 143] and references therein)

$$\Delta x \approx \frac{\hbar}{\Delta p} + \frac{G\hbar}{c^3} \frac{\Delta p}{\hbar}.$$

These relations are sometimes taken as hinting at a “fundamental discretum”. If so, does this not contradict the asymptotic safety scenario, where in the fixed point regime the microscopic spacetimes become self-similar?

**A5** No, the arguments assume that Newton's constant  $G$  is constant. (We momentarily write  $G$  for  $G_N$  in  $(16\pi G_N)^{-1} \int dx \sqrt{g} R(g)$ .) If  $G$  is treated as a running coupling the derivations of

the uncertainty relations break down. As an example consider a photon-electron scattering process as in [143, 163]:  $G$  refers to gravity in the ('photon'  $k$  – 'electron'  $\Delta p$ ) interaction region with a pointlike 'electron'. If viewed as running one expects  $G(k) \approx G_*/k^2$  in the fixed point regime. Hence in the above relation one should replace  $G$  by  $\frac{G_* c^3}{(\Delta p)^2}$ . This gives

$$\Delta x \Delta p \approx \hbar(1 + G_*),$$

and there is no limit on the spatial resolution. One can of course decide to choose units in which  $G$  is constant by definition (see [172]) in which case the derivations go through. Our conclusion is that the perceived dichotomy between a fundamentally 'discrete' versus 'continuum' geometry may itself not be fundamental.

Each of the issues raised clearly deserves much further investigation. For the time being we conclude however that the asymptotic safety scenario is conceptually self-consistent. It remains to assemble hard computational evidence for the existence of a non-Gaussian fixed point with a nontrivial and regular unstable manifold. This task will be taken up in Sections 3 and 4.



## 2 Renormalizing the Non-Renormalizable

The modern view on renormalization has been shaped by Kadanoff and Wilson. The more familiar perturbative notion of renormalizability is neither sufficient (e.g.  $\phi^4$  theory in  $d = 4$ ) nor necessary (e.g. Gross–Neveu model in  $d = 3$ ) for renormalizability in the Kadanoff–Wilson sense. For the convenience of the reader we summarize the main ideas in Appendix A, which also serves to introduce the terminology. The title of this section is borrowed from a paper by Gawedzki and Kupiainen [88].

In the present context the relevance of a Kadanoff–Wilson view on renormalization is two-fold: First it allows one to formulate the notion of renormalizability without reference to perturbation theory, and second it allows one to treat to a certain extent renormalizable and non-renormalizable on the same footing. The mismatch between the perturbative non-renormalizability of the Einstein–Hilbert action and the presumed asymptotic safety of a functional measure constructed by other means can thus be systematically explored.

In a gravitational context also the significance of renormalizability is less clear cut, and one should presumably go back to the even more fundamental property for which renormalizability is believed to be instrumental, namely the existence of a genuine continuum limit, roughly in the sense outlined in Section 1.3. Since rigorous results based on controlled approximations are unlikely to be obtained in the near future, we describe in the following criteria for the plausible existence of a genuine continuum limit based on two uncontrolled approximations: renormalized perturbation theory and the functional renormalization group approach. Such criteria are ‘implicit wisdom’ and are hardly ever spelled out. In the context of Quantum Gravidynamics, however, the absence of an obvious counterpart of the correlation length and non-renormalizability of the Einstein–Hilbert action makes things more subtle. In Sections 2.1 and 2.2 we therefore try to make the implicit explicit and to formulate criteria for the existence of a genuine continuum limit which are applicable to Quantum Gravidynamics as well.

In Section 2.3 we describe the renormalization problem for Quantum Gravidynamics and in Section 2.4 the dimensional reduction phenomenon outlined before.

For a summary of basic renormalization group concepts we refer to Appendix A and for a review of the renormalization group for the effective average action to Appendix C.

### 2.1 Perturbation theory and continuum limit

Perturbatively renormalizable field theories are a degenerate special case of the Wilson–Kadanoff framework. The main advantage of perturbation theory is that the UV cutoff  $\Lambda$  can be removed exactly and independently of the properties of the coupling flow. The existence of a  $\Lambda \rightarrow \infty$  limit with the required Properties (PTC1) can often be rigorously proven, in contrast to most non-perturbative techniques where this can only be established approximately by assembling evidence. With Criterion (PTC1) satisfied, the coupling flow then can be studied in a second step and used to probe whether or not the Criterion (PTC2) for the existence of a genuine continuum limit as anticipated in Section 1.3 is also satisfied. The main disadvantage of perturbation theory is that everything is initially defined as a formal power series in the loop counting parameter. Even if one trades the latter for a running coupling, the series in this coupling remains a formal one, typically non-convergent and not Borel-summable. It is generally believed, however, that provided Criterion (PTC2) is satisfied for a perturbative Gaussian fixed point, the series is asymptotic to the (usually unknown) exact result. In this case the perturbative analysis should indicate the existence of a genuine continuum limit based on an underlying Gaussian fixed point proper. Our main reason for going through this in some detail below is to point out that in a situation with several couplings the very same rationale applies if the perturbative fixed point is a non-Gaussian rather than a Gaussian one.

As mentioned, in perturbation theory one initially only aims at defining the expectations (A.1) as a formal power series in the loop counting parameter  $\lambda$ , where the sum of all  $\ell$ -loop contributions to a quantity is assigned a factor  $\lambda^\ell$ . For the reasons explained in [106] the loop expansion does not necessarily coincide with an expansion in powers of Planck's constant  $\hbar$ . For example when massless fields are involved, 1-loop diagrams can contribute to the classical limit  $O(\hbar^0)$ . The loop counting parameter  $\lambda$  refers to a set of free fields of mass  $\mu$  such that the formal expansion of the exponential in Equation (A.1) gives expectations whose computation can be reduced to the evaluation of Gaussians. We denote the (quadratic) action of this set of free fields by  $S_{*,\mu}[\chi]$ . The interaction is described by a set of monomials  $P_i[\chi]$ ,  $i \in E_{p.c.}$ , which are ‘‘power counting renormalizable’’. The latter means that their mass dimension  $-d_i$  is such that  $d_i \geq 0$ . It is also assumed that the  $P_i$  are functionally independent, so that the corresponding couplings are essential. The so-called ‘‘bare’’ action functional then is  $S_\Lambda = S_{*,\mu} + \sum_{i \in E_{p.c.}} u_i(\Lambda) P_i$ , where  $u_i(\Lambda)$  are the essential ‘‘bare’’ couplings (including masses) corresponding to the interaction monomials  $P_i[\chi]$ . Inessential parameters are generated by subjecting  $S_\Lambda$  to a suitable class of field redefinitions. In more detail one writes

$$\begin{aligned} u_i(\Lambda) &= u_i(\mu) V_{i,0}(\mu) + \sum_{\ell \geq 1} \lambda^\ell V_{i,\ell}(u(\mu), \Lambda, \mu), \\ \chi_\Lambda &= \chi_\mu + \sum_{\ell \geq 1} \lambda^\ell \Xi_\ell(\chi_\mu; u(\mu), \Lambda, \mu). \end{aligned} \quad (2.1)$$

Here  $u_i(\mu)$  are the renormalized couplings which are  $\Lambda$ -independent and the  $V_i(u(\mu), \Lambda, \mu)$  are counterterms which diverge in the limit  $\Lambda \rightarrow \infty$ . This divergence is enforced by very general properties of QFTs. Similarly the  $\chi_\mu$  are called renormalized fields and the  $\Xi_\ell(\chi_\mu; u(\mu), \Lambda, \mu)$  are local functionals of the  $\chi_\mu$  with coefficients depending on  $u(\mu), \Lambda, \mu$ ; the coefficients again diverge in the limit  $\Lambda \rightarrow \infty$ . Often one aims at ‘‘multiplicative renormalizability’’, which means the ansatz for the  $\Xi_\ell$  is taken to be linear in the fields  $\Xi_\ell(\chi_\mu; u(\mu), \Lambda, \mu) = Z_\ell(u(\mu), \Lambda, \mu) \chi_\mu$  and  $Z_\ell$  is the  $\ell$ -loop ‘‘wave function renormalization’’ constant. One should emphasize, however, that multiplicative renormalizability can often not be achieved, and even in field theories where it can be achieved, it evidently will work only with a particular choice of field coordinates (see [41] for a discussion).

The normalizations in Equation (2.1) can be chosen such that  $u_i(\mu = \Lambda) = u_i(\Lambda)$  and  $\chi_{\mu=\Lambda} = \chi_\Lambda$ , but one is really interested in the regime where  $\mu \ll \Lambda$ . Inserting these parameterizations into  $S_\Lambda[\chi_\Lambda]$  gives an expression of the form

$$S_\Lambda[\chi_\Lambda] = S_{*,\mu}[\chi_\mu] + \sum_\alpha \left( \sum_{\ell \geq 0} \lambda^\ell u_{\alpha,\ell}(u(\mu), \Lambda, \mu) \right) P_\alpha[\chi_\mu], \quad (2.2)$$

where the sum over  $\alpha$  includes terms of the form appearing on the right-hand-side of Equation (A.8). Often the  $\mu$ -dependence in the fields can be traded for one carried by (inessential) parameters  $z_i(\mu)$ ,  $i \in I$ . Then Equation (2.2) takes the form  $S_\Lambda[\chi_\Lambda] = \sum_{\alpha'} u_{\alpha'}(g(\mu), z(\mu), \Lambda, \mu) P_{\alpha'}[\chi]$ , with some  $\mu$ -independent fields,  $\chi = \chi_{\mu_0}$ , say. The right-hand-side clearly resembles Equation (A.7) with the difference that modulo field redefinitions only power counting renormalizable interaction monomials occur.

So far the counterterms in Equation (2.1) have been left unspecified. The point of introducing them is of course as a means to absorb the cut-off dependence generated by the regularized functional integral in Equation (A.1). Specifically, one replaces the Boltzmann factor by its power series expansion in  $\lambda$ , i.e.  $\exp\{-S_\Lambda[\chi_\Lambda]\} = \exp\{-S_{*,\mu}[\chi_\mu]\} (1 + \sum_{\ell \geq 1} \lambda^\ell Q_\ell[\chi_\mu])$ , and aims at an evaluation of multipoint functions  $\langle \chi_\Lambda(x_1) \dots \chi_\Lambda(x_n) \rangle_{S_\Lambda}$  as formal power series in  $\lambda$ . After inserting Equation (2.1) and the expansion of  $e^{-S_\Lambda[\chi_\Lambda]}$  this reduces the problem to an evaluation of the free multipoint functions  $\langle \chi_\mu(x_1) \dots \chi_\mu(x_n) Q_\ell[\chi_\mu] \rangle_{S_{*,\mu}}$  computed with the quadratic action

$S_{*,\mu}$  on the field space with cutoff  $\Lambda$ . The free multipoint functions will contain contributions which diverge in the limit  $\Lambda \rightarrow \infty$ . On the other hand via the parameterization (2.1, 2.2) the coefficients carry an adjustable  $\Lambda$  dependence. In a renormalizable QFT the  $\Lambda$  dependence in the coefficients can be chosen such as to cancel (for  $\mu \ll \Lambda$ ) that generated by the multipoint functions  $\langle \chi_\mu(x_1) \dots \chi_\mu(x_n) \mathcal{Q}_l[\chi_\mu] \rangle_{S_{*,\mu}}$ . With this adjustment the limits

$$\sum_{\ell \geq 0} \lambda^\ell \lim_{\Lambda \rightarrow \infty} \langle \chi_\Lambda(x_1) \dots \chi_\Lambda(x_n) \rangle_{S_{\Lambda,\ell}} =: \sum_{\ell \geq 0} \lambda^\ell \langle \chi_\mu(x_1) \dots \chi_\mu(x_n) \rangle_{S_{\mu,\ell}}, \quad (2.3)$$

exist and define the renormalized multipoint functions. As indicated they can be interpreted as referring to the renormalized action  $\lim_{\Lambda \rightarrow \infty} S_\Lambda[\chi_\Lambda] = S_\mu[\chi_\mu]$ . Equation (2.3) highlights the main advantage of renormalized perturbation theory: The existence of the infinite cutoff limit (2.3) is often a provable property of the system, while this is not the case for most nonperturbative techniques. In the terminology introduced in Section 1.3 the Criterion (PTC1) is then satisfied. In order for this to be indicative for the existence of a genuine continuum limit, however, the additional Condition (PTC2) must be satisfied, whose rationale we proceed to discuss now.

Since the renormalization scale  $\mu$  is arbitrary, changing its value must not affect the values of observables. The impact of a change in  $\mu$  can most readily be determined from Equation (2.1). The left-hand-sides are  $\mu$  independent, so by differentiating these relations with respect to  $\mu$  and extracting the coefficients in a power series in (say)  $\Lambda$  and/or  $\log \Lambda$  consistency conditions arise for the derivatives  $\mu \frac{d}{d\mu} u_i$  and  $\mu \frac{d}{d\mu} \chi_\mu$ . The ones obtained from the leading order are the most interesting relations. For the couplings one obtains a system of ordinary differential equations which define their renormalization flow under a change of  $\mu$ . As usual it is convenient to work with dimensionless couplings  $g_i := u_i \mu^{-d_i}$ , where  $d_i$  is the mass dimension of  $u_i$ . The flow equations then take the form

$$\mu \frac{d}{d\mu} g_i = \beta_i(g(\mu)), \quad (2.4)$$

where the  $\beta_i$  are the perturbative beta functions. The flow equations for the renormalized fields are familiar only in the case of multiplicatively renormalizable fields, where one can work with scale independent fields and have the scale dependence carried by the wave function renormalization constant. In general however the fields are scale dependent. For example this ensures that the renormalized action evaluated on the renormalized fields is scale independent:  $\mu \frac{d}{d\mu} S_\mu[\chi_\mu] = 0$ .

By construction the perturbative beta functions have a fixed point at  $g_i^* = 0$ , which is called the *perturbative Gaussian fixed point*. Nothing prevents them from having other fixed points, but the Gaussian one is built into the construction. This is because a free theory has vanishing beta functions and the couplings  $g_i = u_i \mu^{-d_i}$  have been introduced to parameterize the deviations from the free theory with action  $S_{*,\mu}$ . Not surprisingly the stability matrix  $\Theta_{ij} = \partial \beta_i / \partial g_j |_{g^*=0}$  of the perturbative Gaussian fixed point just reproduces the information which has been put in. The eigenvalues come out to be  $-d_i$  modulo corrections in the loop coupling parameter, where  $-d_i$  are the mass dimensions of the corresponding interaction monomials. For the eigenvectors one finds a one-to-one correspondence to the unit vectors in the ‘coupling direction’  $g_i$ , again with power corrections in the loop counting parameter. One sees that the couplings  $u_i$  not irrelevant with respect to the stability matrix  $\Theta$  computed at the perturbative Gaussian fixed point are the ones with mass dimensions  $d_i \geq 0$ , i.e. just the power counting renormalizable ones.

The attribute ‘perturbative Gaussian’ indicates that whenever in a nonperturbative construction of the renormalization flow in the same ‘basis’ of interaction monomials  $g_i^* = 0$  is also a fixed point (called the *Gaussian fixed point*), the perturbatively defined expectations are *believed* to provide an asymptotic (nonconvergent) expansion to the expectations defined nonperturbatively based on the Gaussian fixed point, schematically

$$\langle \mathcal{O} \rangle_{\text{Gaussian FP}} \sim \sum_{\ell \geq 0} \lambda^\ell \langle \mathcal{O} \rangle_\ell. \quad (2.5)$$

Here  $\langle \mathcal{O} \rangle_\ell$  is the perturbatively computed  $\ell$ -loop contribution after a so-called renormalization group improvement. Roughly speaking the latter amounts to the following procedure: One assigns to the loop counting parameter  $\lambda$  a numerical value (ultimately related to the value of Planck's constant in the chosen units; see however [106]) and solves  $\mu \frac{dg_i}{d\mu} = \beta_i(g(\mu))$  as an ordinary differential equation. One of the functions obtained, say  $g_1(\mu)$ , is used to eliminate  $\lambda$  in favor of  $\mu$  and an integration constant  $\Lambda_{\text{beta}}$  (not to be confused with the cutoff, which is gone for good). So  $\langle \mathcal{O} \rangle_L := \sum_{\ell \leq L} \bar{\lambda}(\mu, \Lambda_{\text{beta}})^\ell \langle \mathcal{O} \rangle_{\ell, \mu}$  at this point carries a two-fold  $\mu$ -dependence, the one which comes out of the renormalization procedure (2.3) and the one carried now by  $\bar{\lambda}(\mu, \Lambda_{\text{beta}})$ . For an observable quantity  $\mathcal{O}$  both dependencies cancel out, modulo terms of higher order, leaving behind a dependence on the integration constant  $\Lambda_{\text{beta}}$ . We write  $\langle \mathcal{O} \rangle_L(\Lambda_{\text{beta}})$  to indicate this dependence. One then uses the expectation of one, suitably chosen, observable  $\mathcal{O}_0$  to match its value  $\langle \mathcal{O}_0 \rangle$  (measured or otherwise known) with that of  $\langle \mathcal{O}_0 \rangle_L(\Lambda_{\text{beta}})$  to a given small loop order  $L$  (typically not larger than 2). For a well chosen  $\mathcal{O}_0$  this allows one to replace  $\Lambda_{\text{beta}}$  by a physical mass scale  $m_{\text{phys}}$ . Eliminating  $\Lambda_{\text{beta}}$  in favor of  $m_{\text{phys}}$  gives the perturbative predictions for all other observables. Apart from residual scheme dependencies (which are believed to be numerically small) this defines the right-hand-side of Equation (2.5) unambiguously as a functional over the observables.

Nevertheless, except for some special cases, it is difficult to give a mathematically precise meaning to the ' $\sim$ ' in Equation (2.5). Ideally one would be able to prove that perturbation theory is asymptotic to the (usually unknown) exact answer for the same quantity. For lattice theories on a finite lattice this is often possible; the problems start when taking the limit of infinite lattice size (see [159] for a discussion). In the continuum limit a proof that perturbation theory is asymptotic has been achieved in a number of low-dimensional quantum field theories: the superrenormalizable  $P_2(\phi)$  and  $\phi_3^4$  theories [69, 43] and the two-dimensional Gross–Neveu model, where the correlation functions are the Borel sum of their renormalized perturbation expansion [87, 89]. Strong evidence for the asymptotic correctness of perturbation theory has also been obtained in the  $O(3)$  nonlinear sigma-model via the form factor bootstrap [22]. In four or higher-dimensional theories unfortunately no such results are available. It is still believed that whenever the above  $g_1$  is asymptotically free in perturbation theory, that the corresponding series is asymptotic to the unknown exact answer. On the other hand, to the best of our knowledge, a serious attempt to establish the asymptotic nature of the expansion has never been made, nor are plausible strategies available. The pragmatic attitude usually adopted is to refrain from the attempt to theoretically understand the domain of applicability of perturbation theory. Instead one interprets the ' $\sim$ ' in Equation (2.5) as an approximate numerical equality, to a suitable loop order  $L$  and in a benign scheme, as long as it works, and attributes larger discrepancies to the 'onset of nonperturbative physics'. This is clearly unsatisfactory, but often the best one can do. Note also that some of the predictive power of the QFT considered is wasted by this procedure and that it amounts to a partial immunization of perturbative predictions against (experimental or theoretical) refutation.

So far the discussion was independent of the nature of the running of  $\bar{\lambda}(\mu, \Lambda_{\text{beta}})$  (which was traded for  $g_1$ ). The chances that the vague approximate relation ' $\sim$ ' in Equation (2.5) can be promoted to the status of an asymptotic expansion are of course way better if  $\bar{\lambda}(\mu, \Lambda_{\text{beta}})$  is driven towards  $\bar{\lambda} = 0$  by the perturbative flow. Only then is it reasonable to expect that an asymptotic relation of the form (2.5) holds, linking the perturbative Gaussian fixed point to a genuine Gaussian fixed point defined by nonperturbative means. The perturbatively and the nonperturbatively defined coupling  $g_1$  can then be identified asymptotically and lie in the unstable manifold of the fixed point  $g_1 = 0$ . On the other hand the existence of a Gaussian fixed point with a nontrivial unstable manifold is thought to entail the existence of a genuine continuum limit in the sense discussed before. In summary, if  $g_1$  is traded for a running  $\bar{\lambda}(\mu, \Lambda_{\text{beta}})$ , a perturbative criterion for the existence of a genuine continuum limit is that the perturbative flow of  $g_1$  is regular with  $\lim_{\mu \rightarrow \infty} g_1(\mu) = 0$ . Since the beta functions of the other couplings are formal power series in  $\lambda$

without constant coefficients, the other couplings will vanish likewise as  $g_1 \rightarrow 0$ , and one recovers the local quadratic action  $S_{*,\mu}[\chi]$  at the fixed point. The upshot is that the coupling with respect to which the perturbative expansion is performed should be asymptotically free in perturbation theory in order to render the existence of a nonperturbative continuum limit plausible.

The reason for going through this discussion is to highlight that it applies just as well to a *perturbative non-Gaussian fixed point*. This sounds like a contradiction in terms, but it is not. Suppose that in a situation with several couplings  $g_1, \dots, g_n$  the perturbative beta functions (which are formal power series in  $\lambda$  without constant coefficients) admit a nontrivial zero,  $g_1^*(\lambda), \dots, g_n^*(\lambda)$ . Suppose in addition that *all* the couplings lie in the unstable manifold of that zero, i.e. the flows  $g_i(\mu)$  are regular and  $\lim_{\mu \rightarrow \infty} g_i(\mu) = g_i^*$ . We shall call a coupling with this property *asymptotically safe*, so that the additional assumption is that all couplings are asymptotically safe. As before one must assign  $\lambda$  a numerical value in order to define the flow. Since the series in  $\lambda$  anyhow has zero radius of convergence, the ‘smallness’ of  $\lambda$  is not off-hand a measure for the reliability of the perturbative result (the latter intuition in fact precisely presupposes Equation (2.5)). Any one of the deviations  $\delta g_i = g_i - g_i^*$ , which is of order  $\lambda$  at some  $\mu$  can be used as well to parameterize the original loop expansion. By a relabeling or reparameterization of the couplings we may assume that this is the case for  $\delta g_1$ . The original loop expansion can then be rearranged to read  $\sum_{\ell \geq 0} (\delta g_1)^\ell \langle \mathcal{O} \rangle_\ell$ . However, if there is an underlying nonperturbative structure at all, it is reasonable to assume that it refers to a non-Gaussian fixed point,

$$\langle \mathcal{O} \rangle_{\text{non-Gaussian FP}} \sim \sum_{\ell \geq 0} (\delta g_1)^\ell \langle \mathcal{O} \rangle_\ell. \quad (2.6)$$

The rationale for Equation (2.6) is exactly the same as for Equation (2.5). What matters is not the value of the couplings at a perturbative fixed point, but their flow pattern. For a nontrivial fixed point the couplings  $g_i^*$  in the above basis of interaction monomials are nonzero, but any one of the deviations  $\delta g_i = g_i - g_i^*$  can be made arbitrarily small as  $\mu \rightarrow \infty$ . The relation ‘ $\sim$ ’ in Equation (2.6) then again plausibly amounts to an asymptotic expansion for the unknown exact answer, where the latter this time is based on a non-Gaussian fixed point.

Summarizing: In perturbation theory the removal of the cutoff can be done independently of the properties of the coupling flow, while in a non-perturbative setting both aspects are linked. Only if the coupling flow computed from the perturbative beta functions meets certain conditions is it reasonable to assume that there exists an underlying non-perturbative framework to whose results the perturbative series is asymptotic. Specifically we formulate the following criterion:

**Criterion** (Continuum limit via perturbation theory):

**(PTC1)** Existence of a formal continuum limit, i.e. removal of the UV cutoff is possible and the renormalized physical quantities are independent of the scheme and of the choice of interpolating fields, all in the sense of formal power series in the loop counting parameter.

**(PTC2)** The perturbative beta functions have a Gaussian or a non-Gaussian fixed point and the dimension of its unstable manifold (as computed from the perturbative beta functions) equals the number of independent essential couplings. Equivalently, all essential couplings are asymptotically safe in perturbation theory.

## 2.2 Functional flow equations and UV renormalization

The technique of functional renormalization group equations (FRGEs) does not rely on a perturbative expansion and has been widely used for the computation of critical exponents and the flow of generalized couplings. For a systematic exposition of this technique and its applications we refer to the reviews [146, 21, 166, 229, 29]. Here we shall mainly use the effective average action  $\Gamma_k$  and

its ‘exact’ FRGE. We refer to Appendix C for a summary of this formulation, and discuss in this section how the UV renormalization problem presents itself in an FRGE [157].

In typical applications of the FRG the ultraviolet renormalization problem does not have to be addressed. In the context of the asymptotic safety scenario this is different. By definition the perturbative series in a field theory based on an asymptotically safe functional measure has a dependence on the UV cutoff which is not strictly renormalizable (see Section 1.3). The perturbative expansion of an FRGE must reproduce the structure of these divergencies. On the other hand in an exact treatment or based on different approximation techniques a reshuffling of the cutoff dependence is meant to occur which allows for a genuine continuum limit. We therefore outline here how the UV renormalization problem manifests itself in the framework of the functional flow equations. The goal will be to formulate a criterion for the plausible existence of a genuine continuum limit in parallel to the one above based on perturbative indicators.

Again we illustrate the relevant issues for a scalar quantum field theory on flat space. For definiteness we consider here the flow equation for the *effective average action*  $\Gamma_{\Lambda,k}[\phi]$ , for other types of FRGEs the discussion is similar though. The effective average action interpolates between the bare action  $S_{\Lambda}[\phi]$  and the above, initially regulated, effective action  $\Gamma_{\Lambda}$ , according to

$$S_{\Lambda}[\phi] \xleftarrow{k \rightarrow \Lambda} \Gamma_{\Lambda,k}[\phi] \xrightarrow{k \rightarrow 0} \Gamma_{\Lambda}[\phi]. \quad (2.7)$$

Roughly speaking one should think of  $\Gamma_{\Lambda,k}[\phi]$  as the conventional effective action but with only the momentum modes in the range  $k^2 < p^2 < \Lambda^2$  integrated out. The  $k \rightarrow \Lambda$  limit in Equation (2.7) will in fact differ from  $S_{\Lambda}$  by a 1-loop determinant  $\ln \det[S_{\Lambda}^{(2)} + \mathcal{R}_{\Lambda}]$  (see Appendix C.2). For the following discussion the difference is inessential and for (notational) simplicity we will identify  $\Gamma_{\Lambda,\Lambda}$  with  $S_{\Lambda}$ . Equation (2.7) also presupposes that for fixed UV cutoff  $\Lambda$  the limit  $k \rightarrow 0$  exists, which for theories with massless degrees of freedom is nontrivial.

The conventional effective action obeys a well-known functional integro-differential equation which implicitly defines it (see Equation (B.8) below). Its counterpart for  $\Gamma_{\Lambda,k}[\phi]$  reads

$$\exp\{-\Gamma_{\Lambda,k}[\phi]\} = \int [\mathcal{D}\chi]_{\Lambda,k} \exp\left\{-S_{\Lambda}[\chi] + \int dx (\chi - \phi)(x) \frac{\delta \Gamma_{\Lambda,k}[\phi]}{\delta \phi(x)}\right\}, \quad (2.8)$$

where the functional measure  $[\mathcal{D}\chi]_{\Lambda,k}$  includes mostly momentum modes in the range  $k^2 < p^2 < \Lambda^2$ . This can be done by multiplying the kinematical measure by a suitable mode suppression factor

$$[\mathcal{D}\chi]_{\Lambda,k} = \mathcal{D}\chi \exp\{-C_{\Lambda,k}[\chi - \phi]\}, \quad (2.9)$$

with a suitable quadratic form  $C_{\Lambda,k}$ . From Equation (2.8) one can also directly verify the alternative characterization (1.12).

The precise form of the mode suppression is inessential. In the following we outline a variant which is technically convenient. Here  $C_{\Lambda,k}$  is a quadratic form in the fields defined in terms of a kernel  $\mathcal{R}_{\Lambda,k}$  chosen such that both  $\mathcal{R}_{\Lambda,k}$  and  $k\partial_k \mathcal{R}_{\Lambda,k}$  define integral operators of trace-class on the function space considered. We write  $[\mathcal{R}_{\Lambda,k}\chi](x) := \int dy \mathcal{R}_{\Lambda,k}(x,y)\chi(y)$  for the integral operator and  $\text{Tr}[\mathcal{R}_{\Lambda,k}] := \int dx \mathcal{R}_{\Lambda,k}(x,x) < \infty$  for its trace. The other properties of the kernel are best described in Fourier space, where  $\mathcal{R}_{k,\Lambda}$  acts as  $[\mathcal{R}_{\Lambda,k}\hat{\chi}](p) = \int \frac{dq}{(2\pi)^d} \mathcal{R}_{\Lambda,k}(p,q)\hat{\chi}(q)$ , with  $\hat{\chi}(p) = \int dx \chi(x) \exp(-ipx)$ , the Fourier transform of  $\chi$  and similarly for the kernel (where we omit the hat for notational simplicity). The UV cutoff  $0 \leq p^2 < \Lambda$  renders Euclidean momentum space compact and Mercer’s theorem then provides simple sufficient conditions for an integral operator to be trace-class [157]. We thus take the kernel  $\mathcal{R}_{\Lambda,k}(p,q)$  to be smooth, symmetric in  $p,q$ , and such that

$$C_{\Lambda,k}[\chi] := \frac{1}{2} \int \frac{dp}{(2\pi)^d} \frac{dq}{(2\pi)^d} \hat{\chi}(q)^* \mathcal{R}_{\Lambda,k}(p,q) \hat{\chi}(p) = \frac{1}{2} \int dx \chi(x)^* [\mathcal{R}_{\Lambda,k}\chi](x) \geq 0 \quad (2.10)$$



for all continuous functions  $\chi$  (similarly for  $k\partial_k\mathcal{R}_{\Lambda,k}(p,q)$ ). The trace-class condition is then satisfied and one can adjust the other features of the kernel to account for the mode suppression. These features are arbitrary to some extent; what matters is the limiting behavior for  $p^2, q^2 \gg k^2$  and (with foresight)  $\Lambda \rightarrow \infty$ . We refer to Appendix C for more details.

The presence of the extra scale  $k$  allows one to convert Equation (2.8) into a functional differential equation [228, 229, 29],

$$\begin{aligned} k\frac{\partial}{\partial k}\Gamma_{\Lambda,k}[\phi] &= \frac{1}{2}\text{Tr}\left[k\frac{\partial}{\partial k}\mathcal{R}_{\Lambda,k}\left(\Gamma_{\Lambda,k}^{(2)}[\phi] + \mathcal{R}_{\Lambda,k}\right)^{-1}\right] \\ &= \frac{1}{2}\int\frac{dq_1}{(2\pi)^d}\frac{dq_2}{(2\pi)^d}k\partial_k\mathcal{R}_{\Lambda,k}(q_1,q_2)\left(\Gamma_{\Lambda,k}^{(2)}[\phi] + \mathcal{R}_{\Lambda,k}\right)^{-1}(q_2,q_1), \end{aligned} \quad (2.11)$$

known as the *functional renormalization group equation* (FRGE) for the *effective average action*. For convenience we include a quick derivation of Equation (2.11) in Appendix C. In the second line of Equation (2.11) we spelled out the trace using that  $k\partial_k\mathcal{R}_{\Lambda,k}$  is trace-class. Further  $\Gamma_{\Lambda,k}^{(2)}[\phi]$  is the integral operator whose kernel is the Hessian of the effective average action, i.e.  $\Gamma_{\Lambda,k}^{(2)}(x,y) := \delta^2\Gamma_{\Lambda,k}[\phi]/(\delta\phi(x)\delta\phi(y))$ , and  $\mathcal{R}_{\Lambda,k}$  is the integral operator in Equation (2.8).

For finite cutoffs  $(\Lambda,k)$  the trace of the right-hand-side of Equation (2.11) will exist as the potentially problematic high momentum parts are cut off. In slightly more technical terms, since the product of a trace-class operator with a bounded operator is again trace-class, the trace in Equation (2.11) is finite as long as the inverse of  $\Gamma_{\Lambda,k}^{(2)}[\phi] + \mathcal{R}_{\Lambda,k}$  defines a bounded operator. For finite UV cutoff one sees from the momentum space version of Equation (B.2) in Section 3.4 that this will normally be the case. The trace-class property of the mode cutoff operator (for which Equation (2.10) is a sufficient condition) also ensures that the trace in Equation (2.11) can be evaluated in any basis, the momentum space variant displayed in the second line is just one convenient choice.

Importantly the FRGE (2.11) is independent of the bare action  $S_\Lambda$ , which enters only via the initial condition  $\Gamma_{\Lambda,\Lambda} = S_\Lambda$  (for large  $\Lambda$ ). In the FRGE approach the calculation of the functional integral for  $\Gamma_{\Lambda,k}$  is replaced by the task of integrating this RG equation from  $k = \Lambda$ , where the initial condition  $\Gamma_{\Lambda,\Lambda} = S_\Lambda$  is imposed, down to  $k = 0$ , where the effective average action equals the ordinary effective action  $\Gamma_\Lambda$ .

All this has been for a fixed UV cutoff  $\Lambda$ . The removal of the cutoff is of course the central theme of UV renormalization. In the FRG formulation one has to distinguish between two aspects: first, removal of the explicit  $\Lambda$  dependence in the trace on the right-hand-side of Equation (2.11), and second removal of the UV cutoff in  $\Gamma_{\Lambda,k}$  itself, which was needed in order to make the original functional integral well-defined.

The first aspect is unproblematic: The trace is manifestly finite as long as the inverse of  $\Gamma_{\Lambda,k}^{(2)}[\phi] + \mathcal{R}_{\Lambda,k}$  defines a bounded operator. If now  $\Gamma_{\Lambda,k}^{(2)}[\phi]$  is *independently* known to have a finite and nontrivial limit as  $\Lambda \rightarrow \infty$ , the explicit  $\Lambda$  dependence carried by the  $\mathcal{R}_{\Lambda,k}$  term is harmless and the trace always exists. Roughly this is because the derivative kernel  $k\partial_k\mathcal{R}_{\Lambda,k}$  has support mostly on a thin shell around  $p^2 \approx k^2$ , so that the (potentially problematic) large  $p$  behavior of the other factor is irrelevant (cf. Appendix C.2).

The second aspect of course relates to the traditional UV renormalization problem. Since  $\Gamma_{\Lambda,k}$  came from a regularized functional integral it will develop the usual UV divergencies as one attempts to send  $\Lambda$  to infinity. The remedy is to carefully adjust the bare action  $S_\Lambda[\phi]$  – that is, the initial condition for the FRGE (2.11) – in such a way that functional integral – viz. the solution of the FRGE – is asymptotically independent of  $\Lambda$ . Concretely this could be done by fine-tuning the way how the parameters  $u_\alpha(\Lambda)$  in the expansion  $S_\Lambda[\chi] = \sum_\alpha u_\alpha(\Lambda)P_\alpha[\chi]$  depends on  $\Lambda$ . However the FRGE method in itself provides *no means* to find the proper initial functional  $S_\Lambda[\chi]$ . Identification of the fine-tuned  $S_\Lambda[\chi]$  lies at the core of the UV renormalization problem,

irrespective of whether  $\Gamma_{\Lambda,k}$  is defined directly via the functional integral or via the FRGE. Beyond perturbation theory the only known techniques to identify the proper  $S_\Lambda$  start directly from the functional integral and are ‘constructive’ in spirit (see [195, 48]). Unfortunately four-dimensional quantum field theories of interest are still beyond constructive control.

One may also ask whether perhaps the cutoff-dependent FRGE (2.11) itself can be used to show that a limit  $\lim_{\Lambda \rightarrow \infty} \Gamma_{\Lambda,k}[\phi]$  exists. Indeed using other FRGEs and a perturbative ansatz for the solution has led to economic proofs of perturbative renormalizability, i.e. of the existence of a formal continuum limit in the sense of Criterion (PTC1) discussed before (see [200, 123]). Unfortunately so far this could not be extended to construct a nonperturbative continuum limit of fully fledged quantum field theories (see [145] for a recent review of such constructive uses of FRGEs). For the time being one has to be content with the following *if...then* statement:

*If* there exists a sequence of initial actions  $S_{n\Lambda_0}[\chi]$ ,  $n \in \mathbb{N}$ , such that the solution  $\Gamma_{n\Lambda_0,k}[\phi]$  of the FRGE (2.11) remains finite as  $n \rightarrow \infty$ , *then* the limit  $\Gamma_k[\phi] := \lim_{n \rightarrow \infty} \Gamma_{n\Lambda_0,k}[\phi]$  has to obey the cut-off independent FRGE

$$k \frac{\partial}{\partial k} \Gamma_k[\phi] = \frac{1}{2} \text{Tr} \left[ k \frac{\partial}{\partial k} \mathcal{R}_{k,\infty} \left( \Gamma_k^{(2)}[\phi] + \mathcal{R}_{k,\infty} \right)^{-1} \right]. \quad (2.12)$$

Conversely, under the above premise, this equation should have at least one solution with a finite limit  $\lim_{k \rightarrow \infty} \Gamma_k[\phi]$ . This limit can now be identified with the renormalized fixed point action  $S_*[\chi]$ . It is renormalized because by construction the cutoff dependencies have been eaten up by the ones produced by the trace in Equation (2.11). It can be identified with a fixed point action because lowering  $k$  amounts to coarse graining, and  $S_*[\chi]$  is the ‘inverse limit’ of a sequence of such coarse graining steps.

So far the positivity or unitarity requirement has not been discussed. From the (Osterwalder–Schrader or Wightman) reconstruction theorems it is known how the unitarity of a quantum field theory on a flat spacetime translates into nonlinear conditions on the multipoint functions. Since the latter can be expressed in terms of the functional derivatives of  $\Gamma_k$ , unitarity can in principle be tested retroactively, and is expected to hold only in the limit  $k \rightarrow 0$ . Unfortunately this is a very indirect and retroactive criterion. One of the roles of the bare action  $S_\Lambda[\chi] = \Gamma_{\Lambda,\Lambda}[\chi]$  is to encode properties which are likely to ensure the desired properties of  $\lim_{k \rightarrow 0} \Gamma_k[\phi]$ . In theories with massless degrees of freedom the  $k \rightarrow 0$  limit is nontrivial and the problem of gaining computational control over the infrared physics should be separated from the UV aspects of the continuum limit as much as possible. However the  $k \rightarrow 0$  limit is essential to probe stability/positivity/unitarity.

One aspect of positivity is the convexity of the effective action. The functional equations (2.11, 2.12) do in itself “not know” that  $\Gamma_k$  is the Legendre transform of a convex functional and hence must be itself convex. Convexity must therefore enter through the initial data and it will also put constraints on the choice of the mode cutoffs. Good mode cutoffs are characterized by the fact that  $\Gamma_k^{(2)} + \mathcal{R}_k$  has positive spectral values for all  $k$  (cf. Equation (C.14)). If no blow-up occurs in the flow the limit  $\lim_{k \rightarrow 0} \Gamma_k^{(2)}$  will then also have non-negative spectrum. Of course this presupposes again that the proper initial conditions have been identified and the role of the bare action is as above.

For flat space quantum field theories one expects that  $S_\Lambda[\chi]$  must be local, i.e. a differential polynomial of finite order in the fields so as to end up with an effective action  $\lim_{k \rightarrow \infty} \lim_{\Lambda \rightarrow \infty} \Gamma_{\Lambda,k}[\phi]$  describing a local/microcausal unitary quantum field theory.

For convenient reference we summarize these conclusions in the following criterion:

**Criterion** (Continuum limit in the functional RG approach):

**(FRGC1)** A solution of the cutoff independent FRGE (2.12) which exists globally in  $k$  (for all  $0 \leq k \leq \infty$ ) can reasonably be identified with the continuum limit of the effective average action  $\lim_{\Lambda \rightarrow \infty} \Gamma_{\Lambda,k}[\phi]$  constructed by other means. For such a solution  $\lim_{k \rightarrow 0} \Gamma_k[\phi]$  is the full quantum



effective action and  $\lim_{k \rightarrow \infty} \Gamma_k[\phi] = S_*[\phi]$  is the fixed point action.

**(FRGC2)** For a unitary relativistic quantum field theory positivity/unitarity must be tested retroactively from the functional derivatives of  $\lim_{k \rightarrow 0} \Gamma_k[\phi]$ .

We add some comments:

Since the FRGE (2.12) is a differential equation in  $k$ , an initial functional  $\Gamma_{\text{initial}}[\phi]$  has to be specified for some  $0 < k_{\text{initial}} \leq \infty$ , to generate a local solution near  $k = k_{\text{initial}}$ . The point is that for ‘almost all’ choices of  $\Gamma_{\text{initial}}[\phi]$  the local solution *cannot* be extended to all values of  $k$ . Finding the rare initial functionals for which this is possible is the FRGE counterpart of the UV renormalization problem. The existence of the  $k \rightarrow 0$  limit is itself not part of the UV problem; in conventional quantum field theories the  $k \rightarrow 0$  limit is however essential to probe unitarity/positivity/stability.

It is presently not known whether the above criterion can be converted into a theorem. Suppose for a quantum field theory on the lattice (with lattice spacing  $\Lambda^{-1}$ ) the effective action  $\Gamma_{\Lambda,k}^{\text{latt}}[\phi]$  has been constructed nonperturbatively from a transfer operator satisfying reflection positivity and that a continuum limit  $\lim_{\Lambda \rightarrow \infty} \Gamma_{\Lambda,k}^{\text{latt}}$  is assumed to exist. Does it coincide with a solution  $\Gamma_k[\phi]$  of Equation (2.12) satisfying the Criteria (FRGC1) and (FRGC2)? Note that this is ‘only’ a matter of controlling the limit, for finite  $\Lambda$  also  $\Gamma_{\Lambda,k}^{\text{latt}}$  will satisfy the flow equation (2.11).

For an application to quantum gravity one will initially only ask for Criterion (FRGC1), perhaps with even only a partial understanding of the  $k \rightarrow 0$  limit. As mentioned, the  $k \rightarrow 0$  limit should also be related to positivity issues. The proper positivity requirement replacing Criterion (FRGC2) yet has to be found, however some constraint will certainly be needed. Concerning Criterion (FRGC1) the premise in the *if...then* statement preceding Equation (2.12) has to be justified by external means or taken as a working hypothesis. In principle one can also adopt the viewpoint that the quantum gravity counterpart of Equation (2.12) discussed in Section 4 simply defines the effective action for quantum gravity whenever a solution meets Criterion (FRGC1). The main drawback with this proposal is that it makes it difficult to include information concerning Criterion (FRGC2). However difficult and roundabout a functional integral construction is, it allows one to incorporate ‘other’ desirable features of the system in a relatively transparent way.

We shall therefore also in the application to quantum gravity assume that a solution  $\Gamma_k$  of the cutoff independent FRGE (2.12) satisfying Criterion (FRGC1) comes from an underlying functional integral. This amounts to the assumption that the renormalization problem for  $\Gamma_{k,\Lambda}$  defined in terms of a functional integral can be solved and that the limit  $\lim_{\Lambda \rightarrow \infty} \Gamma_{k,\Lambda}$  can be identified with  $\Gamma_k$ . This is of course a rather strong hypothesis, however its self-consistency can be tested within the FRG framework.

To this end one truncates the space of candidate continuum functionals  $\Gamma_k^{\text{trunc}}[\phi]$  to one where the initial value problem for the flow equation (2.12) can be solved in reasonably closed form. One can then by ‘direct inspection’ determine the initial data for which a global solution exists. Convexity of the truncated  $\lim_{k \rightarrow 0} \Gamma_k^{\text{trunc}}[\phi]$  can serve as guideline to identify good truncations. If the set of these initial data forms a nontrivial unstable manifold of the fixed point  $S_*^{\text{trunc}}[\phi] = \lim_{k \rightarrow \infty} \Gamma_k^{\text{trunc}}[\phi]$ , application of the above criterion suggests that  $\Gamma_k^{\text{trunc}}$  can approximately be identified with the projection of the continuum limit  $(\lim_{\Lambda \rightarrow \infty} \Gamma_{\Lambda,k})^{\text{trunc}}$  of some  $\Gamma_{\Lambda,k}$  computed by other means. The identification can only be an approximate one because in the  $\Gamma_k^{\text{trunc}}$  evolution one first truncates and then evolves in  $k$ , while in  $(\lim_{\Lambda \rightarrow \infty} \Gamma_{\Lambda,k}[\phi])^{\text{trunc}}$  one first evolves in  $k$  and then truncates. Alternatively one can imagine to have replaced the original dynamics by some ‘hierarchical’ (for want of a better term) approximation implicitly defined by the property that  $(\lim_{\Lambda \rightarrow \infty} \Gamma_{\Lambda,k}^{\text{hier}}[\phi])^{\text{trunc}} = \lim_{\Lambda \rightarrow \infty} \Gamma_{\Lambda,k}^{\text{hier}}[\phi]$  (see [75] for the relation between a hierarchical dynamics and the local potential approximation). The existence of an UV fixed point with a nontrivial unstable manifold for  $\Gamma_k^{\text{trunc}}$  can then be taken as witnessing the renormalizability of the ‘hierarchical’ dynamics.

## 2.3 Towards Quantum Gravidynamics

The application of renormalization group ideas to quantum gravity has a long history. In accordance with the previous discussion, we focus here on the aspects needed to explain the apparent mismatch between the perturbative non-renormalizability and the presumed nonperturbative renormalizability. In fact, when looking at higher derivative theories renormalizability *can* already be achieved on a perturbative level in several instructive ways.

### 2.3.1 The role of Newton's constant

Before turning to renormalization aspects proper, let us describe the special role of Newton's constant in a diffeomorphism invariant theory with a dynamical metric. Let  $S[g, \text{matter}]$  be any local action, where  $g = (g_{\alpha\beta})_{1 \leq \alpha, \beta \leq d}$  is the metric and the “matter” fields are not scaled when the metric is. Scale changes in the metric then give rise to a variation of the Lagrangian which vanishes on shell:

$$\frac{d}{d\omega^2} S[\omega^2 g, \text{matter}] \Big|_{\omega=1} = \int dx \sqrt{g} g_{\alpha\beta} \frac{\delta S[g, \text{matter}]}{\delta g_{\alpha\beta}}. \quad (2.13)$$

As a consequence one of the coupling parameters which in the absence of gravity would be *essential* (i.e. a genuine coupling) becomes *inessential* (i.e. can be changed at will by a redefinition of the fields). The running of this parameter, like that of a wave function renormalization constant, has no direct significance. If the pure gravity part contains the usual Ricci scalar term  $Z_N \sqrt{g} R(g)$ , the parameter that becomes inessential may be taken as its prefactor, i.e. may be identified with the inverse of Newton's constant, via

$$Z_N^{-1} = 2 \frac{d-2}{d-3} \text{Vol}(S^{d-2}) G_{\text{Newton}} =: c_d G_{\text{Newton}}. \quad (2.14)$$

The normalization factor  $c_d$ ,  $d \geq 4$  [196], is chosen such that the coefficient in the nonrelativistic force law, as computed from  $Z_N \sqrt{g} R(g) + L_{\text{matter}}$ , equals  $G_{\text{Newton}} \text{Vol}(S^{d-2})$ . For  $d = 2, 3$  a different normalization has to be adopted.

The physics interpretation of the inessential parameter  $\omega$  is that it also sets the absolute momentum or spectral scale. To see this we can think of  $g_{\alpha\beta}$  as a reference metric in the background field formalism. For example for the spectral values  $\nu(g)$  of the covariant Laplacian  $\Delta_g$  associated with  $g_{\alpha\beta}$  one has

$$\nu(\omega^2 g) = \omega^{-2} \nu(g), \quad (2.15)$$

since  $\Delta_{\omega^2 g} = \omega^{-2} \Delta_g$ . The spectral values play the role of a covariant momentum squared. Indeed, if the metric is taken dimensionless,  $\nu(g)$  carries dimension 2 (since  $\Delta_g$  does) and for a flat metric  $g_{\alpha\beta} = \eta_{\alpha\beta}$  they reduce to  $-\nu(\eta) = k^2$ , for plane waves labeled by  $k$ . From Equation (2.15) one sees that rescaling of the metric and rescaling of the spectral values amount to the same thing. Since the former parameter is inessential the latter is too. Hence in a theory with a dynamical metric the three (conceptually distinct) inessential parameters – overall scale of the metric  $\omega$ , the inverse of Newton's constant  $Z_N^{-1} = c_d G_{\text{Newton}}$ , and the overall normalization of the spectral/momentum values – are in one-to-one correspondence. For definiteness we take Newton's constant as the variant under consideration.

Being inessential the quantum field theoretical running of  $G_{\text{Newton}}$  has significance only relative to the running coefficient of some reference operator. The most commonly used choice is a cosmological constant term  $\Lambda \int dx \sqrt{g}$ . Indeed

$$G_{\text{Newton}} \Lambda^{\frac{d-2}{d}} = \text{const } \tau(\mu)^{2/d} \quad (2.16)$$

is dimensionless and invariant under constant rescalings of the metric [116]. One usually switches to dimensionless parameters via

$$c_d G_{\text{Newton}} = \mu^{2-d} g_{\text{N}}(\mu), \quad \Lambda = 2\mu^d \frac{\lambda(\mu)}{g_{\text{N}}(\mu)}, \quad (2.17)$$

where  $\mu$  is some dimension one parameter which will be taken as ‘renormalization group time’. The Einstein–Hilbert action then reads

$$\frac{\mu^{d-2}}{g_{\text{N}}(\mu)} \int dx \sqrt{g} [R(g) - 2\mu^2 \lambda(\mu)]. \quad (2.18)$$

Being dimensionless one expects the running of  $g_{\text{N}}(\mu)$  and  $\lambda(\mu)$  to be governed by flow equations without explicit  $\mu$  dependence

$$\mu \frac{\partial}{\partial \mu} g_{\text{N}} = \gamma_g(g_{\text{N}}, \lambda), \quad \mu \frac{\partial}{\partial \mu} \lambda = \beta_\lambda(g_{\text{N}}, \lambda). \quad (2.19)$$

For the essential parameter  $\tau(\mu) = g_{\text{N}}(\mu)\lambda(\mu)^{(d-2)/2}$  obtained from Equation (2.16) this gives

$$\mu \frac{\partial}{\partial \mu} \tau = \tau \left[ \frac{\gamma_g}{g_{\text{N}}} + \frac{d-2}{2} \frac{\beta_\lambda}{\lambda} \right]. \quad (2.20)$$

Within an asymptotically safe Quantum Gravidynamics this should be an asymptotically safe coupling, i.e.

$$\sup_{\mu_0 \leq \mu \leq \infty} \tau(\mu) < \infty, \quad \lim_{\mu \rightarrow \infty} \tau(\mu) = \tau_* < \infty, \quad (2.21)$$

where here  $0 < \tau_* < \infty$ . Given Equation (2.21) there are two possibilities. First, the various scheme choices are such that the parameters  $g_{\text{N}}(\mu)$  and  $\lambda(\mu)$  are both nonsingular and approach finite values  $g_*$  and  $\lambda_*$  for  $\mu \rightarrow \infty$ . Second, the scheme choices are such that one of them becomes singular and the other vanishes for  $\mu \rightarrow \infty$ . Usually the first possibility is chosen; then the  $g_{\text{N}}(\mu)$  flow defined by the first equation in Equation (2.19) has all the properties required for an essential asymptotically safe coupling. This ‘nonsingular parametric representation’ of the  $\tau(\mu)$  coupling flow is advantageous for most purposes.

The second possibility is realized when inserting a singular solution of the equation for  $g_{\text{N}}(\mu)$  into the equation for  $\lambda(\mu)$ . This naturally occurs when working in Planck units. One makes use of the fact that an inessential parameter can be frozen at a prescribed value. Specifically fixing

$$[G_{\text{Newton}}]^{1/(d-2)} = M_{\text{Pl}} \approx 1.4 \times 10^{19} \text{ GeV}, \quad (2.22)$$

amounts to working with Planck units. From Equation (2.17) it then follows that

$$g_{\text{N}}(\mu) = c_d \left( \frac{\mu}{M_{\text{Pl}}} \right)^{d-2}, \quad (d-2)c_d \left( \frac{\mu}{M_{\text{Pl}}} \right)^{d-2} = \gamma_g(g, \lambda). \quad (2.23)$$

We may assume that the second equation has a local solution  $g_{\text{N}}(\mu) = f(\lambda, \mu/M_{\text{Pl}})$ . Reinserting into the  $\lambda$  equation gives a flow equation

$$\mu \frac{\partial}{\partial \mu} \lambda(\mu) = \tilde{\beta}_\lambda(\lambda, \mu/M_{\text{Pl}}), \quad (2.24)$$

which now explicitly depends on  $\mu$ . Writing similarly  $\tilde{\tau}_* := \tau_*(f(\lambda, \mu/M_{\text{Pl}}), \lambda)$  the condition defining the  $\tau(\mu)$  fixed point becomes

$$\tilde{\beta}_\lambda \Big|_{\tilde{\tau}_*} = -2\lambda. \quad (2.25)$$

Both formulations are mathematically equivalent to the extent the inversion formula  $g_{\text{N}}(\mu) = f(\lambda, \mu/M_{\text{Pl}})$  is globally defined. For definiteness we considered here the cosmological constant term as a reference operator, but the principle clearly generalizes.

### 2.3.2 Perturbation theory and higher derivative theories

By higher derivative theories we mean here gravitational theories whose bare action contains, in addition to the Einstein–Hilbert term, scalars built from powers of the Riemann tensor and its covariant derivatives. In overview there are two distinct perturbative treatments of such theories.

The first one, initiated by Stelle [206], uses  $1/p^4$  type propagators (in four dimensions) in which case a higher derivative action containing all (three) quartic derivative terms can be expected to be power counting renormalizable. In this case strict renormalizability with only 4 (or 5, if Newton’s constant is included) couplings can be achieved [206]. However the  $1/p^4$  type propagators are problematic from the point of view of unitarity.

An alternative perturbative treatment of higher derivative theories was first advocated by Gomis–Weinberg [94]. The idea is to try to maintain a  $1/p^2$  type propagator and include all (infinitely many) counterterms generated in the bare action. Consistency requires that quadratic counterterms (those which contribute to the propagator) can be absorbed by field redefinitions. As shown by Anselmi [10] this is the case either in the absence of a cosmological constant term or when the background spacetime admits a metric with constant curvature.

We now present both of these perturbative treatments in more detail. A putative matching to a nonperturbative renormalization flow is outlined in Equation (2.32).

The general classical action in  $d$  dimensions containing up to four derivatives of the metric reads

$$\begin{aligned} S &= \int dx \sqrt{g} \left[ \Lambda - \frac{1}{c_d G_N} R + \frac{1}{2s} C^2 - \frac{\omega}{3s} R^2 + \frac{\theta}{s} E \right] \\ &= \int dx \sqrt{g} \left[ \Lambda - \frac{1}{c_d G_N} R + z R^2 + y R_{\alpha\beta} R^{\alpha\beta} + x R_{\alpha\beta\gamma\delta} R^{\alpha\beta\gamma\delta} \right]. \end{aligned} \quad (2.26)$$

Here  $C^2$  is the square of the Weyl tensor,  $E$  is the integrand of the Gauss–Bonnet term, and a total derivative term  $\nabla^2 R$  has been omitted. The sign of the  $C^2$  coupling,  $s > 0$ , is fixed by the requirement that the Euclidean functional integral is damping. The metric is Euclidean to facilitate comparison with the original literature. The parameterization of the coefficients by couplings  $s, \omega, \theta$  is chosen for later convenience; we follow the conventions of [59]. The parameters in the second line are related to those in the first by

$$sx = \frac{1}{2} + \theta, \quad sy = -\frac{2}{d-2} - 4\theta, \quad sz = -\frac{\omega}{3} + \theta + \frac{1}{(d-1)(d-2)}. \quad (2.27)$$

In  $d = 4$  the Gauss–Bonnet term is negligible; however if dimensional regularization is used,  $d \neq 4$ , it is crucial to keep the term. For  $d = 3$  both  $E$  and  $C^2$  vanish.

The perturbative quantization of Equation (2.26) proceeds as usual. Gauge fixing and ghost terms are added and the total action is expanded in powers of  $h_{\alpha\beta} = g_{\alpha\beta} - \delta_{\alpha\beta}$ . Due to the crucial  $C^2$  term the gauge-fixed propagator read off from the quadratic part of the full action has a characteristic  $1/p^4$  falloff in  $d = 4$ ,

$$G(p) \sim \frac{1}{p^2(p^2 + m^2)} = \frac{1}{m^2} \left( \frac{1}{p^2} - \frac{1}{p^2 + m^2} \right), \quad (2.28)$$

where  $m$  is the mass of a “wrong sign” propagating spin 2 mode [206, 207]. Also spin 0 modes with the “wrong sign” may occur depending on the coefficient of the  $R^2$  term [207, 49]. The one-loop counterterm (minus the divergent part of the effective action) has been computed by a number of authors using different regularizations: dimensional regularization [59, 19], proper time cut-off [83], zeta function [18]. The resulting one-loop flow equations in  $d = 4$  for  $s, \omega, \theta$  are agreed upon and

read

$$\begin{aligned}
(4\pi)^2 \mu \frac{d}{d\mu} s &= -\frac{133}{10} s^2, \\
(4\pi)^2 \mu \frac{d}{d\mu} \omega &= -\frac{25 + 1098\omega + 200\omega^2}{60} s, \\
(4\pi)^2 \mu \frac{d}{d\mu} \theta &= \frac{7(56 - 171\theta)}{90} s.
\end{aligned}
\tag{2.29}$$

These equations have a trivial fixed point  $s_* = 0$ ,  $\omega_* = \text{const}$ ,  $\theta_* = \text{const}$ , and a nontrivial fixed point  $s_* = 0$ ,  $\omega_* = -(549 \pm 7\sqrt{6049})/200$ ,  $\theta_* = 56/171$ . Importantly the  $C^2$  coupling  $s$  is asymptotically free.

To describe the flow of the Newton and cosmological constants one switches to the dimensionless parameters  $g_N$  and  $\lambda$  as in Section 2.3.1. The result obtained in Berrodo–Peixoto and Shapiro [59] via dimensional regularization reads in our conventions

$$\begin{aligned}
\gamma_g &= 2g_N - \frac{1}{(4\pi)^2} \frac{3 + 26\omega - 40\omega^2}{12\omega} s g_N - \gamma_1 g_N^2, \\
\beta_\lambda &= -2\lambda + \frac{1}{(4\pi)^2} \left[ \frac{1 + 20\omega^2}{8\omega^2 g_N} s^2 + \frac{1 + 86\omega + 40\omega^2}{12\omega} s \lambda \right] + \gamma_1 \lambda g_N.
\end{aligned}
\tag{2.30}$$

The function  $\gamma_1$  depends on the choice of gauge, but in the combination (2.16), i.e.  $g_N \lambda$  here, it drops out. The somewhat surprising  $1/g_N$  term in  $\beta_\lambda$  comes from a counterterm proportional to the volume but not to  $\Lambda$ . Observe that whenever  $\gamma_1$  is independent of  $g_N$  and  $s$ , the flow equations (2.20, 2.30) are compatible with the existence of a non-Gaussian fixed point,

$$g_N^* = \frac{2}{\gamma_1}, \tag{2.31}$$

with  $\lambda_*$  unspecified. A recent study [54] uses a specific momentum space cutoff  $\mathcal{R}_k$  and evaluates the  $\Gamma_k$  effective average action to one loop using known heat kernel coefficients. The resulting flow equations are of the above form with  $\gamma_1 = \gamma_1(\omega) = (83 + 70\omega + 8\omega^2)/(18\pi)$ ; further two additional terms in  $\beta_\lambda$  are found which fix also the value of  $\lambda_*$  in terms of  $\gamma_1$ . The difference in these non-universal terms can be understood [54] from the fact that dimensional regularization discards quadratic and quartic divergencies, while a momentum space cutoff gives  $\mathcal{R}_k$ -dependent but nonzero results for their coefficients.

The flow equations (2.20, 2.30) of course also admit the Gaussian fixed point  $g_N^* = 0 = \lambda_*$ , and one may be tempted to identify the ‘realm’ of perturbation theory (PT) with the ‘expansion’ around a Gaussian fixed point. As explained in Section 2.1, however, the conceptual status of PT referring to a non-Gaussian fixed point is not significantly different from that referring to a Gaussian fixed point. In other words there is no reason to take the perturbative non-Gaussian fixed point (2.31) any less serious than the perturbative Gaussian one. This important point will reoccur in the framework of the 2 + 2 truncation in Section 3, where a non-Gaussian fixed point is also identified by perturbative means.

The fact that a non-Gaussian fixed point can already be identified in PT is important for several reasons. First, although the value of  $g_N^*$  in Equation (2.31) is always non-universal, the anomalous dimension  $\eta_N = \gamma_g/g_N - 2$  is exactly  $-2$  at the fixed point (2.31). The general argument for the dimensional reduction of the residual interactions outlined after Equation (1.5) can thus already be based on PT alone! Second the result (2.31) suggests that the interplay between the perturbative and the nonperturbative dynamics might be similar to that of non-Abelian gauge theories, where the nonperturbative dynamics is qualitatively and quantitatively important mostly in the infrared.

It is instructive [157] to compare the perturbative one-loop flow (2.20, 2.30) with the linearization of the  $(g_N, \lambda)$  flow obtained from the FRGE framework described in Section 4. In the so-called

Einstein–Hilbert truncation using an optimized cutoff and a limiting version of the gauge-fixing parameter, the ‘beta’ functions  $\gamma_g, \beta_\lambda$  reduce to ratios of polynomials in  $g_N, \lambda$  [136]. Upon expansion to quadratic order one finds

$$\begin{aligned}\mu \frac{d}{d\mu} g_N &= 2g_N - \frac{1}{(4\pi)^2} g_N^2 + O(g_N^2 \lambda), \\ \mu \frac{d}{d\mu} \lambda &= -2\lambda + \frac{1}{(4\pi)^2} \frac{g_N}{2} (1 + 2\lambda) + \frac{5}{12} \frac{1}{(4\pi)^4} g_N^2 + O(g_N^2 \lambda).\end{aligned}\tag{2.32}$$

This is of the form (2.30) at  $s = 0$  with  $\gamma_1 = 1/(4\pi)^2$ , and two additional terms in the second equation. The nontrivial fixed point for  $g_N$  remains of the form (2.31), while the one for  $\lambda$  is best seen in the  $\tau = g_N \lambda$  evolution equation,  $\mu d\tau/d\mu = (1 - 8\lambda^2)g_N^2/(48(4\pi)^4) + O(\lambda^3 g_N^2)$ .

The most important drawback of the perturbatively renormalizable theories based on Equation (2.26) are the problems with unitarity entailed by the propagator (2.28). As already mentioned these problems are absent in an alternative perturbative formulation where a  $1/p^2$  type propagator is used throughout [94]. We now describe this construction in slightly more detail following the presentation in [10].

Starting from the  $d = 4$  Lagrangian  $-\frac{1}{c_d G_N} \sqrt{g} R(g)$  without cosmological constant the one-loop divergencies come out in dimensional regularization as [210]

$$\frac{1}{8\pi^2(4-d)} \sqrt{g} \left( \frac{1}{120} R^2 + \frac{7}{20} R_{\alpha\beta} R^{\alpha\beta} \right).\tag{2.33}$$

They can be removed in two different ways. One is by adding new couplings so that a higher derivative action of the form (2.26) arises with parameters

$$\Lambda = 0, \quad x = 0, \quad z_B = \mu^{d-4} \left( z - \frac{1}{8\pi^2(4-d)} \frac{1}{120} \right), \quad y_B = \mu^{d-4} \left( y - \frac{1}{8\pi^2(4-d)} \frac{7}{20} \right).\tag{2.34}$$

The renormalizability of the resulting theory is mostly due to the modified propagator which can be viewed as a resummed graviton propagator in a power series in  $z, y$ . The unphysical singularities are of order  $1/z, 1/y$ . The second option to remove Equation (2.33) is by a singular field redefinition

$$g_{\alpha\beta} \mapsto g_{\alpha\beta} + \frac{c_d G_N}{8\pi^2(4-d)} \frac{1}{20} \left( -7R_{\alpha\beta} + \frac{11}{3} g_{\alpha\beta} R \right).\tag{2.35}$$

This restores the original  $\sqrt{g} R(g)$  Lagrangian up to two- and higher loop contributions. However this feature is specific to one loop. As shown in [95, 222] at two loops there is a divergence proportional to  $R_{\alpha\beta} \gamma^\delta R^{\alpha\beta}{}_{\rho\sigma} R^{\rho\sigma}{}_{\gamma\delta}$ , which cannot be absorbed by a field redefinition. A counterterm proportional to it must thus be added to  $\sqrt{g} R(g)$ . Importantly, when re-expanded in powers of  $h_{\alpha\beta} = g_{\alpha\beta} - \delta_{\alpha\beta}$ , this counterterm, however, produces only terms quadratic in  $h$  that are proportional to the Ricci tensor or the Ricci scalar. These can be removed by a covariant field redefinition, so that the initial  $1/p^2$  type propagator does not receive corrections. A simple argument [10] shows that this property also holds for all higher order counterterms that can be expected to occur. Explicitly, consider a Lagrangian of the form

$$L = -\frac{1}{c_d G_N} \sqrt{g} R(g) + \sum_{i \geq 1} (c_d G_N)^{\frac{d_i}{2-d} - d} g_i L_i(g),\tag{2.36}$$

where  $L_i(g)$  are local curvature invariants of mass dimension  $-d_i$ , the  $g_i$  are dimensionless couplings, and the power of  $c_d G_N$  (normalized as in Equation (2.15)) gives each term in the sum mass dimension  $-d$ .



Let us briefly recap the power counting and scaling dimensions of local curvature invariants. These are integrals  $P_i[g] = \int d^d x L_i(g)$  over densities  $L_i(g)$  which are products of factors of the form  $\nabla_{\alpha_1} \dots \nabla_{\alpha_{l-4}} R_{\alpha_1 \dots \alpha_l}$ , suitably contracted to get a scalar and then multiplied by  $\sqrt{g}$ . One easily checks  $L_i(\omega^2 g) = \omega^{s_i} L_i(g)$ ,  $\omega > 0$ , with  $s_i = d - 2p - q$ , where  $p$  is the total power of the Riemann tensor and  $q$  is the (necessarily even) total number of covariant derivatives. This scaling dimension matches minus the mass dimension of  $P_i(g)$  if  $g$  is taken dimensionless. For the mass dimension  $d_i$  of the associated coupling  $u_i$  in a product  $u_i P_i[g]$  one thus gets  $d_i = s_i = d - 2p - q$ . For example, the three local invariants in Equation (1.14) have mass dimensions  $-d_0 = -d$ ,  $-d_1 = -(d - 2)$ ,  $-d_2 = -(d - 4)$ , respectively. There are three other local invariants with mass dimension  $-(d - 4)$ , namely the ones with integrands  $C^2 = R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} - 2R^{\alpha\beta} R_{\alpha\beta} + R^2/3$  (the square of the Weyl tensor),  $E = R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} - 4R^{\alpha\beta} R_{\alpha\beta} + R^2$  (the generalized Euler density), and  $\nabla^2 R$ . Then there is a set of dimension  $-(d - 6)$  local invariants, and so on. Note that in  $d = 4$  the integrands of the last two of the dimensionless invariants are total divergencies so that in  $d = 4$  there are only 4 local invariants with non-positive mass dimension (see Equation (2.26)).

A generic term in  $P_i$  will be symbolically of the form  $\nabla^q R^p$ , where all possible contractions of the  $4p + q$  indices may occur. Since the Ricci tensor is schematically of the form  $R = \nabla^2 f + O(f^2)$ , the piece in  $P_i$  quadratic in  $f$  is of the form  $\nabla^{q+4} R^{p-2} f^2$ . The coefficient of  $f^2$  is a tensor with 4 free indices and one can verify by inspection that the possible index contractions are such that the Ricci tensor or Ricci scalar either occurs directly, or after using the contracted Bianchi identity. In summary, one may restrict the sum in Equation (2.36) to terms with  $-d_i = -d + 2p + q$ ,  $p \geq 3$ , and the propagator derived from it will remain of the  $1/p^2$  type to all loop orders. This suggests that Equation (2.36) will give rise to a renormalizable Lagrangian. A proof requires to show that after gauge fixing and ghost terms have been included all counter terms can be chosen local and covariant and has been given in [94].

Translated into Wilsonian terminology the above results then show the existence of a “weakly renormalizable” but “propagator unitary” Quantum Gravidynamics based on a perturbative Gaussian fixed point. The beta functions for this infinite set of couplings are presently unknown. If they were known, expectations are that at least a subset of the couplings would blow up at some finite momentum scale  $\mu = \mu_{\text{term}}$  and would be unphysical for  $\mu > \mu_{\text{term}}$ . In this case the computed results for physical quantities are likely to blow up likewise at some (high) energy scale  $\mu = \mu_{\text{term}}$ . In other words the couplings in Equation (2.36) are presumably not all asymptotically safe.

Let us add a brief comment on the relevant-irrelevant distinction in this context, if only to point out that it is no longer useful. Recall from Section 1.3 that the notion of a relevant or irrelevant coupling applies even to flow lines *not* connected to a fixed point. This is the issue here. All but a few of the interaction monomials in Equation (2.36) are power counting irrelevant with respect to the  $1/p^2$  propagator. Equivalently all but a few couplings  $u_i(\mu) = \mu^{d_i} g_i(\mu)$  have non-negative mass dimensions  $d_i \geq 0$ . These are the only ones not irrelevant with respect to the stability matrix  $\Theta$  computed at the perturbative Gaussian fixed point. However in Equation (2.36) these power counting irrelevant couplings with  $d_i < 0$  are crucial for the absorption of infinities and thus are converted into practically relevant ones. In the context of Equation (2.36) we shall therefore discontinue to use the terms relevant/irrelevant.

Comparing both perturbative constructions one can see that the challenge of Quantum Gravidynamics lies not so much in achieving renormalizability, but to reconcile asymptotically safe couplings with the absence of unphysical propagating modes. This program is realized in Section 3 for the  $2 + 2$  reduction; the results of Section 4 for the  $R + R^2$  type truncation likewise are compatible with the absence of unphysical propagating modes.

In order to realize this program without reductions or truncations, a mathematically controllable nonperturbative definition of Quantum Gravidynamics is needed. Within a functional integral formulation this involves the following main steps: definition of a kinematical measure, setting up a coarse graining flow for the dynamical measures, and then probing its asymptotic safety.

### 2.3.3 Kinematical measure

For a functional integral over geometries even the kinematical measure, excluding the action dependent factor, is nontrivial to obtain. A geometric construction of such a measure has been given by Bern, Blau, and Mottola [31] generalizing a similar construction in Yang–Mills theories [20]. It has the advantage of separating the physical and the gauge degrees of freedom (at least locally in field space) in a way that is not tied to perturbation theory. The functional integral aimed at is one over geometries, i.e. equivalence classes of metrics modulo diffeomorphisms. For the subsequent construction the difference between Lorentzian and Riemannian signature metrics is inessential; for definiteness we consider the Lorentzian case and correspondingly have an action dependence  $\exp iS[g]$  in mind.

A geometry can be described by picking a representative  $\hat{g}_{\alpha\beta}$  described by a  $d(d-1)/2$  parametric metric. Here  $\hat{g}_{\alpha\beta}$  can be specified by picking an explicit parameterization or by imposing a gauge fixing condition  $(F \circ \hat{g})_\alpha = 0$ . Typical choices are a harmonic gauge condition with respect to some reference metric connection, or a proper time gauge  $(F \circ \hat{g})_\alpha = n_\alpha$ , for a fixed timelike co-vector  $n_\alpha$ . Once  $\hat{g}_{\alpha\beta}$  has been fixed, the push forward with a generic diffeomorphism  $V$  will generate the associated orbit,

$$g_{\alpha\beta}(x) = (V_*\hat{g})_{\alpha\beta}(x) = \frac{\partial V^\gamma}{\partial x^\alpha} \frac{\partial V^\delta}{\partial x^\beta} \hat{g}_{\gamma\delta}(V(x)). \quad (2.37)$$

The functional integral over the metrics  $g_{\alpha\beta}$  should factorize into one over the geometries  $\hat{g}$  and one over the diffeomorphisms,

$$\mathcal{D}g_{\alpha\beta} = J(\hat{g}) \mathcal{D}g_{\alpha\beta} \mathcal{D}V_\alpha, \quad (2.38)$$

with a  $V$  independent Jacobian  $J(\hat{g}) = J(g)$ . To compute the Jacobian one views the space of metrics locally as a pseudo-Riemannian manifold  $(\mathcal{M}, \mathcal{G})$  and uses the fact that the Jacobian  $J(g)$  at  $g \in \mathcal{M}$  only contains information about the cotangent space  $\mathcal{H}_T(g)$  at  $g$ . Moreover picking coordinates on the cotangent space the Jacobian produced by a linear change of coordinates in the cotangent space will coincide with the Jacobian induced by a corresponding nonlinear change of coordinates on  $\mathcal{M}$ . In analogy with the finite-dimensional case  $1 = \int d\mu_x(\delta x) \exp\{\frac{i}{2}\delta x^\alpha g_{\alpha\beta}(x)\delta x^\beta\}$ ,  $d\mu_x(\delta x) = \sqrt{\det g(x)}/(2\pi)^{d/2}$ , the measure  $\mathcal{D}g_{\alpha\beta}$  can then be defined by imposing a normalization condition in terms of a Gaussian functional integral on the cotangent space

$$1 = \int \mathcal{D}f_{\alpha\beta} \exp\left\{\frac{i}{2}(f, f)_T\right\}, \quad (2.39)$$

where the cotangent space  $\mathcal{H}_T(g)$  to  $\mathcal{M}$  at  $g$  is equipped with the metric

$$\begin{aligned} \langle f, f \rangle_T &= \int dx \sqrt{-g} f_{\alpha\beta}(x) \mathcal{G}^{\alpha\beta, \gamma\delta}(g) f_{\gamma\delta}(x), \\ \mathcal{G}^{\alpha\beta, \gamma\delta}(g) &= \frac{1}{2}(g^{\alpha\gamma} g^{\beta\delta} + g^{\alpha\delta} g^{\beta\gamma} + C g^{\alpha\beta} g^{\gamma\delta}), \end{aligned} \quad (2.40)$$

and the 1-forms  $f_{\alpha\beta} := \delta g_{\alpha\beta}$  span  $\mathcal{H}_T(g)$  (with ‘T’ mnemonic for “tensor”). The measure defined by Equation (2.39) will be formally diffeomorphism invariant provided the metric on  $\mathcal{M}$  is covariant. Requiring the metric to be ultralocal in addition fixes  $\mathcal{G}$  to be of the above “deWitt” form up to an overall normalization and the undetermined constant  $C$ . The latter determines the signature of the metric on  $\mathcal{M}$ , it is of Riemannian type for  $C > -d/2$ , Lorentzian for  $C < -d/2$ , and degenerate for  $C = -d/2$ . Here we take  $C \neq -d/2$  but leave  $C$  unspecified otherwise.

On the tangent space the parameterization (2.37) amounts to

$$f_{\alpha\beta} = (Lv)_{\alpha\beta} + (V_*\hat{f})_{\alpha\beta}, \quad (2.41)$$

where  $V^\alpha(x) = x^\alpha + v^\alpha(x) + O(v^2)$  and  $v^\alpha \partial_\alpha$  is the associated vector field. Further the Lie derivative  $\mathcal{L}_v g_{\alpha\beta} =: (Lv)_{\alpha\beta}$  is for the given  $\hat{g}$  regarded a linear map  $L$  from vectors to symmetric tensors. Its kernel are the Killing vectors of  $g$ , while the kernel of its adjoint  $L^\dagger$  describes the genuine variations in the geometry. Here

$$L^\dagger : \mathcal{H}_T(g) \longrightarrow \mathcal{H}_V(g), \quad (v, v)_V := \int dx \sqrt{g} v_\alpha(x) g^{\alpha\beta}(x) v_\beta(x) \quad (2.42)$$

maps symmetric tensors to (co-)vectors, the space of which is equipped with the obvious invariant inner product. With the normalization  $(Lv, Lv)_T = (v, L^\dagger Lv)_V$  the vector Laplacian  $L^\dagger L$  is given by

$$(L^\dagger L)_\alpha^\beta = -2 [\delta_\alpha^\beta \nabla^2 + (1 + C) \nabla_\alpha \nabla^\beta + R_\alpha^\beta(g)]. \quad (2.43)$$

Under mild technical conditions  $L^\dagger L$  will have a well-defined inverse,  $(L^\dagger L)^{-1} : \mathcal{H}_V(g) \rightarrow \mathcal{H}_T(g)$ . This allows one to replace Equation (2.41) by a decomposition orthogonal with respect to  $(\cdot, \cdot)_T$ , viz.

$$\begin{aligned} f_{\alpha\beta} &= (Lv')_{\alpha\beta} + (PV_* \hat{f})_{\alpha\beta}, & v' &:= v + (L^\dagger L)^{-1} L^\dagger \hat{f}, \\ P &:= \mathbb{1} - L(L^\dagger L)^{-1} L^\dagger, & P^2 &= P, \quad PL = 0. \end{aligned} \quad (2.44)$$

Using  $\mathcal{D}f = J(\hat{g}) \mathcal{D}v \mathcal{D}\hat{f}$  and Equation (2.39) the computation of the Jacobian in Equation (2.38) then reduces to that of two Gaussians

$$1 = J(\hat{g}) \int \mathcal{D}v \exp \left\{ \frac{i}{2} (v, L^\dagger Lv)_V \right\} \int \mathcal{D}\hat{f}_{\alpha\beta} \exp \left\{ \frac{i}{2} (\hat{f}, P\hat{f})_T \right\}, \quad (2.45)$$

which, suitably regularized, we take as the definition of  $J(\hat{g})$ . If  $\hat{g}$  is defined through a gauge fixing condition  $(F \circ \hat{g})_\alpha = 0$  the result (2.45) can be rewritten as [31]

$$J(\hat{g}) = [\det_\sqrt{V}(F \circ F^\dagger)]^{-1/2} \det_\sqrt{V}(F \circ L). \quad (2.46)$$

The subscript  $V$  denotes a vector determinant defined by  $\det_V W = \exp\{-\sum_{k \geq 1} \frac{1}{k} \text{Tr}_V(1 - W)^k\}$ , where  $W$  maps vectors to vectors and the trace refers to the inner product in Equation (2.42). We remark that the second factor in Equation (2.46) is the Faddeev–Popov determinant for the gauge  $(F \circ \hat{f})_\alpha = 0$ , while the first factor is an  $\hat{f}_{\alpha\beta}$  independent normalization factor. Within perturbation theory the above construction is equivalent to the familiar BRST formulation with ghosts.

In summary one arrives at the following proposal for a kinematical measure over geometries:

$$[\text{Vol Diff}]^{-1} \int_{\mathcal{M}} \mathcal{D}g_{\alpha\beta} e^{iS[g]} \quad \mapsto \quad \int_{\mathcal{M}/\text{Diff}} \mathcal{D}\hat{g}_{\alpha\beta} \det_\sqrt{V}(F \circ L)(\hat{g}) e^{iS[\hat{g}]}. \quad (2.47)$$

Here we omitted the normalization factor and for illustration included the factor  $\exp(iS[g])$  (with an invariant action,  $S[V_* \hat{g}] = S[\hat{g}]$ ) that would specify the dynamical measure. The kinematical measure on the right-hand-side can also directly be verified to be diffeomorphism invariant and is hence well-defined on the equivalence classes. Of course the latter presupposes an invariant regulator which is why Equation (2.47) can only serve as a useful guideline. There are two ways to proceed from Equation (2.47). One would work with a noninvariant regulator, maintain the original notion of diffeomorphisms, and use conventional field theoretical techniques to restore the diffeomorphism invariance through Ward identities at the end. Alternatively one can replace the right-hand-side of Equation (2.47) directly with a discretized version, in which case of course diffeomorphism invariance cannot be tested on this level. Both strategies are complementary and have been widely used. For completeness let us also mention at this point the well-known feature

of functional integrals that once the regulator is removed the kinematical measure and the action factor do not have a mathematical meaning individually. In an interacting theory this is also related to the renormalization problem.

In the above discussion we did not split off the conformal factor in the geometries. Doing this however only requires minor modifications and was the setting used in [31, 142, 149]. In Equation (2.37) then  $\hat{g}_{\alpha\beta}$  is written as  $e^\sigma g_{\alpha\beta}^\perp$ , where now  $g_{\alpha\beta}^\perp$  is subject to a gauge condition  $(F \circ g^\perp)_\alpha = 0$ . On the cotangent space this leads to a York-type decomposition [235] replacing (2.41), where the variations  $f_\sigma$  of the conformal factor and that of the tracefree part  $f_{\alpha\beta}^\perp$  of  $\hat{f}_{\alpha\beta}$  describe the variations of the geometry, while the tracefree part,  $(L^{\text{TF}}v)_{\alpha\beta} := \nabla_\alpha v_\beta + \nabla_\beta v_\alpha - \frac{2}{d}g_{\alpha\beta}\nabla^\gamma v_\gamma$ , and the trace part of the Lie derivative  $(Lv)_{\alpha\beta}$  describe the gauge variations. Writing  $\mathcal{D}f_{\alpha\beta} = \mathcal{D}f_\sigma \mathcal{D}f_{\alpha\beta}^\perp \mathcal{D}v$  the computation of the Jacobian proceeds as above and leads to Equation (2.47) with the following replacements:  $\mathcal{D}\hat{g}_{\alpha\beta}$  is replaced with  $\mathcal{D}g_{\alpha\beta}^\perp \mathcal{D}\sigma$ ,  $L$  with  $L^{\text{TF}}$ , and  $\hat{g}$  with  $e^\sigma g^\perp$  in the integrand. By studying the dependence of  $\det_V(F \circ L)(e^\sigma g^\perp)$  on the conformal factor it has been shown in [142] that in the Gaussian approximation of the Euclidean functional integral the instability associated with the unboundedness of the Euclidean Einstein–Hilbert action is absent, due to a compensating contribution from the determinant. It can be argued that this mechanism is valid also for the interacting theory. From the present viewpoint however the (Euclidean or Lorentzian) Einstein–Hilbert action should not be expected to be the proper microscopic action. So the “large field” or “large gradient” problem has to be readdressed anyhow in the context of Quantum Gravidynamics. Note also that once the conformal mode of the metric has been split off the way how it enters a microscopic or an effective action is no longer constrained by power counting considerations. See [12] for an effective dynamics for the conformal factor only.

Once a kinematical measure on the equivalence classes of metrics (or other dynamical variables) has been defined, the construction of an associated dynamical measure will have to rest on renormalization group ideas. Apart from the technical problems involved in setting up a computationally useful coarse graining flow for the measure on geometries, there is also the apparent conceptual problem how diffeomorphism invariance can be reconciled with the existence of a scale with respect to which the coarse graining is done. However no problem of principle arises here. First, similar as in a lattice field theory, where one has to distinguish between the external lattice spacing and a dynamically generated correlation length, a distinction between an external scale parameter and a dynamically co-determined resolution scale has to be made. A convenient way to achieve compatibility of the coarse graining with diffeomorphism invariance is by use of the background field formalism. The initially generic background metric serves as a reference to discriminate modes, say in terms of the spectrum of a covariant differential operator in the background metric (see Section 4.1). Subsequently the background is self-consistently identified with the expectation value of the quantum metric as in the discussion below.

The functional integral over “all geometries” should really be thought of as one over “all geometries subject to suitable boundary conditions”. Likewise the action is meant to include boundary terms which indirectly specify the state of the quantum system.

After a coarse graining flow for the dynamical measures has been set up the crucial issue will be whether or not it has a fixed point with a nontrivial finite-dimensional unstable manifold, describing an interacting system. In this case it would define an asymptotically safe functional measure in the sense defined in Section 1.3. For the reasons explained there the existence of an asymptotically safe functional integral measure is however neither necessary nor sufficient for a physically viable theory of Quantum Gravidynamics. For the latter a somewhat modified notion of a safe functional measure is appropriate which incorporates the interplay between couplings and observables:

1. The choice of couplings has to be based on observables; this will pin down the physically relevant notion of positivity/unitarity.
2. The number of essential or relevant couplings is not a-priori finite.

3. What matters is not so much the dimension of the unstable manifold than how observables depend on the relevant couplings.

### 2.3.4 Effective action and states

Unfortunately, at present little is known about generic quantum gravity observables, so that the functional averages whose expansion would define physical couplings are hard to come by. For the time being we therefore adopt a more pragmatic approach and use as the central object to formulate the renormalization flow the background effective action  $\Gamma[g_{\alpha\beta}, \bar{g}_{\alpha\beta}, \dots]$  as described in Appendix B. Here  $g_{\alpha\beta}$  is interpreted as an initially source-dependent “expectation value of the quantum metric”,  $\bar{g}_{\alpha\beta}$  is an initially independently prescribed “background metric”, and the dots indicate other fields, conjugate to sources, which are inessential for the following discussion. For clarity sake let us add that it is *not* assumed that the metric exists as an operator, or that the metric-like “conjugate sources”  $g_{\alpha\beta}, \bar{g}_{\alpha\beta}$  are necessarily the best choice.

The use of an initially generic background geometry  $\bar{g}_{\alpha\beta}$  has the advantage that one can define propagation and covariant mode-cutoffs with respect to it. A background effective action of this type has an interesting interplay with the notion of a state [156, 157]. An effective action implicitly specifies an expectation functional  $\mathcal{O} \mapsto \langle \mathcal{O} \rangle_{\bar{g}}$  (“a state”) which depends parameterically on the background metric. The background metric is then self-consistently identified with the expectation value of the metric

$$\langle g_{\alpha\beta} \rangle_{\bar{g}_*[g]} = \bar{g}_{\alpha\beta}, \quad (2.48)$$

in a way that encodes information about the full quantum dynamics. Importantly this self-consistent background  $\bar{g}_*[g]_{\alpha\beta}$  is no longer prescribed externally, and to the extent one has access to nontrivial solutions  $\bar{g}_{\alpha\beta} = \bar{g}_*[g]_{\alpha\beta}$ , Equation (2.49) gives rise to a formulation with a ‘state-dependent dynamically adjusted reference metric’. The functional  $g \mapsto \bar{g}_*[g]$  is defined by

$$\frac{\delta}{\delta g_{\alpha\beta}} \Gamma[g, \bar{g}, \dots] = 0 \quad \text{iff} \quad \bar{g}_{\alpha\beta} = \bar{g}_*[g]_{\alpha\beta}. \quad (2.49)$$

We postpone the question how solutions of Equation (2.49) can be found. Note that after the identification has been made  $\bar{\Gamma}[g, \dots] := \Gamma[g, \bar{g}_*[g], \dots]$  is a functional of a single metric only, which obeys

$$\frac{\delta \bar{\Gamma}[g, \dots]}{\delta g_{\alpha\beta}} = 0. \quad (2.50)$$

Vertex functions are defined by functional differentiation at fixed background with subsequent identification (2.48), i.e.

$$\Gamma^{(n)}(x_1, \dots, x_n; g) := \frac{\delta}{\delta g(x_1)} \dots \frac{\delta}{\delta g(x_n)} \Gamma[g, \bar{g}, \dots] \Big|_{\bar{g}=\bar{g}_*[g]}. \quad (2.51)$$

The set of these vertex functions in principle contains the same information as the original functional measure including the state. One would expect them to be related to S-matrix elements on a self-consistent background (2.48), but their precise physics significance remains to be understood.

The condition (2.49) is equivalent to the vanishing of the extremizing sources  $J_*^{\alpha\beta}[g, \bar{g}, \dots]$  in the definition of Legendre transform (see Appendix B). Evidently Equation (2.49) also amounts to the vanishing of the one-point functions in Equation (2.51). Usually the extremizing sources  $J_*^{\alpha\beta}[g, \bar{g}, \dots]$  are constructed by formal inversion of a power series in  $\bar{f}_{\alpha\beta} := g_{\alpha\beta} - \bar{g}_{\alpha\beta}$ . Then  $\bar{f}_{\alpha\beta} = 0$  always is a solution of Equation (2.49) and the functional  $g \mapsto \bar{g}_*[g]$  is simply the identity. In this case the self-consistent background coincides with the naive prescribed background. To find nontrivial solutions of Equation (2.49) one has to go beyond the formal series inversions and the uniqueness assumptions usually made.

Due to the highly nonlocal character of the effective action the identification of *physical* solutions of Equation (2.49) is a nontrivial problem. The interpretation via Equation (2.48) suggests an indirect characterization, namely those solutions of Equation (2.49) should be regarded as physical which come from physically acceptable states [157].

The notion of a state is implicitly encoded in the effective action. Recall that the standard effective action, when evaluated at a given time-independent function  $\phi^i = \langle \chi^i \rangle$ , is proportional to the minimum value of the Hamiltonian  $H$  in that part of the Hilbert space spanned by normalizable states  $|\psi\rangle$  satisfying  $\langle \psi | \chi^i | \psi \rangle = \phi^i$ . A similar interpretation holds formally for the various background effective actions [51]. In conventional quantum field theories there is a clear-cut notion of a ground state and of the state space based on it. In a functional integral formulation the information about the state can be encoded in suitable boundary terms for the microscopic action. Already in quantum field theories on curved but non-dynamical spacetimes a preferred vacuum is typically absent and physically acceptable states have to be selected by suitable conditions (like, for example, the well-known Hadamard condition in the case of a Klein–Gordon field). In quantum gravity the formulation of analogous selection criteria is an open problem. As a tentative example we mention the condition formulated after Equation (2.53) below. On the level of the effective action one should think of  $\Gamma$  as a functional of both the selected state and of the fields. The selected state will indirectly (co-)determine the space of functionals on which the renormalization flow acts. For example the type of nonlocalities which actually occur in  $\Gamma$  should know about the fact that  $\Gamma$  stems from a microscopic action suited for the appropriate notion of positivity and from a physically acceptable state.

### 2.3.5 Towards physical quantities

Finally one will have to face the question of what generic physical quantities are and how to compute them. Although this is of course a decisive issue in any approach to quantum gravity, surprisingly little work has been done in this direction. In classical general relativity Dirac observables do in principle encode all intrinsic properties of the spacetimes, but they are nonlocal functionals of the metric and implicitly refer to a solution of the Cauchy problem. In a canonical formulation quantum counterparts thereof should generate the physical state space, but they are difficult to come by, and a canonical formulation is anyhow disfavored by the asymptotic safety scenario. S-matrix elements with respect to a self-consistent background (2.48) or similar objects computed from the vertex functions (2.51) might be candidates for generic physical quantities, but have not been studied so far.

For the time being a pragmatic approach is to consider quantities which are of interest in a quantum field theory on a fixed but generic geometry and then perform an average over geometries with the measure previously constructed. On a perturbative level interesting possible effects have been studied in [202, 215]. On a nonperturbative level this type of correlations have been discussed mostly in discretized formulations but the principle is of course general. To fix ideas we note the example of a geodesic two point correlator of a scalar field  $\phi$  [58],

$$G(R) = \int \mathcal{D}g \mathcal{D}\phi e^{iS[g,\phi]} \int dx dy \sqrt{g(x)} \sqrt{g(y)} \phi(x) \phi(y) \delta(\Sigma_g(x, y) - R), \quad (2.52)$$

where  $\Sigma_g(x, y)$  is the minimal geodesic distance between the the points  $x$  and  $y$ . The first integral is the heuristic geometry and matter functional integral; all configurations are taken into account which produce the given geodesic distance  $R$ . A nontrivial prediction of the present scenario is that if Equation 2.52) is based on an asymptotically safe functional measure, a  $\log(R)$  behavior for  $R \rightarrow 0$  is expected [157].

If one wants to probe the functional measure over geometries only, an interesting operator



insertion is the trace of the heat kernel [115, 118, 8],

$$G(T) = \int \mathcal{D}g e^{iS[g]} \int dx \sqrt{g(x)} \exp(T\Delta_g)(x, x). \quad (2.53)$$

Here  $\Delta_g := g^{\alpha\beta} \nabla_\alpha \nabla_\beta = \sqrt{g}^{-1} \partial_\alpha (\sqrt{g} g^{\alpha\beta} \partial_\beta)$  is the Laplace–Beltrami operator, and the heat kernel  $\exp(T\Delta_g)(x, x')$  associated with it is the symmetric (in  $x, x'$ ) bi-solution of the heat equation  $\partial_T K = \Delta_g K$  with initial condition  $\lim_{T \rightarrow 0} \exp(T\Delta_g)(x, x') = \delta(x, x')$ . The  $T \rightarrow \infty$  limit will then probe the large scale structure of the typical geometries in the measure and the  $T \rightarrow 0$  limit will probe the micro aspects. Both expressions (2.52, 2.53) are here only heuristic, in particular normalization factors have been omitted and the functional measure over geometries would have to be defined as previously outlined. The condition that the  $T \rightarrow \infty$  behavior of  $G(T) = \langle P_g(T) \rangle$  is like that in flat space,  $\langle P_g(T) \rangle \sim T^{-d/2}$  for  $T \rightarrow \infty$ , is an example for a (rather weak) selection criterion for states [157]. The states obeying it should favor geometries that are smooth and almost flat on large scales.

In a lattice field theory the discretized functional measure typically generates an intrinsic scale, the (dimensionless) correlation length  $\xi$ , which allows one to convert lattice distances into a physical standard of length, such that say,  $\xi$  lattice spacings equal 1 fm. A (massive) continuum limit is eventually defined by sending  $\xi$  to infinity in a way such that physical distances (number of lattice spacings)/ $\xi$  fm are kept fixed and a ‘nonboring’ limit arises. In a functional measure over geometries  $d\mu_{k,\Lambda}(g)$ , initially defined with an UV cutoff  $\Lambda$  and an external scale parameter  $k$ , it is not immediate how to generalize the concept of a correlation length. Exponents extracted from the decay properties of Equation (2.52) or Equation (2.53) are natural candidates, but the ultimate test of the fruitfulness of such a definition would lie in the successful construction of a continuum limit. In contrast to a conventional field theory it is not even clear what the desired/required properties of such a continuum system should be. The working definitions proposed in Section 1.3 tries to identify some salient features.

## 2.4 Dimensional reduction of residual interactions in UV

As highlighted in the introduction an important qualitative feature of an asymptotically safe functional integral can be inferred without actually evaluating it, namely that in the extreme ultraviolet the residual interactions appear two-dimensional. There are a number of interconnected heuristic arguments for this phenomenon which we present here.

### 2.4.1 Scaling of fixed point action

Consider a candidate for a quasilocal microscopic action

$$S_k[g] = \sum_i u_i(k) P_i[g], \quad (2.54)$$

where the  $u_i(k)$  are running couplings of mass dimension  $d_i$  and  $P_i[g]$  are local invariants of mass dimension  $-d_i$ . By quasilocal we mean here that the sum may be infinite and off hand arbitrary high derivative terms may occur. For example such an action arises in the perturbative framework advocated by Gomis and Weinberg [94]. When viewed as a renormalized action perturbatively defined in the above sense (with the UV cutoff strictly removed) the running of the  $u_i^{\text{PT}}(k)$  is unknown but expectations are that  $g_i^{\text{PT}}(k) = k^{-d_i} u_i^{\text{PT}}(k)$  are not uniformly bounded functions in  $k$ ; then the dimensionless couplings are not asymptotically safe but blow up at various ( $i$ -dependent) intermediate scales. The situation is drastically different if all the couplings are assumed to be asymptotically safe. Then  $u_i(k) = g_i(k) k^{d_i} \sim g_i^* k^{d_i}$  as  $k \rightarrow \infty$  and if one uses the fact that  $s_i = d_i$

(see the discussion after Equation (2.36)) for all local invariants one gets

$$S_k[g] \sim \sum_i g_i^* P_i[k^2 g] = S_*[k^2 g] \quad (2.55)$$

for  $k \rightarrow \infty$ , with  $S_*[g] = \sum_i g_i^* S_i[g]$  the candidate fixed point action. The overall scale of the metric is an inessential parameter (see Equations (1.1, A.8)), and as discussed in Section 2.1 a fixed point action always refers to an equivalence class modulo possibly running inessential parameters.

One sees that in the fixed point regime  $g_i(k) \sim g_i^*$  the  $k$ -dependence enters only through the combination  $k^2 g_{\alpha\beta}$ , a kind of self-similarity. This simple but momentous fact eventually underlies all the subsequent arguments. It is ‘as if’ in the fixed point regime only a rescaled metric  $\tilde{g}_{\alpha\beta} = k^2 g_{\alpha\beta}$  entered which carries dimension two. This has consequences for the ‘effective dimensionality’ of Newton’s constant: Recall that conventionally the Ricci scalar term,  $\int dx \sqrt{g} R(g)$ , has mass dimension  $2 - d$  in  $d$  dimensions. Upon substitution  $g_{\alpha\beta} \mapsto \tilde{g}_{\alpha\beta}$  one quickly verifies that  $\int dx \sqrt{\tilde{g}} R(\tilde{g})$  is dimensionless. Its prefactor, i.e. the inverse of Newton’s constant, then can be taken dimensionless – as it is in two dimensions. Compared to the infrared regime it looks ‘as if’ Newton’s constant changed its effective dimensionality from  $d - 2$  to zero, i.e. at the fixed point there must be a large anomalous dimension  $\eta_N = 2 - d$ .

Formally what is special about the Einstein–Hilbert term is that the kinetic (second derivative) term itself carries a dimensionful coupling. To avoid the above conclusion one might try to assign the metric a mass dimension 2 from the beginning (i.e. not just in the asymptotic regime). However this would merely shift the effect from the gravity to the matter sector, as we wish to argue now.

In addition to the dimensionful metric  $\tilde{g}_{\alpha\beta} := k^2 g_{\alpha\beta}$ , we introduce a dimensionful vielbein by  $\tilde{E}_\alpha^m := k E_\alpha^m$ , if  $g_{\alpha\beta} = E_\alpha^m E_\beta^n \eta_{mn}$  is the dimensionless metric. With respect to a dimensionless metric  $\int dx \sqrt{g} R(g)$  has mass dimension  $2 - d$  in  $d$  dimensions, while the mass dimensions  $d_\chi$  of a Bose field  $\chi$  and that  $d_\psi$  of a Fermi field  $\psi$  are set such that their kinetic terms are dimensionless, i.e.  $d_\chi = (d - 2)/2$  and  $d_\psi = (d - 1)/2$ . Upon substitution  $g_{\alpha\beta} \mapsto \tilde{g}_{\alpha\beta}$  the gravity part  $\int dx \sqrt{\tilde{g}} R(\tilde{g})$  becomes dimensionless, while the kinetic terms of a Bose and Fermi field pick up a mass dimension of  $d - 2$  and  $d - 1$ , respectively. This means their wave function renormalization constants  $Z_\chi(k)$  and  $Z_\psi(k)$  are now dimensionful and should be written in terms of dimensionless parameters as  $Z_\chi(k) = k^{d-2}/g_\chi(k)$  and  $Z_\psi(k) = k^{d-1}/g_\psi(k)$ , say. For the dimensionless parameters one expects finite limit values  $\lim_{k \rightarrow \infty} g_\chi(k) = g_\chi^* > 0$  and  $\lim_{k \rightarrow \infty} g_\psi(k) = g_\psi^* > 0$ , since otherwise the corresponding (free) field would simply decouple. Defining the anomalous dimension as usual.  $\eta_\chi = -k \partial_k \ln Z_\chi$  and  $\eta_\psi = -k \partial_k \ln Z_\psi$ , the argument presented after Equation (1.5) can be repeated and gives that  $\eta_\chi^* = 2 - d$ ,  $\eta_\psi^* = 1 - d$  for the fixed point values, respectively. The original large momentum behavior  $1/p^2$  for bosons and  $1/p$  for fermions is thus modified to a  $1/p^d$  behavior in the fixed point regime, in both cases.

This translates into a logarithmic short distance behavior which is universal for all (free) matter. Initially the propagators used here should be viewed as “test propagators”, in the sense that one transplants the information in the  $\eta$ ’s derived from the gravitational functional integral into a conventional propagator on a (flat or curved) background spacetime. Since the short distance asymptotics is the same on any (flat or curved) reference spacetime, this leads to the prediction anticipated in Section 2.3: The short distance behavior of the quantum gravity average of the geodesic two-point correlator (2.52) of a scalar field should be logarithmic.

On the other hand the universality of the logarithmic short distance behavior in the matter propagators also justifies to attribute the phenomenon to a modification in the underlying random geometry, a kind of “quantum equivalence principle”.

### 2.4.2 Anomalous dimension at non-Gaussian fixed point

The “anomalous dimension argument” has already been sketched in the introduction. Here we present a few more details and relate it to Section 2.4.1.

Suppose again that the unknown microscopic action of Quantum Gravidynamics is quasilocal and reparameterization invariant. The only term containing second derivatives then is the familiar Einstein–Hilbert term  $Z_N \int dx \sqrt{g} R(g)$  of mass dimension  $2 - d$  in  $d$  dimensions, if the metric is taken dimensionless. As explained in Section 2.3.2 the dimensionful running prefactor multiplying it  $Z_N(k)$  ( $N$  for “Newton”) can be treated either as a wave function renormalization or as a quasi-essential dimensionless coupling  $g_N$ , where

$$c_d G_{\text{Newton}} = Z_N(k)^{-1} = g_N(k) k^{2-d}. \quad (2.56)$$

Here we treat  $g_N$  as running, in which case its running may also be affected by all the other couplings (gravitational and non-gravitational, made dimensionless by taking out a suitable power of  $k$ ). The short distance behavior of the propagator will now be governed by the “anomalous dimension”  $\eta_N = -k \partial_k \ln Z_N(k)$  by general field theoretical arguments. On the other hand the flow equation for  $g_N$  can be expressed in terms of  $\eta_N$  as  $k \partial_k g_N = [d - 2 + \eta(g_N, \text{other})] g_N$ , where we schematically indicated the dependence on the other dimensionless couplings. If this flow equation now has a nontrivial fixed point  $\infty > g_N^* > 0$ , the only other way how the right-hand-side can vanish is for

$$\eta_N(g_N^*, \text{other}) = 2 - d, \quad (2.57)$$

irrespective of the detailed behavior of the other couplings as long as no blow-up occurs. This is a huge anomalous dimension. We can now transplant this anomalous dimension into a “test graviton propagator” on a flat background. The characteristic property of  $\eta_N$  then is that it gives rise to a high momentum behavior of the form  $(p^2)^{-1+\eta_N/2}$  modulo logarithms, or a short distance behavior of the form  $(\sqrt{x^2})^{2-d-\eta_N}$  modulo logarithms. This follows from general field theoretical principles: a Callan–Symanzik equation for the effective action, the vanishing of the beta function at the fixed point, and the decoupling of the low momentum modes. Keeping only the leading part the vanishing power at  $\eta_N = 2 - d$  translates into a logarithmic behavior,  $\ln x^2$ , formally the same as for a massless Klein–Gordon field in a two-dimensional field theory.

The fact that a large anomalous dimension occurs at a non-Gaussian fixed point was initially observed in the context of the  $2 + \epsilon$  expansion [116, 117] and later in computations based on the effective average action [133, 131]. The above argument shows that no specific computational information enters.

Let us emphasize that in general an anomalous dimension is *not* related to the geometry of field propagation and in a conventional field theory one cannot sensibly define a fractal dimension by looking at the high momentum behavior of a two-point function [125]. What is special about gravity is ultimately that the propagating field itself defines distances. One aspect thereof is the universal way matter is affected, as seen in Section 2.4.1. In contrast to an anomalous dimension in conventional field theories, this “quantum equivalence principle” allows one to attribute a geometric significance to the modified short distance behavior of the test propagators, see Section 2.4.4.

### 2.4.3 Strict renormalizability and $1/p^4$ propagators

With hindsight the above patterns are already implicit in earlier work on strictly renormalizable gravity theories. As emphasized repeatedly the benign renormalizability properties of higher derivative theories are mostly due to the use of  $1/p^4$  type propagator (in  $d = 4$  dimensions). As seen in Section 2.3.2 this  $1/p^4$  type behavior goes hand in hand with asymptotically safe couplings. Specifically for the dimensionless Newton’s constant  $g_N$  it is compatible with the existence of a

nontrivial fixed point (see Equation (2.31)). This in turn enforces anomalous dimension  $\eta_N = -2$  at the fixed point which links back to the  $1/p^4$  type propagator.

Similarly in the  $1/N$  expansion [216, 217, 203] a nontrivial fixed point goes hand in hand with a propagator whose high momentum behavior is of the form  $1/(p^4 \ln p^2)$  in four dimensions, and formally  $1/p^d$  in  $d$  dimensions. In position space this amounts to a  $\ln x^2$  behavior, once again.

#### 2.4.4 Spectral dimension and scaling of fixed point action

The scaling (2.55) of the fixed point action also allows one to estimate the behavior of the spectral dimension in the ultraviolet. This leads to a variant [157] of the argument first used in [135, 134]).

Consider the quantum gravity average  $\langle P_g(T) \rangle$  over the trace of the heat kernel  $P_g(T)$  in a class of states to be specified later. Morally speaking the functional average is over compact closed  $d$ -dimensional manifolds  $(\mathcal{M}, g)$ , and the states are such that they favor geometries which are smooth and approximately flat on large scales.

Let us briefly recapitulate the definition of the heat kernel and some basic properties. For a smooth Riemannian metric  $g$  on a compact closed  $d$ -manifold let  $\Delta_g := g^{\alpha\beta} \nabla_\alpha \nabla_\beta = \sqrt{g}^{-1} \partial_\alpha (\sqrt{g} g^{\alpha\beta} \partial_\beta)$  be the Laplace–Beltrami operator. The heat kernel  $\exp(T\Delta_g)(x, x')$  associated with it is the symmetric (in  $x, x'$ ) bi-solution of the heat equation  $\partial_T K = \Delta_g K$  with initial condition  $\lim_{T \rightarrow 0} \exp(T\Delta_g)(x, x') = \delta(x, x')$ . Since  $(\mathcal{M}, g)$  is compact,  $\Delta_g$  has purely discrete spectrum with finite multiplicities. We write  $-\Delta_g \phi_n(g) = \mathcal{E}_n(g) \phi_n(g)$ ,  $n \geq 0$ , for the spectral problem and assume that the eigenfunctions  $\phi_n$  are normalized and the eigenvalues monotonically ordered,  $\mathcal{E}_n(g) \leq \mathcal{E}_{n+1}(g)$ . We write  $V(g) = \int dx \sqrt{g}$  for the volume of  $(\mathcal{M}, g)$  and

$$P_g(T) = \frac{1}{V(g)} \int dx \sqrt{g} \exp(T\Delta_g)(x, x) = \frac{1}{V(g)} \sum_n e^{-\mathcal{E}_n(g)T}, \quad (2.58)$$

for the trace of the heat kernel. In the random walk picture  $P_g(T)$  can be interpreted as the probability of a test particle diffusing away from a point  $x \in \mathcal{M}$  and to return to it after the fictitious diffusion time  $T$  has elapsed. In flat Euclidean space  $(\mathcal{M}, g) = (\mathbb{R}^d, \eta)$  for example  $P_\eta(T) = (4\pi T)^{-d/2}$  for all  $T$ . For a generic manifold the trace of the heat kernel cannot be evaluated exactly. However the short time and the long time asymptotics can to some extent be described in closed form. Clearly the  $T \rightarrow \infty$  limit probes the large scale structure of a Riemannian manifold (small eigenvalues  $\mathcal{E}_n(g)$ ) while the  $T \rightarrow 0$  limit probes the small scales (large eigenvalues  $\mathcal{E}_n(g)$ ).

For  $T \rightarrow 0$  one has an asymptotic expansion  $P_g(T) \sim (4\pi T)^{-d/2} \sum_{n \geq 0} T^n \int dx \sqrt{g} a_n(x)$ , where the  $a_n$  are the Seeley–deWitt coefficients. These are local curvature invariants,  $a_0 = 1$ ,  $a_1 = \frac{1}{6}R(g)$ , etc. The series can be rearranged so as to collect terms with a fixed power in the curvature or with a fixed number of derivatives [225, 17]. Both produces nonlocal curvature invariants. The second rearrangement is relevant when the curvatures are small but rapidly varying (so that the derivatives of the curvatures are more important than their powers). The leading derivative terms then are given by  $P_g(T) \sim (4\pi T)^{-d/2} [V(g) + T \int dx \sqrt{g} a_1 + T^2 N_2(T) + \dots]$ , where  $N_2(T)$  is a known nonlocal quadratic expression in the curvature tensors (see e.g. [225, 17] for surveys). The  $T \rightarrow \infty$  behavior is more subtle as also global information on the manifold enters. For compact manifolds a typical behavior is  $P_g(T) \sim (4\pi T)^{-d/2} [1 + O(\exp(-cT))]$ , where the rate of decay  $c$  of the subleading term is governed by the smallest non-zero eigenvalue.

Returning now to the quantum gravity average  $\langle P_g(T) \rangle$ , one sees that on any state on which all local curvature polynomials vanish the leading short distance behavior of  $\langle P_g(T) \rangle$  will always be  $\sim T^{-d/2}$ , as on a fixed manifold. The same will hold if the nonlocal invariants occurring in the derivative expansion all have vanishing averages in the state considered. A leading short distance behavior of the form

$$\langle P_g(T) \rangle \sim T^{-d_s/2}, \quad T \rightarrow 0, \quad (2.59)$$

with  $d_s \neq d$  will thus indicate that either the operations “taking the average” and “performing the asymptotic expansion for  $T \rightarrow 0$ ” no longer commute, or that the microscopic geometry is very rough so that the termwise averages no longer vanish, or both. Whenever well-defined the quantity  $d_s(T) := -2d \ln \langle P_g(T) \rangle / d \ln T$  is known as the spectral dimension (of the micro-aspects of the random geometries probed by the state  $\mathcal{O} \mapsto \langle \mathcal{O} \rangle$ ). See [115, 118, 8] for earlier uses in random geometry, and [27] for an evaluation of the spectral dimension for diffusion on the Sierpinski gasket based on a principle similar to Equation (2.62) below.

We assume now that the states considered are such that the  $T \rightarrow \infty$  behavior of  $\langle P_g(T) \rangle$  is like that in flat space, i.e.  $\langle P_g(T) \rangle \sim T^{-d/2}$  for  $T \rightarrow \infty$ . This is an indirect characterization of a class of states which favor geometries that are smooth and almost flat on large scales [157]. (A rough analogy may be the way how the short-distance Hadamard condition used for free QFTs in curved spacetime selects states with desirable stability properties.) Recall that in a functional integral formulation the information about the state can be encoded in suitable boundary terms added to the microscopic action. The effective action used in a later stage of the argument is supposed to be one which derives from a microscopic action in which suitable (though not explicitly known) boundary terms encoding the information about the state have been included.

Since  $(4\pi T)^{-d/2} = \int \frac{d^d p}{(2\pi)^d} \exp(-p^2 T)$  one can give the stipulated  $T \rightarrow \infty$  asymptotics an interpretation in terms of the spectrum  $\{p^2, p \in \mathbb{R}^d\}$  of the Laplacian of a ‘typical’ reference metric  $q_{\alpha\beta}$  which is smooth and almost flat at large scales. The spectrum of  $\Delta_q$  must be such that the small spectral values can be well approximated by  $\{p^2 < C, p \in \mathbb{R}^d\}$  for some constant  $C > 0$ . Its unknown large eigenvalues will then determine the short distance behavior of  $\langle P_g(T) \rangle$ . We can incorporate this modification of the spectrum by introducing a function  $F_q(p^2)$  which tends to 1 for  $p^2 \rightarrow 0$ , and whose large  $p^2$  behavior remains to be determined. Thus

$$\langle P_g(T) \rangle \approx \int \frac{d^d p}{(2\pi)^d} \exp\{-p^2 F_q(p^2) T\}. \quad (2.60)$$

The following argument now suggests that within the asymptotic safety scenario  $F_q(p^2) \sim p^2$  for  $p^2 \rightarrow \infty$ . Before turning to the argument let us note that this property of  $F_q(p^2)$  entails

$$\langle P_g(T) \rangle \sim T^{-d/4} \quad \text{for } T \rightarrow 0, \quad \text{i.e. } d_s = d/2. \quad (2.61)$$

The “microscopic” spectral dimension equals half the “macroscopic”  $d$ . Notably this equals 2, as suggested by the “anomalous dimension argument” precisely in  $d = 4$  dimensions.

The argument for  $F_q(p^2) \sim p^2$  for  $p^2 \rightarrow \infty$  goes as follows: We return to discrete description  $P_g(T) = \sum_n e^{-\mathcal{E}_n(g)T}$  for  $(\mathcal{M}, g)$  compact, and consider the average of one term in the sum  $\langle e^{-\mathcal{E}_n(g)T} \rangle$ , with  $\mathcal{E}_n(g)$  being large. The computation of this average is a single scale problem in the terminology of Appendix A. As such it should allow for a good description via an effective field theory at scale  $k$ . One way of doing this is in terms of the effective average action  $\bar{\Gamma}_k[g]$  as described in Section 4.1. Here only the fact is needed that the average  $\langle e^{-\mathcal{E}_n(g)T} \rangle$  can approximately be evaluated as [135, 134]

$$\langle e^{-\mathcal{E}_n(g)T} \rangle \approx e^{-\mathcal{E}_n(\check{g}_k)T}, \quad \text{where } \frac{\delta \bar{\Gamma}_k}{\delta g_{\alpha\beta}}[\check{g}_k] = 0. \quad (2.62)$$

As indicated  $(\check{g}_k)_{\alpha\beta}$  is a stationary point of the effective action  $\bar{\Gamma}_k[g]$  at a certain scale  $k$ . Since the only scale available is  $\mathcal{E}_n$  itself, the relevant scale  $k$  is for given  $n$  determined by the implicit equation  $k^2 = \mathcal{E}_n(\check{g}_k)$ . Next we consider how these spectral values scale in the fixed point regime where the dimensionless couplings are approximately constant,  $g_i(k) \approx g_i$ . Recall from Equation (2.55) the limiting behavior  $\bar{\Gamma}_k[g] \rightarrow S_*[k^2 g]$  as  $k \rightarrow \infty$ . Two stationary points  $(\check{g}_k)_{\alpha\beta}$  for  $\bar{\Gamma}_k$  and  $(\check{g}_{k_0})_{\alpha\beta}$  for  $\bar{\Gamma}_{k_0}$  will thus in the fixed point regime be simply related by  $k^2 \check{g}_k = k_0^2 \check{g}_{k_0}$ . Since  $k^2 \Delta_{k^2 g} = \Delta_g$  this

means for the spectral values  $k^2 \mathcal{E}_n(\check{g}_k) = k_0^2 \mathcal{E}_n(\check{g}_{k_0})$ . In order to make contact to the continuum parameterization in Equation (2.60) we now identify for given  $p$  the  $n$ 's such that for the typical metric  $q_{\alpha\beta}$  entering Equation (2.60) one has  $\mathcal{E}_n(q) \sim p^2$  for large  $n$ . After this reparameterization  $\mathcal{E}_n = \mathcal{E}_p$ ,  $p = \sqrt{p^2}$  one can identify the  $F_q(p^2)$  in Equation (2.60) with  $F_q(p^2) = \mathcal{E}_p(\check{g}_{k=p})/\mathcal{E}_p(\check{g}_{k_0})$ . This scales for  $p \rightarrow \infty$  like  $p^2$ , which completes the argument.

In summary, the asymptotic safety scenario leads to the specific (theoretical) prediction that the (normally powerlike) short distance singularities of all free matter propagators are softened to logarithmic ones – normally a characteristic feature of massless Klein–Gordon fields in two dimensions. In quantum gravity averages like Equation (2.52) this leads to the expectation that they should scale like  $G(R) \sim \ln R$ , for  $R \rightarrow 0$ . On the other hand this universality allows one to shuffle the effect from matter to gravity propagators. This justifies to attribute the effect to a modification in the underlying random geometry. The average of heat of the heat kernel,  $G(T)$  in Equation (2.53), then scales like  $T^{-d/4}$ . This means the spectral dimension of the random geometries probed by a certain class of “macroscopic” states equals  $d/2$ , which (notably!) equals 2 precisely in  $d = 4$  dimensions.

Accepting this dimensional reduction in the extreme ultraviolet as a working hypothesis one is led to the following question: Is there a two-dimensional field theory which provides an effective description of this regime? “Effective” can mean here “approximate” but quantitatively close, or a system which lies in the same universality class as the original one in the relevant regime. “Effective” is of course not meant to indicate that the theory does not make sense beyond a certain energy scale, as in another use of the term “effective field theory”. We don’t have an answer to the above question but some characteristics of the putative field theory can easily be identified:

1. It should be two-dimensional and self-interacting, the latter because of the non-Gaussian nature of the original fixed point.
2. It should not be a conformal field theory in the usual sense, as the extreme UV regime in the original theory is reached from outside the critical surface (“massive continuum limit”).
3. It should have degrees of freedom which can account for the antiscreening behavior presumed to be responsible for the asymptotically safe stabilization of the UV properties.

Note that in principle the identification of such a UV field theory is a well-posed problem. Presupposing that the functional integral has been made well-defined and through suitable operator insertions data for its extreme UV properties have been obtained, for any proposed field theory with the Properties 1–3 one can test whether or not these data are reproduced.



### 3 Asymptotic Safety from Dimensional Reduction

The systems investigated in this section can be looked at in two ways. First as prototype field theories which have the qualitative Properties 1–3 tentatively identified at the end of the last Section 2.4.4 as characteristics which an effective field theory description of the extreme UV regime of Quantum Gravidynamics should have. Second, they can be viewed as a symmetry reduction of the gravitational functional integral whose embedding into the full theory is left open for the time being. Technically one starts off from the usual gravitational functional integral but restricts it from “4-geometries modulo diffeomorphisms” to “4-geometries constant along a  $2 + 2$  foliation modulo diffeomorphisms”. This means instead of the familiar  $3 + 1$  foliation of geometries one considers a foliation in terms of two-dimensional hypersurfaces  $\Sigma$  and performs the functional integral only over configurations that are constant as one moves along the stack of two-surfaces. The same can be done with the functional integral over matter configurations.

The truncation can be motivated in various ways. It is complementary to the Eikonal sector and describes gravity with collinear initial data in a sense explained later on. It takes into account the crucial ‘spin 2’ aspect, that is, linear and nonlinear gravitational waves are included in this sector and treated without further approximations. Asymptotic safety in this sector is arguably a necessary condition for asymptotic safety of the full theory. Finally, as already mentioned, the sector can serve as a test bed for the investigation of the renormalization structures needed once the extreme UV regime of has been reached.

#### 3.1 $2 + 2$ truncation of Einstein gravity + matter

In accordance with the general picture the renormalization flow will also dictate here to a certain extent the form of the renormalized actions. As mentioned the truncated  $2 + 2$  functional integral turns out to inherit the lack of perturbative renormalizability (with finitely many couplings) from the gravitational part of the full functional integral. However the restricted functional integral is more benign insofar as it is possible to preserve the *conformal geometry in field space* and insofar as *no higher derivative terms* are required for the absorption of cutoff dependencies. The strategy is similar as in the perturbative treatment of the full theory advocated by Gomis and Weinberg [94]: One works with a propagator free of unphysical poles and takes into account all counter terms enforced, but only those. (For the reasons explained in Section 2.3.2 we deliberately avoid using the ‘relevant/irrelevant’ terminology here.) To emphasize the fact that no higher derivative terms are needed we shall refer to the quantum theory defined that way as the symmetry truncation of Quantum Einstein Gravity. We anticipate this fact in the following by taking a classical gravity + matter action as a starting point which is quadratic in the derivatives only (see Equation (3.1) below).

##### 3.1.1 Gravity theories

Since in the renormalization of the restricted functional integral no higher derivative terms will be generated, it suffices to consider 4D gravity actions with second derivatives only. Specifically we consider 4D Einstein gravity coupled to  $k$  Abelian gauge fields and  $\bar{n}$  scalars in a way they typically arise from higher-dimensional (super-)gravity theories. We largely follow the treatment in [45, 46]. The higher-dimensional origin of their 3D reductions is explored in [56]. The 4D action is of the form

$$S_4 = \int d^4x \sqrt{-g} \left[ R(g) - \frac{1}{2} \langle \bar{J}^\alpha, \bar{J}_\alpha \rangle_{\bar{g}} - \frac{q}{4} F_{\alpha\beta}^T (\mu F^{\alpha\beta} - \nu {}^*F^{\alpha\beta}) \right]. \quad (3.1)$$

Here  $g_{\alpha\beta}$ ,  $1 \leq \alpha, \beta \leq 4$ , is the spacetime metric with eigenvalues  $(-, +, +, +)$ ,  $R(g)$  is its scalar curvature, and indices are raised with  $g^{\alpha\beta}$ . There are  $k$  real Abelian vector fields arranged in a

column  $B_\alpha = (B_\alpha^{\hat{i}})$ ,  $\hat{i} = 1, \dots, k$ , with field strength  $F_{\alpha\beta} = \partial_\alpha B_\beta - \partial_\beta B_\alpha$  and dual field strength  $*F^{\alpha\beta} = \frac{1}{2\sqrt{-g}} \epsilon^{\alpha\beta\gamma\delta} F_{\gamma\delta}$ . The scalars  $\bar{\varphi}^i$ ,  $i = 1, \dots, \bar{n}$ , parameterize a non-compact Riemannian symmetric space  $\bar{G}/\bar{H}$  with metric  $\bar{m}_{ij}(\bar{\varphi})$ . Its  $\dim \bar{G}$  Killing vectors give rise to a Lie algebra valued Noether current  $\bar{J}_\alpha$ . In terms of them the sigma-model Lagrangian for the scalars can be written as  $\langle \bar{J}^\alpha, \bar{J}_\alpha \rangle_{\bar{g}}$ , where  $\langle \cdot, \cdot \rangle_{\bar{g}}$  is an invariant scalar product on the Lie algebra  $\bar{g}$ . Finally the coupling matrices  $\mu = \mu(\bar{\varphi})$  and  $\nu = \nu(\bar{\varphi})$  are symmetric  $k \times k$  matrices that depend on the scalars; the constant  $q > 0$  has been extracted for normalization purposes. The vector fields are supposed to contribute positively to the energy density which requires that  $\mu$  is a positive definite matrix. As such it has a unique positive square root  $\mu^{1/2}$  to be used later. The coupling matrices  $\mu$  and  $\nu$  are now chosen in a way that renders the field equations derived from  $S_4$  – though in general not the action itself –  $\bar{G}$ -invariant.

In brief this is done as follows. The field equation for the gauge fields  $\nabla_\alpha(\mu F^{\alpha\beta} - \nu *F^{\alpha\beta}) = 0$  can be interpreted as the Bianchi identity for a field strength  $G_{\alpha\beta} = \partial_\alpha C_\beta - \partial_\beta C_\alpha$  derived from dual potentials  $C_\alpha$ . For later convenience one chooses  $*G_{\alpha\beta} = \eta(\mu F_{\alpha\beta} - \nu *F_{\alpha\beta})$  with some constant orthogonal matrix  $\eta \in O(k)$ . In view of  $**F = -F$  they satisfy the linear relation

$$\begin{pmatrix} F \\ G \end{pmatrix} = \Upsilon \mathcal{V}_c \mathcal{V}_c^T \begin{pmatrix} *F \\ *G \end{pmatrix} \quad \text{with } \mathcal{V}_c = \begin{pmatrix} \mu^{1/2} & \nu \mu^{-1/2} \\ 0 & \eta \mu^{-1/2} \end{pmatrix}, \quad \Upsilon = \begin{pmatrix} 0 & \eta^T \\ -\eta & 0 \end{pmatrix}, \quad (3.2)$$

where the subscript  $c$  is mnemonic for ‘coupling’. If one now assumes that the column  $\begin{pmatrix} F \\ G \end{pmatrix}$  transforms linearly under a faithful  $2k$ -dimensional real matrix representation  $c$  of  $\bar{G}$ , i.e.  $\begin{pmatrix} F \\ G \end{pmatrix} \mapsto c(\bar{g}^{-1})^T \begin{pmatrix} F \\ G \end{pmatrix}$ ,  $\bar{g} \in \bar{G}$ , one finds that Equation (3.2) transforms covariantly if  $\mathcal{V}_c \mapsto c(\bar{g}) \mathcal{V}_c h_c$ , with an orthogonal matrix  $h_c$  and  $c(\bar{g}^{-1})^T = \Upsilon c(\bar{g}) \Upsilon^{-1}$ . Comparing this with the transformation law of the  $\bar{G}$ -valued coset representatives  $\mathcal{V}_*$  (see e.g. [237]) one sees that these conditions are satisfied if  $c(\mathcal{V}_*) = \mathcal{V}_c$  and  $c(\bar{\tau}(\bar{g})) = c(\bar{g}^{-1})^T$ ,  $\bar{g} \in \bar{G}$ , where  $\bar{\tau}$  is the involution whose set of fixed points defines  $\bar{H}$ . Clearly this restricts the allowed cosets  $\bar{G}/\bar{H}$ . For the admissible ones  $c(\mathcal{V}_*) = \mathcal{V}_c$  then determines the couplings  $\mu(\bar{\varphi})$ ,  $\nu(\bar{\varphi})$  as functions of the scalars. Since  $c$  is faithful the determination is unique for a given choice of section  $\mathcal{V}_*$ . Since  $\mathcal{V}_c \mathcal{V}_c^T = c(\mathcal{V}_* \bar{\tau}(\mathcal{V}_*^{-1}))$  the result does not depend on the choice of section, i.e.  $\tilde{\mathcal{V}}_* = \mathcal{V}_* h$  for some  $H$ -valued function  $h$  determines the same  $\mu(\bar{\varphi})$  and  $\nu(\bar{\varphi})$ . We refrain from presenting more details here, which can be found in [45, 46], since the result of the 2-Killing vector reduction can be understood without them.

### 3.1.2 2-Killing vector reduction

Concerning the reduction, we consider here only the case when both Killing vector fields  $K_1, K_2$  are spacelike everywhere and commuting. The other signature (one timelike and one spacelike Killing vector field) is most efficiently treated by relating it to the spacelike case via an Abelian T-duality transformation (see [154]). Alternatively one can perform the reduction in two steps and perform a suitable Hodge dualization in-between (see [45, 56]).

Thus, from here on we take  $K_a = K_a^\alpha \partial_\alpha$ ,  $a = 1, 2$ , to be Killing vector fields on the Lorentzian manifold  $(\mathcal{M}, g)$  that are spacelike everywhere and commuting:  $\mathcal{L}_{K_a} g_{\alpha\beta} = 0$  and  $[K_1, K_2] = 0$ . Their Lorentzian norms and inner product form a symmetric  $2 \times 2$  matrix  $M = (M_{ab})_{1 \leq a, b \leq 2}$ . For the resulting three scalar fields on the 4D spacetime it is convenient to adopt a lapse-shift type parameterization. This gives

$$M_{ab} := g_{\alpha\beta} K_a^\alpha K_b^\beta = K_a \cdot K_b, \quad M =: \frac{\rho}{\Delta} \begin{pmatrix} \Delta^2 + \psi^2 & \psi \\ \psi & 1 \end{pmatrix}. \quad (3.3)$$

In the general relativity literature the fields  $\Delta > 0$  and  $\psi$  are known as the real and imaginary parts of the ‘Ernst potential’. The parameterization (3.3) is chosen such that the (non-negative)

“area element”  $\det M$  is the square of one of the fields. Taking the positive square root one has

$$\rho := \sqrt{K_1 \cdot K_1 K_2 \cdot K_2 - (K_1 \cdot K_2)^2} \geq 0. \quad (3.4)$$

By definition the metric  $g$  is left unchanged along the flow lines of the Killing vector fields. We denote the space of orbits by  $\Sigma$ . A projection operator onto the (co-)tangent space to each point in  $\Sigma$  is given by  $\gamma_\alpha^\beta := \delta_\alpha^\beta - M^{ab} K_{a\alpha} K_b^\beta$ , where  $M^{ab}$  are the components of  $M^{-1}$  and  $K_{a\alpha} := g_{\alpha\beta} K_a^\beta$ . The associated Lorentzian metric  $\gamma_{\alpha\beta} := \gamma_\alpha^{\alpha'} \gamma_{\beta'}^\beta g_{\alpha'\beta'} = g_{\alpha\beta} - M^{ab} K_{a\alpha} K_{b\beta}$  satisfies

$$\mathcal{L}_{K_a} \gamma_{\alpha\beta} = 0, \quad K_a^\alpha \gamma_{\alpha\beta} = 0. \quad (3.5)$$

Since  $\gamma_\alpha^\beta$  is a projector of rank two, the metric  $\gamma_{\alpha\beta}$  has three independent components (not accounting for diffeomorphism redundancies). Generally one can show [90, 91] that there exists a one-to-one correspondence between tensor fields on the “orbit space”  $(\Sigma, \gamma)$  and tensor fields on  $(\mathcal{M}, g)$  with vanishing Lie derivative along  $K_a^\alpha$  and which are “completely orthogonal” to  $K_a^\alpha$ . This will be used for the matter fields below. Given  $\gamma$  subject to Equation (3.5) and  $M$  one can reconstruct the original metric tensor as

$$g_{\alpha\beta} = \gamma_{\alpha\beta} + M^{ab} K_{a\alpha} K_{b\beta}. \quad (3.6)$$

The 10 components are parameterized by the 3 + 3 independent functions in  $\gamma$  and  $M$ . Each of these functions is constant along the flow lines of the Killing vector fields but may vary arbitrarily within  $\Sigma$ .

We deliberately refrained from picking coordinates so far to emphasize the geometric nature of the reduction. As usual however the choice of adapted coordinates is advantageous. We now pick (“Killing”) coordinates in which  $K_a$  acts as  $\partial/\partial y^a$ , for  $a = 1, 2$ . In these coordinates the components of  $\gamma$  and  $M$  are independent of  $y^1$  and  $y^2$  and thus are functions of the remaining (nonunique “non-Killing”) coordinates  $x^0$  and  $x^1$  only. We write  $\gamma_{\mu\nu}(x)$ ,  $\mu = 0, 1$ , for the components of  $\gamma$  in such a coordinate system. Since both Killing vectors are spacelike,  $\gamma_{\mu\nu}$  has eigenvalues  $(-, +)$  and can be brought into the form  $\gamma_{\mu\nu}(x) = e^{\sigma(x)} \eta_{\mu\nu}$ , by a change of the non-Killing coordinates, where  $\eta$  is the metric of flat 1 + 1-dimensional Minkowski space. This can be taken to define  $\sigma$ . Alternatively one can introduce  $\sigma$  by

$$\partial\rho \cdot \partial\rho = \Delta(x) \gamma^{\mu\nu}(x) \partial_\mu \rho \partial_\nu \rho = \Delta(x) e^{-\sigma(x)} \partial^\mu \rho \partial_\mu \rho. \quad (3.7)$$

On the left-hand-side is the (coordinate-independent) Lorentzian norm of the gradient of the 4D scalar field  $\rho$ ; on the right-hand-side we evaluated this norm in the Killing coordinates where it must be proportional to  $\gamma^{\mu\nu}(x) \partial_\mu \rho \partial_\nu \rho$ . Adjusting also the non-Killing coordinates then gives the rightmost expression in Equation (3.7), which could also be taken to define  $\sigma$ . The upshot is that the most general 4D metric with two commuting Killing vectors is parameterized by four scalar fields,  $\rho, \sigma$  and  $\Delta, \psi$ . In the adapted coordinates the 4D metric then reads

$$dS^2 = e^\sigma [-(dx^0)^2 + (dx^1)^2] + \frac{\rho}{\Delta} (dy^1 + \psi dy^2)^2 + \rho \Delta (dy^2)^2. \quad (3.8)$$

As already mentioned, the fields  $\Delta$  and  $\psi$  are known as the real and the imaginary part of the “Ernst potential”; we shall refer to  $\rho$  as the “area radius” associated with the two Killing vectors, and to  $\sigma$  as the conformal factor. To motivate the latter note that a Weyl transformation  $g_{\alpha\beta} \rightarrow e^\omega g_{\alpha\beta}$  of the 4D metric compatible with the 2-Killing vectors amounts to the simultaneous rescalings  $\gamma_{\mu\nu}(x) \rightarrow e^{\omega(x)} \gamma_{\mu\nu}(x)$  and  $\rho(x) \rightarrow e^{\omega(x)} \rho(x)$ .

The matter content in Equation (3.1) consists of the  $k$  Abelian gauge fields and the sigma-model scalars  $\bar{\varphi}^i$ ,  $i = 1, \dots, \bar{n}$ . For the scalars the reduction is trivial, and simply amounts to considering configurations constant in the Killing coordinates. For the gauge fields it turns out that the  $4k$

components of  $B_\alpha^{\hat{i}}$ ,  $\hat{i} = 1, \dots, k$ , give rise to  $2k$  fields  $A^I$ ,  $I = 1, \dots, 2k$ , which transform as scalars under a change of the non-Killing coordinates  $(x^0, x^1)$ . In brief this comes about as follows. The field equation  $\nabla_\alpha(\mu F^{\alpha\beta} - \nu^* F^{\alpha\beta}) = 0$  for the gauge fields in Equation (3.1) can be interpreted as the Bianchi identity for a field strength  $G_{\alpha\beta} = \partial_\alpha C_\beta - \partial_\beta C_\alpha$  derived from dual potentials  $C_\alpha^{\hat{i}}$ ,  $\hat{i} = 1, \dots, k$ . We can take one of the Killing vectors, say  $K = K_1$ , and build  $2k$  spacetime scalars by contraction  $B^{\hat{i}} := B_\alpha^{\hat{i}} K^\alpha$  and  $C^{\hat{i}} := C_\alpha^{\hat{i}} K^\alpha$ . Reduction with respect to the other Killing vector  $K_2$  just requires that these scalars are constant in the corresponding Killing coordinate  $y^2$ . The dependence on  $y^1$  is constrained by gauge invariance. If  $B_\alpha^{\hat{i}} \mapsto B_\alpha^{\hat{i}} + \partial_\alpha b^{\hat{i}}$  and  $C_\alpha^{\hat{i}} \mapsto C_\alpha^{\hat{i}} + \partial_\alpha c^{\hat{i}}$ , the scalars change by a term  $K^\alpha \partial_\alpha b^{\hat{i}}$  and  $K^\alpha \partial_\alpha c^{\hat{i}}$ , respectively, and hence are invariant under  $y^1$  independent gauge transformations. Thus, if the 4D gauge potentials and their duals, together with the corresponding transformations are taken to be independent of  $y^1, y^2$ , a set of gauge invariant scalars  $B^{\hat{i}}(x)$  and  $C^{\hat{i}}(x)$  arises. As a matter of fact a constant remnant of the gauge transformations remains and gives rise to a symmetry of the reduced system (see the discussion after Equation (3.11) below). We arrange the  $2k$  fields  $B^{\hat{i}}, C^{\hat{i}}$  in a column vector  $A^I$ ,  $I = 1, \dots, 2k$ . For convenience we summarize the field content of the 2-Killing vector subsector of Equation (3.1) in Table 1.

4D fields	fields in 2-Killing subsector
$g_{\alpha\beta}$ metric	$\Delta, \psi, \rho, \sigma$
$B_\alpha^{\hat{i}}$ Abelian gauge fields	$A^I, I = 1, \dots, 2k$
$\bar{\varphi}^{\hat{i}}$ KK scalars	$\bar{\varphi}^{\hat{i}}, \hat{i} = 1, \dots, \bar{n}$

Table 1: Field content of the 2-Killing vector subsector of the gravity theories 3.1.

We combine all but  $\rho$  and  $\sigma$  into an  $n := 2 + \bar{n} + 2k$  components scalar field

$$\varphi = (\Delta, \psi, \bar{\varphi}^1, \dots, \bar{\varphi}^{\bar{n}}, A^1, \dots, A^{2k}) \quad (3.9)$$

on the 2D orbit space with metric  $\gamma_{\mu\nu}$  and coordinates  $(x^0, x^1)$ . A lengthy computation (which is best done in a two step procedure; see [45, 56]) gives the form of the action (3.1) on the field configurations compatible with the two Killing vectors. The result is

$$S = \frac{1}{2\lambda} \int d^2x \rho \sqrt{\gamma} [2R(\gamma) + \gamma^{\mu\nu} \rho^{-2} \partial_{\mu\rho} \partial_{\nu\rho} - \gamma^{\mu\nu} \mathbf{m}_{ij}(\varphi) \partial_\mu \varphi^i \partial_\nu \varphi^j]. \quad (3.10)$$

Here  $\lambda$  is the gravitational constant per unit volume of the internal space. One sees that the reduced action has the form of a 2D nonlinear sigma-model non-minimally coupled to 2D gravity via the area radius  $\rho$  of the two Killing vectors. The target space of the sigma-model has dimension  $n = 2 + \bar{n} + 2k$ , we take Equation (3.9), viewed as a column, as field coordinates. With the normalization  $\langle \bar{J}_{\hat{\alpha}}, \bar{J}_{\hat{\beta}} \rangle_{\bar{g}} = \bar{\mathbf{m}}_{ij}(\bar{\varphi}) \partial_{\hat{\alpha}} \bar{\varphi}^i \partial_{\hat{\beta}} \bar{\varphi}^j$  the metric then comes out as

$$\mathbf{m}(\varphi) = \left( \begin{array}{cc|c|c} \frac{1}{\Delta^2} & 0 & & 0 \\ 0 & \frac{1}{\Delta^2} & & -\frac{q}{2\Delta^2} A^T \Upsilon \\ \hline & & \bar{\mathbf{m}}(\bar{\varphi}) & \\ \hline 0 & \frac{q}{2\Delta^2} \Upsilon A & & \frac{q}{\Delta} \mathcal{V}_c \mathcal{V}_c^T - \frac{q^2}{4\Delta^2} \Upsilon A \otimes A^T \Upsilon \end{array} \right). \quad (3.11)$$

Here  $\Upsilon$  and  $\mathcal{V}_c$  are as in Equation (3.2) and  $q > 0$  is a parameter used to adjust normalizations. The metric (3.11) has Riemannian signature (if the reduction was performed with respect to one spacelike and one timelike Killing vector it had  $2k$  negative eigenvalues).

We briefly digress on the isometries of Equation (3.11). By virtue of the  $\overline{G}$  invariance of the action  $\mathfrak{m}$  has  $\dim \overline{G}$  Killing vectors of which  $\bar{n} = \dim \overline{G}/\overline{H}$  are algebraically independent. Interestingly, the action (3.10) is also invariant under  $A \mapsto A + a$ ,  $\psi \mapsto \psi - \frac{q}{2} A^T \Upsilon a$ , with a constant  $2k$  column  $a$ . These symmetries can be viewed as residual gauge transformations; note however that a compensating transformation of the gravitational potential  $\psi$  is needed. Finally constant translations in  $\psi$  and scale transformations  $(\Delta, \psi, \overline{\varphi}, A) \mapsto (s\Delta, s\psi, \overline{\varphi}, s^{1/2}A)$ ,  $s > 0$ , are obvious symmetries of the action. The associated Killing vectors  $\mathbf{e}, \mathbf{h}$  of  $\mathfrak{m}$  generate a Borel subalgebra of  $sl_2$ , i.e.  $[\mathbf{h}, \mathbf{e}] = -2\mathbf{e}$ . Together the metric (3.11) always has  $\dim \overline{G} + 2k + 2$  Killing vectors.

In contrast the last  $sl_2$  generator  $\mathbf{f}$  is only a Killing vector of  $\mathfrak{m}$  under certain conditions on  $\overline{G}/\overline{H}$ . If these are satisfied a remarkable ‘symmetry enhancement’ takes place in that  $\mathfrak{m}$  is the metric of a much larger symmetric space  $G/H$ , where  $G$  is a non-compact real form of a simple Lie group with  $\dim G = \dim \overline{G} + 4k + \dim \text{SL}(2)$ . The point is that if  $\mathbf{f}$  exists as Killing vector its commutator with the gauge transformations is nontrivial and yields  $2k$  additional symmetries (generalized ‘Harrison transformations’). Since  $\mathfrak{m}$  always has  $\dim \overline{G} + 2k + 2$  Killing vectors, the additional  $1 + 2k$  then match the dimension of  $G$ . For the number of dependent Killing vectors, i.e. the dimension of the putative maximal subgroup  $H \subset G$  one expects  $\dim H = \dim \overline{H} + 1 + 2k$ . Indeed under the conditions stated the symmetric space  $G/H$  exists and is uniquely determined by  $\overline{G}/\overline{H}$  (and the signature of the Killing vectors). See [45] for a complete list. Evidently the gauge fields are crucial for the symmetry enhancement. Among the systems in [45] only pure gravity has  $k = 0$ .

From now on we restrict attention to the cases where such a symmetry enhancement takes place. The scalars  $\varphi^i$  can then be arranged into a coset nonlinear sigma-model whose  $2k + \bar{n} + 2$ -dimensional target space is of the form  $G/H$ . Here  $G$  is always a simple noncompact Lie group and  $H$  a maximal subgroup; the coset is a Riemannian space with metric  $\mathfrak{m}_{ij}(\varphi)$ . Being the metric of a symmetric space  $\mathfrak{m}$  enjoys the properties

$$\nabla_m R_{ijkl}(\mathfrak{m}) = 0, \quad R_{ij}(\mathfrak{m}) = \zeta_1 \mathfrak{m}_{ij}, \quad (3.12)$$

which will be important later on. By construction  $G/H$  contains as subcosets the space  $\text{SL}(2, \mathbb{R})/\text{SO}(2)$  on which the gravitational potentials are coordinates and the original coset  $\overline{G}/\overline{H}$  of the scalars  $\overline{\varphi}$ :

$$\text{SL}(2, \mathbb{R})/\text{SO}(2) \subset G/H \supset \overline{G}/\overline{H}. \quad (3.13)$$

A brief list of examples of 4D theories (3.1) and the cosets  $G/H$  they give rise to in Equation (3.10) is: Pure gravity in 4D corresponds to  $\text{SL}(2, \mathbb{R})/\text{SO}(2)$ , Einstein–Maxwell theory gives rise to  $\text{SU}(2, 1)/S[\text{U}(2) \times \text{U}(1)]$ , a 4D Einstein–dilaton–axion theory gives a coset  $\text{SO}(3, 2)/\text{SO}(3) \times \text{SO}(2)$ , and the reduction of  $N = 8$  supergravity leads to a bosonic sector with  $E_{8(+8)}/\text{SO}(16)$  coset.

The reduced classical field theories (3.10) have some remarkable properties which we discuss now:

1. First, the field equations and the symplectic structure derived from Equation (3.11) coincide with the restriction of the field equations and the symplectic structure derived from Equation (3.1). In fact, this is a general feature of this type of symmetry reductions, and can be understood in terms of the ‘principle of symmetric criticality’ [76].
2. The field equations are classical integrable in the sense that they can be written as the compatibility condition of a pair of first order matrix-valued differential operators, depending on a free complex parameter. As a consequence large classes of solutions can be constructed analytically (see the books [26, 121, 98] and [103, 4] for detailed expositions).
3. The classical integrability also entails that an infinite number of nonlocal conserved charges can be constructed explicitly. These are generalizations of the Lüscher–Pohlmayer charges

for the  $O(N)$  model. Moreover, these charges Poisson commute with the Hamiltonian and the diffeomorphism constraint that arise in a Hamiltonian (“Arnowitt–Deser–Misner-type”) analysis of the covariant system (3.11). In other words, an infinite system of Dirac observables can be constructed explicitly as functionals of the metric and the matter fields! Given the fact that (apart from mass and angular momentum) not a single Dirac observable is known explicitly in full general relativity, this is a most remarkable feature.

4. The system captures the crucial “spin-two” aspect of gravity. For example without matter the classical solutions comprise various types of (nonlinear) gravitational waves with two independent polarizations (per spacetime point).
5. In conformal gauge,  $\gamma_{\mu\nu}(x) = e^{\sigma(x)}\eta_{\mu\nu}$ , the curvature term in Equation (3.11) is proportional to  $\partial^\mu\rho\partial_\mu\sigma$ . Upon diagonalization,  $\partial^\mu(\rho+\sigma)\partial_\mu(\rho+\sigma) - \partial^\mu(\rho-\sigma)\partial_\mu(\rho-\sigma)$ , this is proportional to a sum of two standard kinetic terms, one of which invariably has the ‘wrong sign’. Since a Weyl rescaling  $g_{\alpha\beta} \mapsto e^{\omega(x)}g_{\alpha\beta}$  of the metrics (3.3) amounts to  $\rho \mapsto e^\omega\rho$ ,  $\sigma \mapsto \sigma + \omega$  this appears to reflect a conformal factor instability of the 4D gravitational action(s). Upon closer inspection it signals the absence of a genuine instability (see Section 3.2).

Taken together these properties make the 2-Killing vector reductions a compelling laboratory to study the quantum aspect of the gravitational field.

### 3.1.3 Hamiltonian formulation

For later reference we also briefly outline the Hamiltonian formulation of the system. In a Hamiltonian formulation of this two-dimensional diffeomorphism invariant system one fixes  $\gamma_{\mu\nu} = e^\sigma\eta_{\mu\nu}$  at the expense of a Hamiltonian constraint  $\mathcal{H}_0$  and a diffeomorphism constraint  $\mathcal{H}_1$ . The properly normalized constraints come out of a lapse and shift decomposition of the form

$$\gamma_{\mu\nu} = e^\sigma \begin{pmatrix} -n^2 + s^2 & s \\ s & 1 \end{pmatrix}. \quad (3.14)$$

The lapse  $n$  and the shift  $s$  here are spatial densities of weight  $-1$ , while  $e^\sigma$  carries weight 2. In Section 3.1.4 we collect some useful formulas which allows one to streamline the Hamiltonian analysis of 2D gravity theories.

Performing a standard Hamiltonian (“ADM type”) analysis based on Equation (3.10) and Equation (3.14), using the formulas of Section 3.2.3 one finds

$$\begin{aligned} \mathcal{H}_0 = T_{00} &= \frac{1}{2}(T_{++} + T_{--}) \\ &= \frac{\lambda}{2\rho} \mathbf{m}^{ij}(\varphi)\pi_i(\varphi)\pi_j(\varphi) + \frac{\rho}{2\lambda} \mathbf{m}_{ij}(\varphi)\partial_1\varphi^i\partial_1\varphi^j + \lambda\pi_\rho\pi_\sigma + \frac{1}{\lambda}\partial_1\sigma\partial_1\rho - \frac{2}{\lambda}\partial_1^2\rho, \\ \mathcal{H}_1 = T_{10} &= \frac{1}{2}(T_{++} - T_{--}) \\ &= \pi_i(\varphi)\partial_1\varphi^i + \pi_\sigma\partial_1\sigma + \pi_\rho\partial_1\rho - 2\partial_1\pi_\sigma, \end{aligned} \quad (3.15)$$

with a self-explanatory notation for the canonical momenta. The linear combinations  $T_{\pm\pm}$  are introduced for later use. The action (3.10) can then be recast in Hamiltonian form, where the Hamiltonian is, up to possible boundary terms, given by  $\int dx (n\mathcal{H}_0 + s\mathcal{H}_1)$ . If boundary terms are present they are separately conserved local charges (see [197]). Note also that a constant  $\rho$  field trivializes the system in that  $\rho = \text{const}$  implies  $\pi_\sigma = 0$ , in which case the only solutions of the constraints are  $\varphi^i = \text{const}$ ,  $\pi_i(\varphi) = 0$ . Then  $\sigma$  is harmonic function and  $e^\sigma$  can be set to unity by a conformal transformation. In other words, for  $\rho = \text{const}$  the only solutions of the classical field equations is Minkowski space with constant matter fields.



A shortcut to arrive at the constraints (3.15) is to start from the Lagrangian (3.10) in conformal gauge ( $\gamma_{\mu\nu} = e^\sigma \eta_{\mu\nu}$ ,  $\eta = \text{diag}(-1, 1)$ ), using  $R(e^\sigma \eta) = -e^{-\sigma} \partial^2 \sigma$ ,

$$L(\varphi, \rho, \sigma) = -\frac{1}{2\lambda} [\rho \mathbf{m}_{ij}(\varphi) \partial^\mu \varphi^i \partial_\mu \varphi^j - 2\partial^\mu \rho \partial_\mu \sigma - \rho^{-1} \partial^\mu \rho \partial_\mu \rho], \quad (3.16)$$

and to work out the ‘would be’ energy momentum tensor  $T_{\mu\nu}$  of the Poincaré invariant Lagrangian (3.16). The relation between the velocities and the momenta then is of course different from that based on the lapse and shift analysis. However the Hamiltonian constraint  $\mathcal{H}_0$  and the 1D diffeomorphism constraint  $\mathcal{H}_1$ , regarded as functions of the momenta, coincide with  $T_{00}$  and  $T_{01}$ , respectively, as derived from Equation (3.16) as the components of the energy momentum tensor. Note that the trace of the ‘would-be’ energy momentum tensor  $\lambda T^\mu{}_\mu = -2\partial^\mu \partial_\mu \rho$  vanishes if the equation of motion  $\partial^\mu \partial_\mu \rho = 0$  of  $\rho$  is imposed. We shall freely switch back and forth between both interpretations of the constraints.

For the computation of the Poisson algebra it is convenient to put  $\rho$  on-shell throughout (as its equation of motion  $\partial^\mu \partial_\mu \rho = 0$  is trivially solved) and to interpret the improvement terms in Equation (3.15) (here, those linear in the canonical variables) such that second time derivatives are eliminated. As expected, the  $\mathcal{H}_1$  generates infinitesimal spatial reparameterizations and the covariance of the fields is a merely kinematical property. Explicitly a spatial density  $d(x)$  of weight  $s$  transforms as  $\{\mathcal{H}_1(x), d(y)\} = \partial_1 d \delta(x - y) - s d(y) \delta'(x - y)$ , the right-hand-side being the infinitesimal version of  $d(x) \rightarrow \bar{d}(\tilde{x}) = [f'(\tilde{x})]^{-s} d(f(\tilde{x}))$ , under  $x \rightarrow \tilde{x} = f^{-1}(x)$ . The canonical momenta  $\pi_\rho, \pi_\sigma, \pi_i$  are spatial densities of weight  $s = 1$ , while  $\mathcal{H}_0, \mathcal{H}_1$ , and  $e^\sigma$  are densities of weight 2. The Hamiltonian constraint on the other hand resumes its usual kinematical-dynamical double role.

The advantage of having the constraints defined with respect to the densitized lapse and shift functions (3.14) is that the Poisson algebra generated by  $\mathcal{H}_0$  and  $\mathcal{H}_1$  is a Lie algebra on-shell, and equivalent to the algebra of 2D conformal transformations. (Otherwise it yields the algebra of ‘surface deformations’; cf. [214]). To illustrate the difference let us note that with only the  $\rho$  equation of motion imposed one computes

$$\begin{aligned} \{T_{\pm\pm}(x), T_{\pm\pm}(y)\} &= \mp 2 [T_{\pm\pm}(x) + T_{\pm\pm}(y)] \delta'(x - y), \\ \{T_{++}(x), T_{--}(y)\} &= 0. \end{aligned} \quad (3.17)$$

If also  $\partial_\mp T_{\pm\pm} = 0$  is imposed the first equation can be rewritten as

$$\{T_{\pm\pm}(x), T_{\pm\pm}(y)\} = \partial_\pm T_{\pm\pm}(y) \delta(x - y) \mp 4T_{\pm\pm}(y) \delta'(x - y). \quad (3.18)$$

As expected, the  $T_{\pm\pm}$  generate infinitesimal conformal transformations on the basis fields  $\rho, \sigma, \varphi^j$ :  $\{T_{\pm\pm}(x), d(y)\} = \partial_\pm d(y) \delta(x - y)$ .

### 3.1.4 Lapse and shift in 2D gravity theories

Here we collect some useful formulas for 2D gravity theories in a lapse/shift parameterization of the metric, taken from [156]. As a byproduct we obtain a closed expression for the current  $K_\mu$  of the Euler density  $\sqrt{\gamma} R^{(2)}(\gamma) = -\partial_\mu K_\mu$  expressed in terms of the metric only. See [62] for a discussion. Our curvature conventions are the ones used throughout, the metric  $\gamma_{\mu\nu}$  has eigenvalues  $(-, +)$ .

In 2D a ‘densitized’ lapse-shift parameterization is convenient (see e.g. [197]),

$$\gamma_{\mu\nu} = e^\sigma \begin{pmatrix} -n^2 + s^2 & s \\ s & 1 \end{pmatrix}. \quad (3.19)$$

The lapse  $n$  and the shift  $s$  here are spatial densities of weight  $-1$ , while  $e^\sigma$  carries weight 2. The rationale for this densitized lapse and shift parameterization is that the associated constraints

automatically generate the proper Lie algebra of surface deformations (see [214, 197]). Using

$$\sqrt{-\gamma}\gamma^{\mu\nu} = \begin{pmatrix} -\frac{1}{n} & \frac{s}{n} \\ \frac{s}{n} & n - \frac{s^2}{n} \end{pmatrix}, \quad (3.20)$$

one checks that the Hamiltonian associated with a free field Lagrangian  $-\frac{1}{2}\sqrt{-\gamma}\gamma^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$  is  $\frac{n}{2}[\pi_\phi^2 + (\partial_1\phi)^2] + s\partial_1\phi\pi_\phi$ , as it should be. Evaluating  $\sqrt{-\gamma}R^{(2)}(\gamma)$  in this parameterization gives the following expression (which we were unable to locate in the literature):

$$\sqrt{-\gamma}R^{(2)}(\gamma) = \partial_0 \left[ \frac{1}{n}(\partial_0\sigma - s\partial_1\sigma - 2\partial_1s) \right] + \partial_1 \left[ \frac{1}{n}((s^2 - n^2)\partial_1\sigma - s\partial_0\sigma + \partial_1(s^2 - n^2)) \right]. \quad (3.21)$$

It can be rewritten as

$$\begin{aligned} \sqrt{-\gamma}R^{(2)}(\gamma) &= -\partial_\mu K^\mu, \\ K^\mu &:= \sqrt{-\gamma}\gamma^{\mu\nu}\partial_\nu \ln \sqrt{-\gamma} + \partial_\nu [\sqrt{-\gamma}\gamma^{\mu\nu}] + \epsilon^{\mu\nu}\sqrt{-\gamma}\gamma^{01}\partial_\nu \ln(-\sqrt{-\gamma}\gamma^{00}), \end{aligned} \quad (3.22)$$

where  $\epsilon_{01} = -\epsilon_{10} = 1 = -\epsilon^{01} = \epsilon^{10}$ .

This provides an explicit though noncovariant expression for the current  $K^\mu$  in terms of the metric. Related formulas either invoke the zweibein or use an explicit parameterization. The one given in [62] is based on an  $SL(2, \mathbb{R})$  type parameterization of  $\gamma_{\mu\nu}/\sqrt{-\gamma}$  and is equivalent to Equation (3.22). Compared to Equation (2.11) in [62] a curl term  $\epsilon^{\mu\nu}\partial_\nu\phi$  has been added which allows one to express  $K^\mu$  solely in terms of the metric. Another advantage of Equation (3.22) is that the separation in dynamical and nondynamical variables is manifest:  $K^\mu$  is a function of  $\det \gamma$  and the combination  $\sqrt{-\gamma}\gamma^{\mu\nu}$  only; the former is the dynamical variable, the latter can be parameterized in terms of the lapse and shift functions. They can be anticipated to be nondynamical in that no time derivatives of lapse and shift appear in  $K^\mu$ , as is manifest from Equation (3.21).

In the actions considered the term  $\sqrt{-\gamma}R^{(2)}(\gamma)$  always multiplies a scalar field  $\rho$ , say. Using Equation (3.20) the Hamiltonian associated with an Lagrangian of the form

$$L = \rho\sqrt{-\gamma} \left[ R^{(2)}(\gamma) - \frac{1}{2}\gamma^{\mu\nu}\partial_\mu\phi\partial_\nu\phi \right] \quad (3.23)$$

is readily worked out. As a function of the momenta  $\pi_\rho = -\frac{1}{n}(\partial_0\sigma - s\partial_1\sigma - 2\partial_1s)$  and  $\pi_\sigma = \frac{1}{n}(-\partial_0\rho + s\partial_1\rho)$ , one has  $\rho\sqrt{-\gamma}R^{(2)}(\gamma) \simeq n[-\pi_\rho\pi_\sigma + \partial_1\rho\partial_1\sigma - 2\partial_1^2\rho]$ , where ‘ $\simeq$ ’ denotes ‘modulo total derivatives’. For the Hamiltonian this gives

$$\begin{aligned} H &= n\mathcal{H}_0 + s\mathcal{H}_1, \\ \mathcal{H}_0 &= \frac{1}{2\rho}\pi_\phi^2 + \frac{\rho}{2}(\partial_1\phi)^2 - (\pi_\rho\pi_\sigma + \partial_1\rho\partial_1\sigma - 2\partial_1\rho), \\ \mathcal{H}_1 &= \pi_\phi\partial_1\phi + \pi_\rho\partial_1\rho + \pi_\sigma\partial_1\sigma - 2\partial_1\pi_\sigma. \end{aligned} \quad (3.24)$$

The Poisson algebra of the constraints  $\mathcal{H}_0, \mathcal{H}_1$  is the algebra of surface deformations, as required. Alternatively the form (3.22) can be used beforehand to get

$$\rho\sqrt{-\gamma}R^{(2)}(\gamma) \simeq \sqrt{-\gamma}\gamma^{\mu\nu}\partial_\mu\rho\partial_\nu \ln \sqrt{-\gamma} + \partial_\mu\rho\partial_\nu [\sqrt{-\gamma}\gamma^{\mu\nu}] + \sqrt{-\gamma}\gamma^{01}\epsilon^{\mu\nu}\partial_\mu\rho\partial_\nu \ln [-\sqrt{-\gamma}\gamma^{00}], \quad (3.25)$$

where  $\simeq$  denotes again ‘modulo total derivatives’.

### 3.1.5 Symmetries and currents

Later on we aim at a Dirac quantization of the 2-Killing subsector of the theories (3.1). The functional measure is then defined with the reduced Lagrangian (3.16) in conformal gauge, and a quantum version of the constraints  $\mathcal{H}_0 \pm \mathcal{H}_1$  is imposed subsequently. For the renormalization the symmetries of the Lagrangian (3.16) are crucial. The  $G$  invariance of course gives rise to a set of Lie algebra valued Noether currents  $J_\mu$ . Let  $t_a$ ,  $a = 1, \dots, \dim G$ , denote a basis of the Lie algebra with Killing form  $\langle \cdot, \cdot \rangle_g$ . Let further  $Y_a = Y_a^i(\varphi)\partial/\partial\varphi^i$  denote the Killing vectors of  $\mathfrak{m}$ . Then we define  $J_\mu$  through its projection onto the basis  $t_a$ , via

$$\langle t_a, J_\mu \rangle_g := 2Y_a^i(\varphi)m_{ij}(\varphi)\partial_\mu\varphi^j. \quad (3.26)$$

From the equation of motion for  $\varphi$  and the Killing vector equation one readily checks that  $\partial^\mu(\rho J_\mu) = 0$ . Further, the invariance of Equation (3.16) under constant shifts in  $\sigma$  gives the trivial current  $\partial_\mu\rho$ .

More interestingly there are two ‘conformal currents’ which are not conserved on-shell but whose divergence reproduces the Lagrangian (3.16) up to a multiple

$$\begin{aligned} \partial^\mu C_\mu &= L, & \lambda C_\mu &= \rho\partial_\mu\left(\sigma + \frac{1}{2}\ln\rho\right), \\ \partial^\mu D_\mu &= -\ln\rho L, & \lambda D_\mu &= \rho(\sigma\partial_\mu\ln\rho - \ln\rho\partial_\mu\sigma). \end{aligned} \quad (3.27)$$

In fact,  $C_\mu$  and  $D_\mu$  are the only currents with that property. Their origin are the following scaling relations,

$$\begin{aligned} L &\mapsto \omega^2 L & \text{if } (\rho, \sigma) &\mapsto (\omega^2\rho, \sigma + c\ln\omega^2), \quad c \in \mathbb{R}, \\ L &\mapsto \rho^{\omega-1} L & \text{if } (\ln\rho, \sigma + \frac{1}{2}\ln\rho) &\mapsto \left[\omega\ln\rho, \omega^{-1}\left(\sigma + \frac{1}{2}\ln\rho\right)\right], \end{aligned} \quad (3.28)$$

where  $L$  is the Lagrangian (3.16). The parameter  $c$  is arbitrary,  $c = -1/2$  matches the algebra of the conformal Killing vectors in Section 3.3, and  $c = 1$  corresponds to scale  $g_{\alpha\beta} \mapsto \omega^2 g_{\alpha\beta}$  transformations of the original 4D metric (3.6). Since for the action (3.10) the response under such a rescaling vanishes on-shell (see Equation (1.1)), the inessential nature of Newton’s constant is preserved by the reduction and remains visible through the conformal current  $C_\mu$  in the gauge fixed Lagrangian (3.16).

Finally there exists an infinite set of nonlocal conserved currents whose charges are Dirac observables and which be constructed explicitly(!) in terms of the dynamical fields, that is, without having to solve the field equations. These currents can be found by different techniques similar to those used in nonlinear sigma-models [150, 124]. For illustration we present the lowest one which is (for all on-shell configurations) defined in terms of the dual potentials  $\partial_\mu\chi := -\epsilon_{\mu\nu}\rho J^\nu$  and  $\partial_\mu\tilde{\rho} := -\epsilon_{\mu\nu}\rho^\nu$ , with  $\epsilon_{01} = -\epsilon_{10} = 1$ . Then

$$\begin{aligned} \left\{ \int dx \mathcal{J}_0, \mathcal{H}_\mu \right\} &= 0, & \mu &= 0, 1, \\ \mathcal{J}_\mu &= \frac{1}{2}[\rho J_\mu, \chi] + 2\tilde{\rho}\rho J_\mu - \rho^2\epsilon_{\mu\nu}J^\nu. \end{aligned} \quad (3.29)$$

It may be worthwhile to point out what is trivial and what is nontrivial about the relations (3.29). Once the expression for the current  $\mathcal{J}_\mu$  is known it is trivial to verify its conservation using the definition of the potentials  $\chi$  and  $\tilde{\rho}$ . Since  $\mathcal{H}_0$  generates time translations on the basis fields  $\rho, \sigma, \varphi^i$  the associated conserved charge Poisson commutes with  $\mathcal{H}_0$  (and trivially with  $\mathcal{H}_1$ ) and thus qualifies as a genuine Dirac observable. What is nontrivial about Equation (3.29) is that

a Dirac observable can be constructed explicitly in a way that does *not require a solution of the Cauchy problem*. The potentials  $\chi$ ,  $\tilde{\rho}$  are only defined on-shell but one does not need to know how they are parameterized by initial data. In stark contrast the known abstract construction principles for Dirac observables in full general relativity always refer to a solution of the Cauchy problem (see [70] for a recent account). The bonus feature of the 2-Killing vector reduction that allows for this feat is the existence of a solution generating group [90, 91] (“Geroch group”) and, related to it, the existence of a Lax pair. The latter allows one to convert the Cauchy problem into a linear singular integral equation [4, 103] (which is still nontrivial to solve) and at the same time it underlies the techniques used to find an infinite set of nonlocal conserved currents of which the one in Equation (3.29) is the lowest (least nonlinear) one.

In the quantum theory, a construction of observables from first principles has not yet been achieved. Existence of a quantum counterpart of the first charge (3.29) would already be a very nontrivial indication for the quantum integrability of the systems. For its construction the procedure of Lüscher [139] could be adopted. Independent of this, the ‘exact’ bootstrap formulation of [158] shows that the existence of a ‘complete’ set of quantum observables is compatible with the quantum integrability of the system.

## 3.2 Collinear gravitons, Dirac quantization, and conformal factor

In this section, taken from [156], we discuss a number of structural issues of the 2 + 2 truncations and advocate that, as far as the investigation of the renormalization properties is concerned, the use of a proper time or Dirac quantization is the method of choice. We begin by describing what conventional graviton perturbation theory looks like in this sector.

### 3.2.1 Collinear gravitons

Performing a standard expansion  $g_{\alpha\beta} = \eta_{\alpha\beta} + f_{\alpha\beta}$ ,  $0 \leq \alpha, \beta \leq 3$ , around the metric  $\eta_{\alpha\beta}$  of Minkowski space the quadratic part of the Einstein–Hilbert action  $\int d^4x \sqrt{-g}R(g)$  reads

$$\begin{aligned} S''_{\text{EH}}[f] &= -\frac{1}{4} \int d^4x \left[ \partial_\alpha f_{\beta\gamma} \partial^\alpha f^{\beta\gamma} - 2\partial_\alpha f_{\beta\gamma} \partial^\gamma f^{\beta\alpha} + 2\partial_\alpha f^\beta{}_\beta \partial_\gamma f^{\gamma\alpha} - \partial_\alpha f^\beta{}_\beta \partial^\alpha f^\gamma{}_\gamma \right] \\ &= -\frac{1}{4} \int d^4x \left[ f^{\text{TF}\alpha\beta} (2\partial^\gamma \partial_\alpha f_{\gamma\beta}^{\text{TF}} - \partial^2 f_{\alpha\beta}^{\text{TF}} - \partial_\alpha \partial_\beta f) + \frac{3}{8} f \partial^2 f \right]. \end{aligned} \quad (3.30)$$

Indices are raised with  $\eta^{\alpha\beta}$ . In the second part of Equation (3.30) we decomposed  $f_{\alpha\beta}$  into its trace  $f := f^\alpha{}_\alpha$  and a trace free part  $f_{\alpha\beta}^{\text{TF}} := f_{\alpha\beta} - \frac{1}{4}\eta_{\alpha\beta}f$ . A critical discussion of the reconstruction of the nonlinear theory from the linear one can be found in [164]; a proper time formulation of the linearized quantum theory has been given in [141]. In the present conventions a scalar action of the form  $S = -\frac{1}{2} \int d^4x [\partial^\alpha \phi \partial_\alpha \phi + m^2 \phi^2]$  has a positive semidefinite Hamiltonian. The kinetic term for the trace  $f$  thus has the ‘wrong’ sign, which reflects the conformal factor instability. One readily checks the invariance of the action (3.30) under the gauge transformations

$$f_{\alpha\beta} \mapsto f_{\alpha\beta} + \partial_\alpha \xi_\beta + \partial_\beta \xi_\alpha, \quad (3.31)$$

with  $\xi_\alpha$  a being Lorentz covector field. The field equations of Equation (3.30) are

$$\square f_{\alpha\beta} - \partial_\alpha \partial_\gamma f^\gamma{}_\beta - \partial_\beta \partial_\gamma f^\gamma{}_\alpha + \partial_\alpha \partial_\beta f^\gamma{}_\gamma + \eta_{\alpha\beta} [\partial^\gamma \partial^\delta f_{\gamma\delta} - \square f^\gamma{}_\gamma] = 0. \quad (3.32)$$

The 2-Killing vector reduction of the action (3.30) amounts to considering field configurations obeying  $\mathcal{L}_{K_1} f_{\alpha\beta} = 0 = \mathcal{L}_{K_2} g_{\alpha\beta}$  (which is equivalent to  $g_{\alpha\beta} = \eta_{\alpha\beta} + f_{\alpha\beta}$  having two Killing vectors).

In adapted coordinates where  $K_a = \partial/\partial y^a$ ,  $a = 1, 2$ , the metric components depend only on the non-Killing coordinates  $(x^0, x^1)$ . Moreover  $f_{\alpha\beta}$  can be assumed to be block-diagonal

$$f_{\alpha\beta} = \begin{pmatrix} f_{\mu\nu} & 0 \\ 0 & f_{ab} \end{pmatrix}, \quad \eta_{\alpha\beta} = \begin{pmatrix} \eta_{\mu\nu} & 0 \\ 0 & \delta_{ab} \end{pmatrix}, \quad (3.33)$$

where  $\mu, \nu \in \{0, 1\}$ ,  $a, b \in \{2, 3\}$ , and  $\eta_{\mu\nu}$  has eigenvalues  $(-, +)$ . The component fields  $f_{22}$ ,  $f_{33}$ ,  $f_{23}$  retain their interpretation as parameterizing the norms and inner product of the Killing vectors, i.e.  $K_1 \cdot K_1 = 1 + f_{22}$ ,  $K_2 \cdot K_2 = 1 + f_{33}$ ,  $K_1 \cdot K_2 = f_{23}$ . Since  $(K_1 \cdot K_1)(K_2 \cdot K_2) - (K_1 \cdot K_2)^2 = 1 + f_{22} + f_{33} + O(f^2)$  is nonlinear in the perturbations, a (re-)parameterization in terms of a radius field  $\rho$  as in the nonlinear theory is less useful. Nevertheless the combination  $f_{22} + f_{33}$  will play a special role later on.

Entering with the ansatz (3.33) into Equation (3.30) gives the reduced action

$$\begin{aligned} S''[f] &= -\frac{1}{4\lambda} \int d^2x \left[ 2\partial^\mu f^a{}_a (\partial^\nu f_{\nu\mu} - \partial_\mu f^\nu{}_\nu) + \partial^\mu f^{ab} \partial_\mu f_{ab} - \partial^\mu f^a{}_a \partial_\mu f^b{}_b \right] \\ &= -\frac{1}{4\lambda} \int d^2x \left[ f_{\mu\nu}^T \partial_\mu \partial_\nu (k - f) - f^{T\,ab} \partial^2 f_{ab}^T - \frac{1}{4} k \partial^2 f - \frac{1}{8} k \partial^2 k + \frac{3}{8} f \partial^2 f \right]. \end{aligned} \quad (3.34)$$

Here the Greek indices  $\mu, \nu, \dots$  are raised with  $\eta^{\mu\nu}$ , and the latin  $a, b, \dots$  with  $\delta^{ab}$ . As in Equation (3.10) we also included Newton's constant per unit volume of the internal space as a prefactor. The second line of Equation (3.34) is the counterpart of the second line of Equation (3.30) and we wrote

$$f_{\mu\nu}^{\text{TF}} = f_{\mu\nu}^T + \frac{1}{2} \eta_{\mu\nu} k, \quad f_{ab}^{\text{TF}} = f_{ab}^T - \frac{1}{2} \delta_{ab} k, \quad k := f^\mu{}_\mu - f^a{}_a, \quad (3.35)$$

where  $\eta^{\mu\nu} f_{\mu\nu}^T = 0 = \delta^{ab} f_{ab}^T$  are the individually tracefree parts of the blocks in Equation (3.33). Their traces  $f^\mu{}_\mu$  and  $f^a{}_a$  have been replaced by  $k$  and  $f = f^\alpha{}_\alpha = f^\mu{}_\mu + f^a{}_a$  in Equation (3.34). The wrong sign of the trace component  $f$  of course remains.

We add some remarks. As one might expect, the action (3.34) can also be obtained by “first reducing” and then “linearizing”. Indeed, from Equation (3.50) below one has

$$\rho \sqrt{-\gamma} R^{(2)}(\gamma) \Big|_{\gamma_{\mu\nu} = \eta_{\mu\nu} + f_{\mu\nu}} \simeq \partial^\mu \rho - \partial^\nu f_{\nu\mu} + \partial_\mu f^\nu{}_\nu + O(\rho f^2), \quad (3.36)$$

modulo total derivatives, and with the identifications

$$\begin{aligned} \rho &= 1 + \frac{1}{2}(f_{22} + f_{33}) + O(f^2), \\ \Delta &= 1 + \frac{1}{2}(f_{22} - f_{33}) + O(f^2), \\ \psi &= f_{23} + O(f^2). \end{aligned} \quad (3.37)$$

one recovers (3.34) from Equation (3.10). Second, the action (3.34) is invariant under the gauge transformations

$$f_{\mu\nu} \mapsto f_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu, \quad (3.38)$$

which corresponds to Lorentz vector fields  $\xi_\alpha$  of the form  $\xi_\alpha = (\xi_\mu(x^0, x^1), 0, 0)$ . In addition there are the remnants of the  $\text{SL}(2, \mathbb{R})/\text{SO}(2)$  isometries which now generate the isometries group  $\text{ISO}(2)$  of  $\mathbb{R}^2$ : two translations  $f_{22} \mapsto f_{22} + a_1$ ,  $f_{33} \mapsto f_{33} - a_1$  and  $f_{23} \mapsto f_{23} + a_2$ , and one rotation

$$f_{22} \mapsto f_{22} + 2\alpha f_{23}, \quad f_{33} \mapsto f_{33} - 2\alpha f_{23}, \quad f_{23} \mapsto f_{23} - \alpha(f_{22} - f_{33}), \quad (3.39)$$

with  $\alpha$  constant and modulo  $O(\alpha^2)$  terms.

Further the operations “varying the action” and “reduction” are commuting, as expected from the principle of symmetric criticality. Thus, the reduction of the field equations (3.32) coincides with the field equations obtained by varying Equation (3.34). The latter are

$$\begin{aligned}\partial^2 f_{23} &= 0, \\ \partial^2 f_{22} &= \partial^2 f_{33} = \partial^2 f^\nu{}_\nu - \partial^\mu \partial^\nu f_{\mu\nu}, \\ [\partial_\mu \partial_\nu - \eta_{\mu\nu} \partial^2](f_{22} + f_{33}) &= 0,\end{aligned}\tag{3.40}$$

with  $\partial^2 = \eta^{\mu\nu} \partial_\mu \partial_\nu$ . These are equivalent to  $\partial^2 f_{23} = \partial^2 f_{22} = \partial^2 f_{33} = 0$ ,  $(\partial^\mu \partial^\nu - \eta^{\mu\nu} \partial^2) f_{\mu\nu} = 0$ , and

$$f_{22} + f_{33} = x^\mu v_\mu + v_0,\tag{3.41}$$

for integration constants  $v_\mu$  and  $v_0$ . A nonzero  $v_\mu$  (which is not co-rotated) violates Poincaré invariance and is undesirable for most purposes. Equation (3.37) shows that on shell only two of the three free fields  $f_{22}$ ,  $f_{33}$ ,  $f_{23}$  are independent. In fact, in a Hamiltonian formulation of Equation (3.34) the condition  $f_{22} + f_{33} = \text{const}$  can be understood as a constraint associated with the gauge invariance (3.36). Equivalently it can be viewed as a remnant of the constraints (3.15). Indeed, to linear order  $\mathcal{H}_0 = -2\partial_1^2 \rho / \lambda$ ,  $\mathcal{H}_1 = -2\partial_1 \pi_\sigma$ , with  $\pi_\sigma = \partial_0 \rho$ , which combined with Equation (3.37) gives  $f_{22} + f_{33} = \text{const}$ , once again.

For the special case of the Einstein–Rosen waves an on-shell formulation suffices and a related study linking the ‘graviton modes’ to the ‘Einstein–Rosen modes’ can be found in [23].

The linearized theory is well suited to discuss the physics content of the 2-Killing vector reduction. To this end one fixes the gauge and displays the independent on-shell degrees of freedom. The most widely used gauge in the linearized theory (3.30) is the transversal-traceless gauge,

$$\begin{aligned}\partial^\alpha \left[ f_{\alpha\beta} - \frac{1}{2} \eta_{\alpha\beta} f^\gamma{}_\gamma \right] &= 0, \\ f_{\alpha 0} &= 0, \\ f^\gamma{}_\gamma &= 0.\end{aligned}\tag{3.42}$$

The first condition is the “harmonic gauge condition”, the others are “transversal-traceless” conditions. Using the latter in the former gives  $\partial_i f_{ij} = 0$ , with  $i, j \in \{1, 2, 3\}$ . Thus Equation (3.42) removes  $3 + 4 + 1$  degrees of freedom of the 10 components of  $f_{\alpha\beta}$  leaving the familiar 2 graviton degrees of freedom. In the 2-Killing subsector the gauge conditions read

$$\begin{aligned}\partial^\mu [2f_{\mu\nu} - \eta_{\mu\nu} f^\rho{}_\rho + \eta_{\mu\nu} \delta^{ab} f_{ab}] &= 0, \\ f_{00} = f_{01} &= 0, \\ f_{11} + f_{22} + f_{33} &= 0.\end{aligned}\tag{3.43}$$

Using again the second set in the harmonic gauge condition gives  $f_{22} + f_{33} = \text{const}$ . Thus Equation (3.43) removes  $1 + 3$  of the 6 degrees of freedom in  $f_{\mu\nu}$ ,  $f_{ab}$ , leaving again the familiar two ‘graviton’ degrees of freedom,  $f_{22}$ ,  $f_{23}$ , say. The only difference to the physical degrees of freedom of the full linearized theory (3.30) is that all wave vectors are aligned, that is, the gravitons are collinear. To see this recall that the general solution of Equation (3.32) subject to Equation (3.42) is a superposition of plane waves  $f_{\alpha\beta}(x) = A_{\alpha\beta}(k) e^{ik_\gamma x^\gamma} + \text{c.c.}$ , with  $k_\gamma k^\gamma = 0$  and  $A_{\alpha\beta}(k)$  constrained by Equation (3.42). Both the action (3.30) and the field equations (3.31) are Poincaré invariant, so any one wave vector in a superposition can be fixed to have the form  $k = (k_0, k_1, 0, 0)$ , with  $k_\mu k^\mu = 0$ . Identifying  $(x^0, x^1)$  with the non-Killing coordinates and  $(x^2, x^3)$  with the Killing coordinates  $(y^2, y^3)$ , the waves in the 2-Killing subsector are of identically the same form; the only



difference is that now *all* of the wave vectors have the form  $k = (k_0, k_1, 0, 0)$ . In other words, all waves move in the same or in the opposite direction; they are collinear. This collinearity of the wave vectors must of course not be confused with the alignment of the polarization vectors; the polarization tensor  $A_{\alpha\beta}(k)$  here always carries two independent polarizations. Nontrivial scattering is possible despite the collinearity of the waves. It should be interesting to compute this S-matrix and to contrast it with the one in the Eikonal sector [209, 113, 73, 71].

In summary we conclude that the 2-Killing vector subsector comprises the gravitational self-interaction of collinear gravitons, in the same sense as the full Einstein–Hilbert action describes the self-interaction of non-collinear gravitons.

The formulation of the perturbative functional integral in this subsector would now proceed in exact parallel to the non-collinear case: A gauge fixing term

$$\frac{1}{2\alpha} \int d^2x Q^\mu Q_\mu, \quad Q_\nu := \partial^\mu [2f_{\mu\nu} - \eta_{\mu\nu} f^\rho{}_\rho + \eta_{\mu\nu} \delta^{ab} f_{ab}], \quad (3.44)$$

implementing the harmonic gauge condition is added to the action. This renders the kinetic term in Equation (3.34) nondegenerate. The extra propagating degrees of freedom are then ‘canceled out’ by the Faddeev–Popov determinant. In principle a systematic collinear graviton loop expansion could be set up in this way, much in parallel to the generic non-collinear case.

### 3.2.2 Dirac versus covariant quantization

For the systems at hand we now want to argue that this is not the method of choice. To this end we return to the covariant action (3.10) and decompose  $\gamma_{\mu\nu}$  into a conformal factor  $e^\sigma$  and a two-parametric remainder  $\hat{\gamma}_{\mu\nu}$  to be adjusted later,  $\gamma_{\mu\nu} = e^\sigma \hat{\gamma}_{\mu\nu}$ . Using  $R^{(2)}(e^\sigma \hat{\gamma}) = e^{-\sigma} [R^{(2)}(\hat{\gamma}) - \hat{\gamma}^{\mu\nu} \hat{\nabla}^\mu \hat{\nabla}^\nu \sigma]$  we can rewrite Equation (3.10) as a gravity theory for the two-parametric  $\hat{\gamma}_{\mu\nu}$

$$S = \frac{1}{2\lambda} \int d^2x \sqrt{\hat{\gamma}} [2\rho R(\hat{\gamma}) + \hat{\gamma}^{\mu\nu} \partial_\mu \rho \partial_\nu (2\sigma + \ln \rho) - \hat{\gamma}^{\mu\nu} (\partial_\mu \Delta \partial_\nu \Delta + \partial_\mu \psi \partial_\nu \psi) \rho \Delta^{-2}]. \quad (3.45)$$

There are two instructive choices for  $\hat{\gamma}_{\mu\nu}$ ,

$$\hat{\gamma}_{\mu\nu} = \begin{pmatrix} -n^2 + s^2 & s \\ s & 1 \end{pmatrix}, \quad (3.46)$$

$$\hat{\gamma}_{\mu\nu} = \bar{\gamma}_{\mu\rho} e^{-\bar{\sigma}} (e^{f^T})^\rho{}_\nu. \quad (3.47)$$

The first choice is the densitized lapse-shift parameterization already used before (see Section 3.1.4). This is adapted to a proper time or Dirac quantization: The lapse  $n$  and shift  $s$  are classically nonpropagating degrees of freedom. In a Dirac quantization one can simply fix the temporal gauge ( $n = 1$ ,  $s = 0$ ) and use the nondegenerate gauge-fixed action in Equation (3.45) to define the quantum theory. The Hamiltonian and diffeomorphism constraint arising from the gauge fixing then have to be defined as composite operators.

The choice (3.47) is adapted to a covariant quantization. Here  $\bar{\gamma}_{\mu\nu} e^{\bar{\sigma}}$  is a generic (off-shell) background metric with again the conformal mode  $\bar{\sigma}$  split off. The fluctuation field  $f_{\mu\nu}^T$  is trace-free with respect to the background  $\bar{\gamma}^{\mu\nu} f_{\mu\nu}^T = 0$ . Then Equation (3.45) describes a unimodular gravity theory ( $\det e^{f^T} = 1$ ) and the original metric is parameterized by  $f_{\mu\nu}^T$  and  $f_\sigma := \sigma - \bar{\sigma}$  as  $\gamma_{\mu\nu} = \bar{\gamma}_{\mu\rho} (e^{f^T})^\rho{}_\nu e^{f_\sigma}$ . In a covariant formulation the degrees of freedom in  $f^T$  would be promoted to propagating ones by adding a gauge fixing term to the action (3.45). The associated Faddeev–Popov determinant is designed to cancel out their effect again. In the case at hand this is clearly roundabout as the gauge-frozen Lagrangian (3.45) is already nondegenerate.

This setting can be promoted to a generalization of the one presented in Section 3.1.2 to generic backgrounds. In the terminology of Section B.2 one then gets a non-geodesic background-fluctuation split, which treats the nonpropagating lapse and shift degrees of freedom on an equal footing with the others. In order to contrast it with the geodesic background-fluctuation split for the propagating modes used later on, we spell out here the first few steps of such a procedure.

We wish to expand the Einstein–Hilbert action for the class of metrics (3.6) around a generic background. Technically it is simpler to “first reduce” and then “expand”. This is legitimate since all operations involved are algebraic. Recall that the reduced Lagrangian is defined by inserting the ansatz

$$g_{\alpha\beta} = \begin{pmatrix} \gamma_{\mu\nu} & 0 \\ 0 & M_{ab} \end{pmatrix}, \quad M_{ab} = \frac{\rho}{\Delta} \begin{pmatrix} \Delta^2 + \psi^2 & \psi \\ \psi & 1 \end{pmatrix} \quad (3.48)$$

with  $\gamma_{\mu\nu}$ , for example, in the above lapse-shift parameterization into the Einstein–Hilbert Lagrangian. This gives Equation (3.10). For the 2D metric a convenient parameterization of the fluctuations around a generic background is Equation (3.47) which for  $\gamma_{\mu\nu}$  amounts to

$$\gamma_{\mu\nu} = \bar{\gamma}_{\mu\rho} (e^{f^T})^\rho{}_\nu e^{f^\sigma} = \bar{\gamma}_{\mu\nu} + f_{\mu\nu}^T + f_\sigma \bar{\gamma}_{\mu\nu} + \dots, \quad (3.49)$$

where  $\bar{\gamma}^{\mu\nu} f_{\mu\nu}^T = 0$  and indices are raised and with the (inverse) background metric  $\bar{\gamma}_{\mu\nu}$ . Since  $\bar{\gamma}_{\mu\nu}$  is diffeomorphic to  $e^{\bar{\sigma}}$  one has  $\sigma = \bar{\sigma} + f_\sigma$ .

The counterpart of Equation (3.49) for the lower  $2 \times 2$  block in Equation (3.48) is

$$M_{ab} = \bar{M}_{ac} (e^{f^T})^c{}_b e^{f^\rho} = \bar{M}_{ab} + f_{ab}^T + f_\rho \bar{M}_{ab} + \dots, \quad (3.50)$$

where  $\bar{M}^{ab} f_{ab}^T = 0$ . Writing  $\bar{\rho} = \sqrt{\det \bar{M}}$  one has  $\rho = \bar{\rho} e^{f^\rho}$ . The trace-free part  $f_{ab}^T$  defines an element of the Lie algebra of  $\text{SL}(2, \mathbb{R})$ , though in a nonstandard basis. Writing

$$\Delta = \bar{\Delta} + \bar{\rho}^{-1} f_1, \quad \psi = \bar{\psi} + \bar{\rho}^{-1} f_2, \quad \rho = \bar{\rho}(1 + f_\rho), \quad (3.51)$$

one finds  $M_{ab} = \bar{M}_{ab} + f_{ab}^T + f_\rho \bar{M}_{ab} + O(f^2)$ , with

$$f_{ab}^T = f_1 \frac{1}{\bar{\Delta}^2} \begin{pmatrix} \bar{\Delta}^2 - \bar{\psi}^2 & -\bar{\psi} \\ -\bar{\psi} & -1 \end{pmatrix} + f_2 \frac{1}{\bar{\Delta}} \begin{pmatrix} 2\bar{\psi} & 1 \\ 1 & 0 \end{pmatrix}. \quad (3.52)$$

The decomposition (3.49, 3.50) can be viewed as the counterpart of the York decomposition for the metrics (3.6). Note that conventional linear background fluctuation split would only keep the leading terms in Equations (3.49, 3.50) and base the expansion on this split.

Writing  $L_{\text{EH}}(g)$  for Einstein–Hilbert Lagrangian with the blockdiagonal metric (3.48) one arrives at an expansion of the form

$$L_{\text{EH}}(g) = L_{\text{EH}}(\bar{g}) + \sum_{n \geq 1} L_n(f_{\mu\nu}^T, f_\sigma, f_{ab}^T, f_\rho; \bar{\gamma}_{\mu\nu}, \bar{M}_{ab}), \quad (3.53)$$

where  $L_n$  is of order  $n$  in the fluctuations. This holds both when a York-type decomposition is adopted and when a linear split is used, just the higher order terms  $L_n$ ,  $n \geq 3$ , get reshuffled. We won't need their explicit form; the point relevant here is the invariance of  $L_2$  under the genuine gauge transformations in  $f_{\mu\nu}$  and the background counterpart of the  $\text{ISO}(2)$  rotations (3.39). The genuine gauge invariance has to be gauge-fixed in order to get nondegenerate kinetic terms. The standard gauge fixing is the background covariant harmonic one,  $\bar{\nabla}^\mu (f_{\mu\nu} - \frac{1}{2} \bar{\gamma}_{\mu\nu} \bar{\gamma}^{\rho\sigma} f_{\rho\sigma}) = 0$ . The extra propagating degrees of freedom are canceled by the Faddeev–Popov determinant, and the construction of the perturbative functional integral proceeds as usual. One can take as the central object in the quantum theory the background effective action  $\Gamma_{\text{B}}[\langle \gamma_{\mu\nu} \rangle, \langle M_{ab} \rangle; \bar{\gamma}_{\mu\nu}, \bar{M}_{ab}]$ ,

which as far as the symmetries are concerned of the type characteristic for a linear background-fluctuation split. Note that this feature would not change if a nonperturbative construction of  $\Gamma_{\text{B}}[\langle\gamma_{\mu\nu}\rangle, \langle M_{ab}\rangle; \bar{\gamma}_{\mu\nu}, \bar{M}_{ab}]$  was aimed at, for example via a functional renormalization group equation.

In all cases one sees that the procedure outlined has two drawbacks. First, the split (3.49) ignores the special status of the lapse and shift degrees of freedom in  $\gamma_{\mu\nu}$ ; all components are expanded. We know, however, that there must be two infinite series built from the components of  $\bar{\gamma}_{\mu\nu}$  and  $f_{\mu\nu}$  that enter the left-hand-side of Equation (3.53) anyhow linearly. Concerning the lower  $2 \times 2$  block in Equation (3.48) both the linear and the York-type decomposition will only keep the ISO(2) symmetry more or less manifest. The nonlinear realization of the SL(2,  $\mathbb{R}$ ) symmetry then has to be restored through Ward identities, iteratively in a perturbative formulation or otherwise in a nonperturbative one.

In the following we shall adopt the following remedies. The fact that the lapse and shift degrees of freedom in  $\gamma_{\mu\nu}$  enter the left-hand-side of Equation (3.53) linearly of course just means that they are the Lagrange multipliers of the constraints in a Hamiltonian formulation. The linearity can thus be exploited either by a gauge fixing with respect to these variables before expanding, giving rise to a proper time formulation, or by directly adopting a Dirac quantization procedure. By and large both should be equivalent; in [154, 155] a direct Dirac quantization was used, and we shall describe the results in the next two Sections 3.3 and 3.4. With the lapse and shift in Equation (3.14) ‘gone’ one only needs to perform a background-fluctuation split only for the remaining propagating fields  $\Delta, \psi, \rho, \sigma$ .

To cope with the second of the before-mentioned drawbacks we equip – following deWitt and Vilkovisky – this space of propagating fields with a pseudo-Riemannian metric and perform a normal-coordinate expansion around a (‘background’) point with respect to it. This leads to the formalism summarized in Section B.2.2. The pseudo-Riemannian metric on the space of propagating fields can be read off from Equation (3.16) and converts the gauge-frozen but nondegenerate Lagrangian into that of a (pseudo-)Riemannian nonlinear-sigma model. The renormalization theory of these systems is well understood and we summarize the aspects needed here in Section B.3. In exchange for the gauge-freezing one then has to define quantum counterparts of the constraints (3.15) as renormalized composite operators. This will be done in Equations (3.100) ff.

For the sake of comparison with Equations (3.49, 3.51) we display here the first two terms of the resulting geodesic background-fluctuation split:

$$\begin{aligned}\Delta &= \bar{\Delta} + \xi^1 + \frac{1}{2\bar{\Delta}}(\xi^{12} - \xi^{22}) - \frac{1}{\bar{\rho}}\xi^1\xi^3 + \dots, \\ \psi &= \bar{\psi} + \xi^2 + \frac{1}{\bar{\Delta}}\xi^1\xi^2 - \frac{1}{2\bar{\rho}}\xi^2\xi^3 + \dots, \\ \rho &= \bar{\rho} + \xi^3 + \dots, \\ \sigma &= \bar{\sigma} + \xi^4 + \frac{1}{4b\bar{\Delta}}(\xi^{12} + \xi^{22}) + \frac{a}{4b\bar{\rho}^2}\xi^3{}^2 + \dots\end{aligned}\tag{3.54}$$

Here  $\bar{\phi} = (\bar{\Delta}, \bar{\psi}, \bar{\rho}, \bar{\sigma})$  is the reference point in field space and  $(\xi^1, \xi^2, \xi^3, \xi^4)$  is the tangent vector at  $\bar{\phi}$  to the geodesic connecting  $\bar{\phi}$  and  $\phi = (\Delta, \psi, \rho, \sigma)$ ; the dots indicate terms of cubic and higher order in the  $\xi^i$ . The geodesic in question is defined with respect to the pseudo-Riemannian metric in field space  $\mathfrak{h}_{ij}$  (see Equation (3.60) below). This metric in field space possesses a number of Killing vectors  $Y$  (not to be confused with the Killing vectors of the spacetime geometries considered) and two conformal Killing vectors related to Equation (3.27). A major advantage of the geodesic background fluctuation split is that the associated generalized Ward identities are built into the formalism. Indeed the diffeomorphism Ward identities (B.24, B.39) become Ward identities proper

for the isometries of the target space and “conformal” generalizations thereof for the conformal isometries.

In summary, we find the following differences to standard perturbation theory:

1. Lapse and shift viewed as infinite series in the fluctuation field are not expanded. Only the metric degrees of freedom other than lapse and shift are expanded.
2. Through the use of the background effective action formalism the expectation of the quantum metric  $\langle g_{\alpha\beta} \rangle$  and the background  $\bar{g}_{\alpha\beta}$  are related by the condition

$$\frac{\delta\Gamma_{\text{B}}[\langle g_{\alpha\beta} \rangle; \bar{g}_{\alpha\beta}]}{\delta\langle g_{\alpha\beta} \rangle} = 0. \quad (3.55)$$

One does not expand around a solution of the classical field equations.

3. Through the use of a geodesic background-fluctuation split on the space of propagating fields the resulting background effective action is in principle invariant under arbitrary local reparameterizations of the propagating fields. Among those the ones associated with isometries or conformal isometries on field space are of special interest and give rise to Ward identities associated with the Noether currents (3.26) and the conformal currents (3.27). The latter are built into the formalism, and do not have to be imposed order by order.

The first point entails that the sigma-model perturbation theory we are going to use is *partially non-perturbative* from the viewpoint of a standard graviton loop expansion.

### 3.2.3 Conformal factor

Before turning to the quantum theory of these warped product sigma-models we briefly discuss the status of the conformal factor instability in a covariant formulation of the  $2+2$  truncation. As emphasized by Mazur and Mottola [142] in linearized Euclidean quantum Einstein Gravity (based on the Euclidean version of the action (3.30)) there is really no conformal factor instability. The  $f \partial^2 f$  kinetic term in the second part of Equation (3.30) with the wrong sign receives an extra contribution from the measure which after switching to gauge invariant variables renders both the Gaussian functional integral over the conformal factor and that for the physical degrees of freedom well-defined. They also gave a structural argument why this should be so even on a nonlinear level: As one can see from a canonical formulation the conformal factor in Einstein gravity is really a constrained degree of freedom and should not have a canonically conjugate momentum.

In the  $2+2$  truncation we shall use a Lorentzian functional integral defined through the sigma-model perturbation theory outlined above. So a conformal factor instability proper associated with a Euclidean functional integral anyhow does not arise. Nevertheless it is instructive to trace the fate of the incriminated  $f \partial^2 f$  term.

From the York-type decomposition (3.49, 3.50) one sees that  $f_\sigma + f_\rho = \frac{1}{2}(f^\mu{}_\mu + f^a{}_a) =: \frac{1}{2}f$  plays the role of the (gauge-variant) conformal factor. The wrong sign kinetic term is indeed still present in the second part of Equation (3.34) and  $f$  also appears through a dilaton type coupling in the  $f_{\mu\nu}^T \partial_\mu \partial_\nu (k - f)$  term. In 2D however  $f_{\mu\nu}^T$  has no propagating degrees of freedom and the term could be taken care of promoting  $f_{\mu\nu}^T$  to a dynamical degree of freedom via gauge fixing and then cancelling the effect by a Faddeev–Popov determinant. As already argued before it is better to avoid this and look at the remaining propagating degrees of freedom directly. They simplify when reexpressed in terms of  $f_\rho = (f + k)/4$  and  $f_\sigma = (f - k)/4$ , viz.  $\frac{1}{4}\partial k \partial f + \frac{1}{8}(\partial k)^2 - \frac{3}{8}(\partial f)^2 = -4\partial f_\sigma \partial f_\rho - 2(\partial f_\rho)^2$ . This occurs here on the linearized level but comparing with Equation (3.16) one sees that the same structure is present in the full gauge-frozen action. We thus consider from now on directly the corresponding terms proportional to  $-\partial\rho \partial(\sigma + \frac{1}{2} \ln \rho)$ . By a local redefinition

of  $\sigma$  one can eliminate the term quadratic in  $\rho$  and in dimensional regularization used later no Jacobian arises. One is left with a  $-\partial\rho\partial\sigma$  term which upon diagonalization gives rise to one field whose kinetic term has the wrong sign. However  $\rho$  is a dilaton type field which multiplies all of the self-interacting positive energy scalars in the first term of Equation (3.16), and the dynamics of this mode turns out to be very special (see Section 3.3). Heuristically this can be seen by viewing the  $\sigma$  field in the Lorentzian functional integral simply as a Lagrange multiplier for a  $\delta(\partial^2\rho)$  insertion. The remaining Lorentzian functional integral would allow for a conventional Wick rotation with a manifestly bounded Euclidean action. We expect that roughly along these lines a non-perturbative definition of the functional integral for Equation (3.16) could be given, which would clearly be one without any conformal factor instability. Within the perturbative construction used in Section 3.3 the special status of the  $\partial^2\rho$  field, viewed as a renormalized operator, can be verified. Since the system is renormalizable only with infinitely many couplings, the functional dependence on  $\rho$  in the renormalized Lagrangian and in the  $\partial^2\rho$  field has to be ‘deformed’ in a systematic way; however this does not affect the principle aspect that no instability occurs.

Finally, let us briefly comment on the role of Newton’s constant and of the cosmological constant in the  $2+2$  truncations. The gravity part of the action (3.10) or (3.45) arises from evaluating the Einstein–Hilbert action  $S_{\text{EH}}$  on the class of metrics (3.6). The constant  $1/\lambda$  in Equation (3.10, 3.45) can be identified with  $d^2y/g_{\text{N}}$ , i.e. with Newton’s constant per unit volume of the orbits. As such  $\lambda$  is an inessential parameter and its running is defined only relative to a reference operator. For the  $2+2$  truncations it turns out that the way how the action (3.10) depends on  $\rho$  has to be modified in a nontrivial and scale dependent way by a function  $h(\cdot)$  (see Equation (3.56) below) in order to achieve strict cut-off independence. This modification amounts to the inclusion of infinitely many essential couplings, only the overall scale of  $h(\cdot)$  remains an inessential parameter. It is thus convenient not to renormalize this overall scale and to treat  $\lambda$  in Equation (3.56) as a loop counting parameter.

A similar remark applies to the cosmological constant. Adding a cosmological constant term to the Ricci scalar term results in a  $\Lambda\rho e^\sigma$  type addition to Equation (3.56) below. In the quantum theory one is again forced to replace  $\rho$  with an scale dependent function  $f(\rho)$  in order to achieve strict cutoff independence [156]. The cosmological constant proper can be identified with the overall scale of the function  $f(\cdot)$ . The function  $f$  is subject to a non-autonomous flow equation, triggered by  $h$ , but if its initial value is set to zero it remains zero in the course of the flow [156]. To simplify the exposition we thus set  $f \equiv 0$  from the beginning and omit the cosmological constant term in the following. It is however a nontrivial statement that this can be done in a way compatible with the renormalization flow.

### 3.3 Tamed non-renormalizability

For the reasons explained in the Appendices B.2.2 and B.3 we now study the quantum theory based on the sigma-model Lagrangian (3.16) in the setting of the covariant background field expansion. Since this is a well-tested formalism (see Section B.3) it has the additional advantage that any unexpected findings cannot be blamed on the use of an untested formalism. Technically it is also convenient to use dimensional regularization and minimal subtraction. The analysis can then be done to all orders of sigma-model perturbation theory.

Our first goal thus is to construct the infinite cut-off limit of the background effective action  $\lim_{\Lambda \rightarrow \infty} \Gamma_\Lambda[\langle g_{\alpha\beta} \rangle, \bar{g}_{\alpha\beta}]$  in the covariant background field formalism to all orders of the loop expansion. It turns out that this can be done *only* if infinitely many essential couplings are allowed, so even the truncated functional integral based on Equation (3.16) is not renormalizable in the strict sense. However, once one allows for infinitely couplings strict cutoff independence ( $\Lambda \rightarrow \infty$ ) can be achieved. Remarkably, for Equation (3.1) the generalized beta function for a generating functional of these couplings can be found in closed form (see Equation (3.88) below). This allows

one to study their RG flow in detail and to prove the existence of a non-Gaussian UV stable fixed point. One also finds a Gaussian fixed point which is not UV stable.

The main principle guiding the renormalization are generalized Ward identities for the Noether currents (3.26) and the conformal currents (3.27). The general solution of these generalized Ward identities suggests a space of actions which is stable under the renormalization flow. One finds that by suitable redefinitions the general solution can always be brought into the form

$$L_h(\varphi, \rho, \sigma) = -\frac{1}{2\lambda} h(\rho) [\mathbf{m}_{ij}(\varphi) \partial^\mu \varphi^i \partial_\mu \varphi^j - 2\rho^{-1} \partial^\mu \rho \partial_\mu \sigma - \rho^{-2} \partial^\mu \rho \partial_\mu \rho]. \quad (3.56)$$

Here  $h$  is the generating functional for an infinite set of couplings; it defines a function of one real variable whose argument in Equation (3.56) is the field  $\rho$ . The fact that the renormalization flow does not force one to leave the (infinite-dimensional) space of actions, Equation (3.56) of course has to be justified by explicit construction. Below we shall show this in renormalized perturbation theory to all loop orders. Since the ansatz (3.56) is based on a symmetry characterization it seems plausible, however, that also a nonperturbatively constructed flow would not force one to leave the space (3.56), though one could certainly start off with a more general ansatz containing for example higher derivative terms.

As described above in the sigma-model perturbation theory we use dimensional regularization and minimal subtraction. Counter terms will then have poles in  $d - 2$  rather than containing positive powers of the cutoff  $\Lambda$ . The role of the scale  $k$  is played by the renormalization scale  $\mu$ , the fields and the couplings at the cutoff scale are called the “bare” fields and the “bare” couplings, while the fields and couplings at scale  $\mu$  are referred to as “renormalized”. The fact that the ansatz (3.56) ‘works’ is expressed in the following result:

**Result** (Generalized renormalizability) [154, 155]:

To all orders in the sigma-model loop expansion there exist nonlinear field renormalizations  $(\varphi_B, \rho_B, \sigma_B) \mapsto (\varphi, \rho, \sigma)$  such that for any prescribed bare generating coupling functional  $h_B(\cdot)$  there exists a renormalized  $h(\cdot)$  such that

$$L_{h_B}(\varphi_B, \rho_B, \sigma_B) = L_h(\varphi, \rho, \sigma). \quad (3.57)$$

Both coupling functionals are related by

$$h_B(\cdot) = \mu^{d-2} h(\cdot) \left[ 1 + \frac{1}{2-d} \sum_{l \geq 1} \left( \frac{\lambda}{2\pi} \right)^l H_l[h](\cdot) + \dots \right]. \quad (3.58)$$

The  $H_l[h]$  are explicitly known functionals of the renormalized  $h$  (see Equation (3.70) below). The nonlinear field renormalizations are likewise explicitly known and given in Equations (3.65, 3.70).

A subscript ‘ $B$ ’ denotes the bare fields while the plain symbols refer to the renormalized ones and similarly for  $h$ . Notably no higher order derivative terms are enforced by the renormalization process; strict cutoff independence can be achieved without them. However the fact that  $h_B(\cdot)$  and  $h(\cdot)$  differ marks the deviation from conventional renormalizability. The pre-factor  $h(\rho)/\rho$  also has a physical interpretation: To lowest order it is for pure gravity the conformal factor in a Weyl transformation  $g_{\alpha\beta}(x) \rightarrow e^{\omega(\rho(x))} g_{\alpha\beta}(x)$  of a generic four-dimensional metric  $g_{\alpha\beta}(x)$  with two Killing vectors in adapted coordinates.

The derivation of this result is based on a reformulation of the class of QFTs based on Equation (3.56) as a Riemannian sigma-model in the sense of Friedan [84]. This is a class of two-dimensional QFTs which is also (perturbatively) renormalizable only in a generalized sense, namely by allowing for infinitely many relevant couplings. The generating functional for these couplings in



this case is a (pseudo-)Riemannian metric  $\mathfrak{h}$  on a “target manifold  $\mathcal{M}$ ” of arbitrary dimension  $D$  and field coordinates  $\phi : \Sigma \rightarrow \mathcal{M}$ , where  $\Sigma$  is the two-dimensional “base manifold”. The renormalization theory of these systems is well understood. A brief summary of the results relevant here is given in Appendix B.3.

The systems (3.56) can be interpreted as Riemannian sigma models where the target manifold is a special class of “warped products” (see Equation (3.59) below) and the fields are  $\phi = (\varphi, \rho, \sigma)$ . The relation between the quantum theory of these Riemannian sigma-models and the QFT based on Equation (3.56) will roughly be that one performs an infinite reduction of couplings in a sense similar to [236, 160, 174]. The generating functional  $\mathfrak{h}_{ij}(\phi)$  is parameterized by  $\frac{D(D-1)}{2}$  functions of  $D$  variables, while the generating functional  $h$  in Equation (3.56) amounts to one function of one variable. Thus “ $\frac{D(D-1)}{2} \times D \times \infty$ ” many couplings are reduced to “ $1 \times \infty$ ” many couplings. As always in a reduction of couplings the nontrivial point is that this reduction can be done in a way compatible with the RG dynamics. The original construction in [236, 160] was in the context of strictly renormalizable QFTs with a finite number of relevant couplings. In a QFT with infinitely many relevant couplings (QCD in a lightfront formulation) the reduction principle was used by Perry–Wilson [174]. A general study of an ‘infinite reduction’ of couplings has been performed in [11].

The reduction technique used here is different, but essentially Equation (3.68) below plays the role of the reduction equation. Apart from the different derivation and the fact that the reduction is performed on the level of generating functionals, the main difference to a usual reduction is that Equation (3.68) also involves nonlinear field redefinitions without which the reduction could *not* be achieved here. The reduction equation (3.68) thus *mixes* field redefinitions and couplings. From the viewpoint of Riemannian sigma-models this amounts to the use of *metric dependent* diffeomorphisms on the target manifold, a concept neither needed nor used in the context of Riemannian sigma-models otherwise.

Since Riemannian sigma-models have been widely used in the context of “strings in curved spacetimes” it may be worthwhile to point out the differences to their use here:

1. First, the scalar fields  $(\Delta, \psi, \rho, \sigma)$  in Equation (3.56) parameterize a 4D spacetime metric with 2 Killing vectors (not the position of a string in target space) while the target space metric  $\mathfrak{h}$  here (see Equation (3.60) below) has 4 Killing vectors. It is auxiliary and not interpreted as a physical spacetime metric. From the viewpoint of “strings in curved spacetime” the system (3.56) (without matter), on the other hand, describes strings moving on a spacetime with 4 Killing vectors and signature  $(+, +, +, -)$ .
2. The aim in the renormalization process here is to preserve the conformal geometry in target space, not conformal invariance on the worldsheet (base space)  $\Sigma$ . To achieve this one needs metric dependent diffeomorphisms in target space which, as mentioned before, neither need to be nor have been considered before in the context of Riemannian sigma-models.
3. As a consequence of Difference 2 the renormalized fields  $\rho$  and  $\sigma$  become scale dependent and their renormalization flow backreacts on the coupling flow (see Equations (3.75, 3.84) below). This aspect is absent if one naively specializes the renormalization theory of a generic Riemannian sigma-model to a target space geometry which is a warped product (see [221]).
4. As will become clear later in the class of warped product sigma-models considered here the Weyl anomaly is overdetermined at the fixed point of the coupling flow. In contrast to a generic Riemannian sigma-model one is therefore not free to adjust the renormalized target space metric  $\mathfrak{h}_{ij}(\phi)$  such that the Weyl anomaly vanishes and the system is a conformally invariant 2D field theory.

5. The renormalization flow in Riemannian sigma-models is of the form  $\mu \frac{d}{d\mu} \mathfrak{h}_{ij} = \beta_{ij}(\mathfrak{h})$ , where  $\mathfrak{h}$  is the renormalized generating coupling functional (“target space metric”) with the renormalized quantum fields  $\phi$  inserted. Conceptually the highly nonlinear but local  $\beta_{ij}(\mathfrak{h})$  on the right-hand-side thus is a (very special) composite operator, whose finiteness is guaranteed by the construction (see Section B.3). The fact that this very special composite operator is finite does of course not entail that any other nonlinear composite operator built from  $\phi$  or  $\mathfrak{h}_{ij}(\phi)$  is finite (without introducing additional counterterms). For example  $\phi^2$  or a curvature combination of  $\mathfrak{h}_{ij}(\phi)$  not occurring in  $\beta_{ij}(\mathfrak{h})$  is simply not defined off-hand. This is true no matter how  $\phi \mapsto \mathfrak{h}_{ij}(\phi)$  is chosen, so the folklore that one can restrict attention to functionals  $\mathfrak{h}$  for which the trace or Weyl anomaly vanishes and get a “finite” QFT is incorrect (see [201] for a discussion). Moreover the Weyl anomaly is itself a (very special) composite operator and the condition for its vanishing is not equivalent to a partial differential equation of the same form for any classical metric. By expanding the quantum fields  $\phi$  around a classical background configuration one can convert the condition for a vanishing Weyl anomaly into a condition formulated in terms of a classical metric [220]. However beyond lowest order (that is, beyond the Ricci term) nonlocal terms are generated, and the resulting cumbersome equations are rarely used. As a consequence beyond leading order (beyond Ricci flatness modulo an improvement term) most of the “consistent string backgrounds” (defined by ad-hoc replacing the composite operator  $\mathfrak{h}(\phi)$  by a classical metric in the formula for the Weyl anomaly as a composite operator) are actually not consistent, in the sense that the corresponding metric re-interpreted (ad-hoc) as one with the quantum fields re-inserted does not guarantee the vanishing of the Weyl anomaly in its operator form.
6. Even the Ricci flow equations arising at lowest order have the property that for a generic smooth target space metric the flow is often singular towards the ultraviolet [52]. For generic target spaces the Riemannian sigma-models are therefore unlikely to give rise to genuine (not merely effective) quantum field theories.

The situation changes drastically if one considers Riemannian sigma-models where the target manifold is one the warped products (3.59) below. The Problem 5 is absent on the basis of the following Non-renormalization Lemma, the Problem 6 is evaded because the Ricci-type flow arising at first order is constant [61] while to higher orders the asymptotic safety property to be described strikes:

**Non-renormalization Lemma** [154]:

The field  $\rho$  is nonlinearly renormalized but once it is renormalized arbitrary powers thereof (defined by multiplication pointwise on the base manifold) are automatically finite, without the need of additional counterterms. In terms of the normal product defined in Appendix B.3.  $\llbracket F(\rho) \rrbracket = \mu^{d-2} F(\rho)$  for an arbitrary (analytic) function  $F$ .

Needless to say that the same is *not* true for  $\sigma$  or any other of the quantum fields  $\varphi^i$ . As a consequence of this Non-renormalization Lemma the renormalization flow equations for the generating functional  $h$  (the counterpart of  $\mathfrak{h}$ ) can be consistently interpreted as an equation for a classical field, which we also denote by  $\rho$  since the quantum field can be manipulated as if it was a classical field. The resulting flow equations then take the form of a recursive system of nonlinear partial integro-differential equations, which are studied in Section 3.4.

We now describe the derivation of these results in outline; the full details can be found in [154, 155]. The class of warped product target manifolds relevant for Equation (3.56) is of the form

$$(\mathcal{M}, \mathfrak{h}) = (G/H_{h \times} \mathcal{R}, \mathfrak{m}_{h \times} \mathfrak{r}), \quad (3.59)$$

where  $\mathbf{m}$  is the metric (3.11) on the symmetric space  $G/H$ ,  $h = h(\rho)$  is the ‘warp factor’, and  $\mathcal{R}$  is a flat two-dimensional space with Lorentzian metric given by the lower  $2 \times 2$  block in the metric

$$\mathfrak{h}_{ij} := h(\rho) \left( \begin{array}{c|cc} (\mathbf{m}_{ij})_{1 \leq i, j \leq n} & & 0 \\ \hline 0 & a/\rho^2 & b/\rho \\ & b/\rho & 0 \end{array} \right). \quad (3.60)$$

If  $\zeta_1$  is the scalar curvature of the  $G/H$  normalized as in Equation (3.12) the metric (3.60) has scalar curvature

$$R(\mathfrak{h}) = \frac{\zeta_1 \dim G/H}{h(\rho)} \quad (3.61)$$

so that the warp function parameterizes the inverse curvature radius of the target space.

Here we combined the field vector (3.9) with  $\phi^{n+1} := \rho$ ,  $\phi^{n+2} := \sigma$  to a  $D = 2 + \bar{n} + 2k + 2$ -dimensional vector  $\phi = (\varphi^1, \dots, \varphi^n, \rho, \sigma)$ , and the metric (3.60) refer to this coordinate system. Further  $a, b$  are real parameters kept mainly to illustrate that they drop out in the quantities of interest. The metric is chosen such that for the parameter values  $a = -1$ ,  $b = -1$  the Lagrangian (3.56) can be written in the form

$$L_h(\varphi, \rho, \sigma) = -\frac{1}{2\lambda} \mathfrak{h}_{ij}(\phi) \partial^\mu \phi^i \partial_\mu \phi^j. \quad (3.62)$$

In addition to the Killing vectors associated with  $\mathbf{m}$  the metric (3.60) possesses two conformal Killing vectors  $\mathbf{t}_+ = \rho \partial_\rho - \frac{a}{2b} \partial_\sigma$  and  $\mathbf{d} = -\rho \ln \rho \partial_\rho + (\sigma + \frac{a}{b} \ln \rho) \partial_\sigma$ , which together with  $\mathbf{t}_- = \partial_\sigma$  generate the isometries of  $\mathbb{R}^{1,1}$ , i.e.  $[\mathbf{t}_+, \mathbf{t}_-] = 0$ ,  $[\mathbf{d}, \mathbf{t}_\pm] = \pm \mathbf{t}_\pm$ . Conversely any metric with these conformal isometries can be brought into the above form. Each Killing vector of  $\mathfrak{h}$  of course gives rise to a Noether current; the conformal Killing vectors  $\mathbf{t}_+$  and  $\mathbf{d}$  give rise to currents analogous to those in Equation (3.27). The counterpart of on-shell the relations (3.27) is  $\partial^\mu C_\mu = \rho \partial_\rho \ln h \cdot L$ ,  $\partial^\mu D_\mu = -\ln \rho \rho \partial_\rho \ln h \cdot L$ . Upon quantization  $\lambda$  in Equation (3.62) plays the role of the loop counting parameter. In dimensional regularization ( $\int d^2x \rightarrow \int d^d x$ ) the  $l$ -loop counter terms contain poles of order  $\nu \leq l$  in  $(2-d)$ . We denote the coefficient of the  $\nu$ -th order pole by  $T_{ij}^{(\nu, l)}(\mathfrak{h})$ . In principle the higher order pole terms are determined recursively by the residues  $T_{ij}^{(1, l)}(\mathfrak{h})$  of the first order poles. Taking the consistency of the cancellations for granted one can focus on the residues of the first order poles, which we shall do throughout. One can show [154] that they have the following structure:

$$T_{ij}^{(1, l)}(\mathfrak{h}) = \frac{1}{h(\rho)^{l-1}} \left( \begin{array}{c|cc} \zeta_l (\mathbf{m}_{ij})_{1 \leq i, j \leq n} & & 0 \\ \hline 0 & \frac{n}{\rho^2} S_l(\rho) & 0 \\ & 0 & 0 \end{array} \right), \quad \forall l \geq 1. \quad (3.63)$$

It should be stressed that this is not trivially a consequence of the block-diagonal form of Equation (3.60), rather the properties (3.12) enter in an essential way.

The  $\zeta_l$  are constants defined through the curvature scalars of  $\mathbf{m}_{ij}$ . The  $S_l(\rho)$  are differential polynomials in  $h$  invariant under constant rescalings of  $h$  and normalized to vanish for constant  $h$ .

The first three are

$$\begin{aligned} S_1(h) &= -\frac{1}{2}(\rho\partial_\rho)^2 \ln h + \frac{1}{4}(\rho\partial_\rho \ln h)^2, \\ S_2(h) &= 0, \\ S_3(h) &= -\frac{\zeta_2}{4}(\rho\partial_\rho)^2 \ln h + \frac{\zeta_2}{12}(\rho\partial_\rho \ln h)^2. \end{aligned} \quad (3.64)$$

The counter terms (3.63) ought to be absorbed by nonlinear field renormalizations,

$$\phi_B^j = \phi^j + \frac{1}{2-d}\Xi^j(\phi, \lambda) + \dots, \quad \text{with } \Xi^j = \sum_{l \geq 1} \left(\frac{\lambda}{2\pi}\right)^l \phi_l^j(\phi), \quad (3.65)$$

and a renormalization of the function  $h$ ,

$$h_B(\rho) = \mu^{d-2}h(\rho, \lambda) \left[ 1 + \frac{1}{2-d}H(\rho, \lambda) + \dots \right], \quad H(\rho, \lambda) = \sum_{l \geq 1} \left(\frac{\lambda}{2\pi}\right)^l H_l(\rho), \quad (3.66)$$

where  $\mu$  is the renormalization scale. Note that on *both* sides of Equation (3.66) the argument is the renormalized field. The renormalized  $h$  function is allowed to depend on  $\lambda$ ; specifically we assume it to have the form

$$h(\rho, \lambda) = \rho^p + \frac{\lambda}{2\pi}h_1(\rho) + \left(\frac{\lambda}{2\pi}\right)^2 h_2(\rho) + \dots, \quad (3.67)$$

where the first term ensures standard renormalizability at the 1-loop level – and is determined by this requirement except for the power  $p \neq 0$ . The power has no intrinsic significance; one could have chosen a parameterization of the 4D spacetime metric  $g_{\alpha\beta}$  such that the action (3.10) with  $\rho$  replaced by  $\rho^p$  was the outcome of the classical reduction procedure. In particular the sectors  $p > 0$  and  $p < 0$  are equivalent and we assume  $p > 0$  throughout.

Combining Equations (3.60, 3.65, 3.66) and (3.63) one finds that the first order poles cancel in the renormalized Lagrangian iff the following “reduction condition” holds:

$$\mathcal{L}_\Xi \mathfrak{h}_{ij} + H(\rho, \lambda)\mathfrak{h}_{ij} = \lambda T_{ij}^{(1)}(\mathfrak{h}/\lambda), \quad (3.68)$$

where  $T^{(1)}(\mathfrak{h}) = \sum_{l \geq 1} \left(\frac{1}{2\pi}\right)^l T^{(1,l)}(\mathfrak{h})$  and  $\mathcal{L}_\Xi \mathfrak{h}$  is the Lie derivative of  $\mathfrak{h}$ ,  $\mathcal{L}_\Xi \mathfrak{h}_{ij} := (\mathcal{L}_\Xi \mathfrak{h})_{ij} = \Xi^k \partial_k \mathfrak{h}_{ij} + \partial_i \Xi^k \mathfrak{h}_{kj} + \partial_j \Xi^k \mathfrak{h}_{ki}$ . The  $\rho$ -dependence of  $H$  marks the deviation from conventional renormalizability. Guided by the structure of Equations (3.60) and (3.63) we search for a solution with  $\Xi^j = (0, \dots, 0, \Xi^\rho(\rho, \lambda), \Xi^\sigma(\rho, \lambda))$ , where here and later on we also use  $\rho = n+1$ ,  $\sigma = n+2$  for the index labeling. The Lie derivative term with this  $\Xi^j$  is

$$\begin{aligned} \mathcal{L}_\Xi \mathfrak{h}_{ij} &= \left( \begin{array}{c|cc} \Xi^\rho(\rho)\partial_\rho h(\mathfrak{m}_{ij})_{1 \leq i, j \leq n} & & 0 \\ \hline 0 & \mathcal{L}_\Xi \mathfrak{h}_{\rho\rho} & \mathcal{L}_\Xi \mathfrak{h}_{\rho\sigma} \\ & \mathcal{L}_\Xi \mathfrak{h}_{\rho\sigma} & 0 \end{array} \right), \\ \mathcal{L}_\Xi \mathfrak{h}_{\rho\rho} &= \frac{a}{\rho^2} \left[ \partial_\rho h \Xi^\rho + 2h\rho \partial_\rho \left(\frac{\Xi^\rho}{\rho}\right) \right] + 2b \frac{h}{\rho} \partial_\rho \Xi^\sigma, \\ \mathcal{L}_\Xi \mathfrak{h}_{\rho\sigma} &= b \partial_\rho \left(\frac{h \Xi^\rho}{\rho}\right). \end{aligned} \quad (3.69)$$

The reduction condition (3.68) then is equivalent to a simple system of differential equations whose solution is

$$\begin{aligned} H(h/\lambda) &= -\frac{1}{h(\rho, \lambda)} \rho \partial_\rho \left[ h(\rho, \lambda) \frac{\Xi^\rho(h/\lambda)}{\rho} \right], \\ \Xi^\rho(h/\lambda) &= -\rho \int^\rho \frac{du}{u} B_\lambda \left( \frac{\lambda}{h(u, \lambda)} \right), \\ \Xi^\sigma(h/\lambda) &= -\frac{a}{2b\rho} \Xi^\rho(h/\lambda) + \frac{1}{2b} \int^\rho \frac{du}{u} S(u, \lambda). \end{aligned} \quad (3.70)$$

Here we set

$$B_\lambda(\lambda) := \sum_{l \geq 1} \zeta_l \left( \frac{\lambda}{2\pi} \right)^l, \quad S(\rho, \lambda) := n \sum_{l \geq 1} \left( \frac{\lambda}{2\pi} \right)^l h^{-l} S_l(\rho), \quad (3.71)$$

and slightly adjusted the notation to stress the functional dependence on  $h/\lambda$ . Possibly  $\lambda$ -dependent integration constants have been absorbed into the lower integration boundaries of the integrals. Throughout these solutions should be read as shorthands for their series expansions in  $\lambda$  with  $h$  of the form (3.67). For example

$$\Xi^\rho(\rho, \lambda) = \frac{\lambda}{2\pi} \frac{\zeta_1}{p} \rho^{-p+1} + \left( \frac{\lambda}{2\pi} \right)^2 \rho \int_\rho^\infty \frac{du}{u^{2p+1}} [\zeta_2 - \zeta_1 h_1(u)] + O(\lambda^3). \quad (3.72)$$

For the derivation of Equations (3.68, 3.70) we fixed a coordinate system in which the target space metric takes the form (3.60). Under a change of parameterization  $\phi^j \rightarrow \phi^j(\hat{\phi})$  the reduction condition (3.68) should transform covariantly, and indeed it does. The constituents transform as

$$\begin{aligned} \hat{\mathfrak{h}}_{ij}(\hat{\phi}) &= \frac{\partial \phi^k}{\partial \hat{\phi}^i} \frac{\partial \phi^l}{\partial \hat{\phi}^j} \mathfrak{h}_{kl}(\phi), & T_{ij}^{(1)}(\hat{\mathfrak{h}})(\hat{\phi}) &= \frac{\partial \phi^k}{\partial \hat{\phi}^i} \frac{\partial \phi^m}{\partial \hat{\phi}^j} T_{km}^{(1)}(\mathfrak{h})(\phi), \\ \hat{\Xi}^j(\hat{\phi}) &= \frac{\partial \hat{\phi}^j}{\partial \phi^k} \Xi^k(\phi), & (\mathcal{L}_{\hat{\Xi}} \hat{\mathfrak{h}})_{ij}(\hat{\phi}) &= \frac{\partial \phi^k}{\partial \hat{\phi}^i} \frac{\partial \phi^m}{\partial \hat{\phi}^j} (\mathcal{L}_{\Xi} \mathfrak{h})_{km}(\phi). \end{aligned} \quad (3.73)$$

The covariance of the counter terms as a function of the full field is nontrivial [110, 41] and is one of the main advantages of the covariant background field expansion. The relations (3.73) can be used to convert the solutions (3.70) of the finiteness condition into any desired coordinate system on the target space. The coordinates  $\sigma$  and  $\rho$  used in Equation (3.60) are adapted to the Killing vector  $\mathbf{t}_-$  and the conformal Killing vectors  $\mathbf{t}_+$ ,  $\mathbf{d}$ .

This completes the renormalization of the Lagrangian  $L_h$ . The nonlinear field redefinitions alluded to in Equation (3.57) are explicitly given by Equation (3.70). The function  $\rho \mapsto h(\rho)$  plays the role of a generating function of an infinite set of essential couplings. In principle it could be expanded with respect to a basis of  $\mu$ -independent functions of  $\rho$  with  $\mu$ -dependent coefficients, the couplings. Technically the fact that these couplings are essential (in sense defined in the introduction) follows from Equations (3.27). Since  $\rho$  is a nontrivial function on the base manifold, the Lagrangian is a total divergence on shell if and only if  $h(\rho) = \rho^p$ , or  $h(\rho) = \ln \rho^p$ . The first case corresponds to the classical Lagrangian (3.16), the second case was studied (in a different context) by Tseytlin [221]. In the case  $h(\rho) = \rho^p$  the identity  $pL = \partial^\mu C_\mu$  reflects the fact that the overall scale of the metric is an inessential parameter (see Appendix A). The renormalization flow associated with the coupling functional  $h$  will be studied in the next Section 3.4.

The fields themselves, here to be viewed as a collection of inessential parameters, are likewise subject to flow equations. Recall from Equation (3.65) the relation between the bare and the renormalized fields, where  $\Xi^1 = \dots = \Xi^n = 0$ , while  $\Xi^\rho, \Xi^\sigma$  have been computed in Equation (3.70)

and depend on  $h$ . Since the bare fields are  $\mu$ -independent, the renormalized fields  $\phi^j$  have to carry an implicit  $\mu$ -dependence through  $h$ . (This is analogous to the situation in an ordinary multiplicatively renormalizable quantum field theory, where the coupling dependence of the wave function renormalization induces a compensating  $\mu$ -dependence of the renormalized fields governed by the anomalous dimension function.) The flow equations involve functional derivatives with respect to the  $h$  field. For any functional  $X(\rho) = X[h](\rho)$  of  $h(\rho)$  we set

$$\dot{X}(\rho) := \int du h(u) \frac{\delta X(\rho)}{\delta h(u)}. \quad (3.74)$$

Observe that for any differential or integral polynomial  $X_l$  in  $h$  which is homogeneous of degree  $l$ , the functional derivative (3.74) just measures the degree,  $\dot{X}_l = lX_l$ . From Equations (3.65, 3.70) and the  $\bar{h}$ -flow (3.83) below one derives

$$\mu \frac{d}{d\mu} \bar{\rho} = -\dot{\Xi}^\rho[\bar{h}/\lambda](\bar{\rho}), \quad \mu \frac{d}{d\mu} \bar{\sigma} = -\dot{\Xi}^\sigma[\bar{h}/\lambda](\bar{\rho}), \quad (3.75)$$

where  $\dot{\Xi}^\rho[\bar{h}]$ ,  $\dot{\Xi}^\sigma[\bar{h}]$  refer to Equation (3.76) with the solution of Equation (3.83) inserted for  $h$ . Note that, conceptually, the problems decouple: One first solves the autonomous equation (3.83) to obtain the coupling flow  $\mu \rightarrow \bar{h}(\cdot, \mu)$  which is then used to specify the right-hand-side of the  $\bar{\rho}$ -flow equation whose solution in turn determines the  $\bar{\sigma}$ -flow. The ‘ $\cdot$ ’ derivatives of the solution (3.70) of the reduction condition come out as

$$\begin{aligned} \dot{H} &= -\frac{1}{h} \rho \partial_\rho \left[ \frac{h}{\rho} \dot{\Xi}^\rho \right], \\ \dot{\Xi}^\rho &= \rho \int_\rho^\infty \frac{du}{u} \frac{h(u)}{\lambda} \beta_{G/H} \left( \frac{\lambda}{h(u)} \right), \\ \dot{\Xi}^\sigma &= -\frac{a}{2b\rho} \dot{\Xi}^\rho + \frac{1}{2b} \int_\rho^\infty \frac{du}{u} \dot{S}(u, \lambda). \end{aligned} \quad (3.76)$$

In  $\dot{\Xi}^\rho$  we set  $\beta_{G/H}(\lambda) := -\lambda \sum_{l \geq 1} l \zeta_l \left( \frac{\lambda}{2\pi} \right)^l$  and anticipated in the notation that this is the conventional beta functions of a  $G/H$  coset sigma-model without coupling to gravity. In  $\dot{\Xi}^\sigma$  we absorbed a  $\lambda$ -dependent additive constant into the lower integration boundary and used  $\dot{S}(\rho, \lambda) = -\sum_{l \geq 1} \left( \frac{\lambda}{2\pi} \right)^l l h^{-l} S_l(\rho)$ , as  $\dot{S}_l(\rho) = 0$ .

We proceed with the renormalization of composite operators. Again we borrow techniques from Riemannian sigma models (see Appendix B.3). The normal product of scalar, vector, and tensor operators on the target manifold is defined in Equation (B.70). For generic composite operators of course the bare operator viewed as a function of the bare couplings and of the bare fields will have a different functional form from the renormalized one viewed as a function of the renormalized couplings and fields. An important exception was already described in the above ‘Non-renormalization Lemma’ for functions of  $\rho$  only. This is specific to the system here. Another class of operators for which similar non-renormalization results hold are conserved Noether currents; this is a feature true in general. In the case at hand the relevant Noether currents are Equation (3.26), the current  $R_\mu = \frac{b}{\lambda} h \partial_\mu \ln \rho$ , and the ‘would-be’ energy momentum tensor  $T_{\mu\nu}$ . In terms of the normal product (B.70) the corresponding non-renormalization results read [154]

$$\llbracket J_\mu(h; \phi) \rrbracket = J_\mu(h_B; \phi_B). \quad (3.77)$$

For the current  $R_\mu$  the identity  $\llbracket R_\mu(h; \rho) \rrbracket = \mu^{d-2} R_\mu(h_B; \rho_B)$  follows similarly on general grounds, while the stronger identity

$$\llbracket R_\mu(h; \rho) \rrbracket = \mu^{d-2} R_\mu(h; \rho) = \mu^{d-2} \frac{b}{\lambda} \partial_\mu \int_\rho^\infty \frac{du}{u} h(u, \lambda) \quad (3.78)$$



is a consequence of the non-renormalization Lemma. The result (3.78) will later turn out to reflect a property of the generalized beta function. For the renormalization of constraints in Equation (3.15) improvement terms are crucial. As in Equation (3.15) we wish to identify the constraints  $\mathcal{H}_0$  and  $\mathcal{H}_1$  with the components of the “would be” energy momentum tensor associated with the ‘deformed’ Lagrangian  $L_h$ . To this end we decompose the energy momentum tensor for the Lagrangian  $L_h$  into a symmetric tracefree part  $t_{\mu\nu}$  and an improvement term

$$T_{\mu\nu}(h; \phi) = t_{\mu\nu}(h; \phi) + \Delta_{\mu\nu}f(\rho), \quad f(\rho) = 2b \int^{\rho} \frac{du}{u} h(u), \quad (3.79)$$

with  $\Delta_{\mu\nu} = \partial_{\mu}\partial_{\nu} - \eta_{\mu\nu}\partial^2$ . The improvement term  $\Delta_{\mu\nu}f$  is trivially conserved but its trace  $\eta^{\mu\nu}\Delta_{\mu\nu}f = -\partial^2f$  vanishes only on-shell. In contrast to  $t_{\mu\nu}$  its functional form is not protected by the conservation equation, and for the finiteness of the composite operator  $\llbracket T_{\mu\nu}(h; \phi) \rrbracket$  the improvement potential  $f(\rho)$  has to be renormalized in a way that changes its functional form. This is to say, there is no function  $f(\cdot, \lambda)$  such that the bare and the renormalized improvement potential would merely be related by substituting the bare field  $\rho_B$  and the renormalized one  $\rho$ , respectively, into  $f(\cdot, \lambda)$ . Rather we set  $V_B = f_B(\rho_B) + \mu^{d-2}f_0\sigma_B$  and  $V = f(\rho) + f_0\sigma$ , where  $f_0$  is a potentially  $\lambda$ -dependent constant,  $f_B$  is a function of the bare field  $\rho_B$  and  $\lambda$ , and  $f$  is a function of the renormalized field  $\rho$  and  $\lambda$ . The finiteness of  $\llbracket T_{\mu\nu}(h; \phi) \rrbracket$  can then be achieved by relating  $f_B$  and  $f$  (the functions, not their values) according to

$$f_B(\rho) = \mu^{d-2}f(\rho) + \frac{\mu^{d-2}}{2-d} [k(\rho) - \partial_{\rho}f(\rho)\Xi^{\rho}(\rho) - f_0\Xi^{\sigma}(\rho)] + \dots \quad (3.80)$$

Here  $k(\rho)$  is a differential polynomial in  $h(\rho)$  that can be computed from the counterterms in Equation (B.61) of Appendix B.3. Note that as in Equation (3.66) the argument on *both* sides is the renormalized  $\rho$  field. Starting from the fact that the right-hand-side of Equation (3.80) is  $\mu$ -independent one can derive a non-autonomous flow equation [155]

$$\begin{aligned} \mu \frac{d}{d\mu} \bar{f} &= F_f[\bar{h}] \bar{f}, \\ F_f[h] &= \dot{\Xi}^{\rho}[h]\partial_{\rho} + \dot{\Xi}^{\sigma}[h]f_0 - K[h], \end{aligned} \quad (3.81)$$

where  $K[h]$  is determined by  $k[h]$ . For a given solution  $\bar{h}$  the flow equation (3.81) in principle determines  $\bar{f} = \bar{f}[\bar{h}](\rho, \mu)$ . Finally

$$\llbracket T_{\mu\nu}(h; \phi) \rrbracket := \llbracket t_{\mu\nu}(h; \phi) \rrbracket + \Delta_{\mu\nu} \llbracket V \rrbracket = t_{\mu\nu}(h_B, \phi_B) + \Delta_{\mu\nu}V_B \quad (3.82)$$

can be shown to be a finite composite operator. Upon specification of initial data for  $h(\rho) = \bar{h}(\rho, \mu_0)$  and  $f(\rho) = \bar{f}(\rho, \mu_0)$  satisfying the proper boundary or fall-off conditions in  $\rho$ , the composite operator  $\llbracket T_{\mu\nu}(h; \phi) \rrbracket$  is completely determined. This holds for an arbitrary coupling function  $h$ . In general the operator  $\llbracket T_{\mu\nu}(h; \phi) \rrbracket$  will not be trace free. The interpretation of the components of  $\llbracket T_{\mu\nu} \rrbracket$  as quantum constraints, on the other hand, requires that the trace vanishes as  $\llbracket \mathcal{H}_0 \rrbracket$  should be equal to both  $\llbracket T_{00} \rrbracket$  and  $\llbracket T_{11} \rrbracket$ . We shall return to this condition below (see Equations (3.100), etc.).

### 3.4 Non-Gaussian fixed point and asymptotic safety

As explained before the function  $h$  plays the role of a generating functional for an infinite set of essential couplings. As such it is subject to a flow equation

$$\mu \frac{d}{d\mu} \bar{h} = \lambda \beta_h(\bar{h}/\lambda), \quad (3.83)$$

where  $\mu$  is the renormalization scale and  $\mu \rightarrow \bar{h}(\cdot, \mu)$  is the ‘running’ coupling function. The flow equation can be obtained by the usual procedure starting from the fact that the left-hand-side of Equation (3.58) is  $\mu$ -independent. One finds [154]

$$\begin{aligned} \lambda \beta_h(h/\lambda) &= (2-d)h(\rho) - h(\rho) \int du h(u) \frac{\delta H(\rho)}{\delta h(u)} \\ &= (2-d)h(\rho) - \frac{\hbar^2}{\lambda} \beta_{G/H} \left( \frac{\lambda}{\hbar} \right) + \dot{\Xi}^\rho \left( \frac{\hbar}{\lambda} \right) \partial_\rho h, \end{aligned} \quad (3.84)$$

where we suppress the  $\lambda$  dependence of  $h$ . Inserting Equation (3.69) and setting  $d = 2$  in the result gives

$$\beta_h(h/\lambda) = \rho \partial_\rho \left[ \frac{h(\rho)}{\lambda} \int_\rho^\infty \frac{du}{u} \frac{h(u)}{\lambda} \beta_{G/H} \left( \frac{\lambda}{h(u)} \right) \right]. \quad (3.85)$$

Here  $\beta_{G/H}(\lambda) = -\lambda \sum_{l \geq 1} l \zeta_l \left( \frac{\lambda}{2\pi} \right)^l$  is again the conventional (numerical) beta function of the  $G/H$  nonlinear sigma-model *without* coupling to gravity, computed in the minimal subtraction scheme.  $\beta_h(h)$  can thus be regarded as a ‘gravitationally dressed’ version of  $\beta_{G/H}(\lambda)$ . The term is borrowed from [120] where a similar phenomenon was found in a different context and to lowest order. In contrast to Equation (3.85) the effect of quantum gravity on the running of the couplings in these Liouville-type theories cannot be represented as a simple ‘dressing relation’ beyond lowest order [168]. The flow equation resulting from Equation (3.85) will be studied in more detail below. We anticipate however that the appropriate boundary conditions are such that the solution  $\bar{h}(\rho, \lambda)$  is stationary (constant in  $\mu$ ) for  $\rho \rightarrow \infty$ . This guarantees that the renormalization flow is exclusively driven by the counterterms, as it should. We add some comments on the structure of the beta function (3.85).

An initially puzzling feature of  $\beta_h(h)$  is that it comes out as a total  $\ln \rho$ -derivative. Restoring the interpretation of  $\rho = \rho(x)$  as a field on the 2D base space, however, it has a natural interpretation: An immediate consequence of Equations (3.83, 3.85) is that contour integrals of the form

$$\int_C dx^\mu \partial_\mu \ln \rho \bar{h}(\rho, \mu) \quad (3.86)$$

are  $\mu$ -independent for any closed contour  $C$  in the base space, as can be seen by differentiating Equation (3.86) with respect to  $\mu$ . They are thus invariants of the flow and can be used to discriminate the inequivalent quantum theories parameterized by the  $h_i(\rho)$ . With the initial condition  $\bar{h}(\rho, \mu_0) = h(\rho)$  the  $\mu$ -independence of Equation (3.86) is equivalent to  $\partial^\mu [(\bar{h} - h) \partial_\mu \ln \rho] = 0$ . On the other hand the (classical and quantum) equations of motion for  $\rho$  with respect to the  $h$ -modified action are just  $\partial^\mu (h \partial_\mu \ln \rho) = 0$ . Combining both we find that the significance of  $\beta_h(h)$  being a total  $\ln \rho$ -derivative is that this feature preserves the equations of motion for  $\rho$  under the  $\mu$ -evolution of  $\bar{h}(\cdot, \mu)$ :

$$\int_\rho^\infty \frac{du}{u} h(u) \text{ harmonic} \quad \implies \quad \int_\rho^\infty \frac{du}{u} \bar{h}(u, \mu) \text{ harmonic}. \quad (3.87)$$

This provides an important consistency check as Equation (3.87) is also required by the non-renormalization of the  $R_\mu$  Noether current in Equation (3.78).

Another intriguing property of Equation (3.85) can be seen from the second line in Equation (3.84). The first two terms correspond to the beta function of a  $G/H$  sigma-model without coupling to 2D gravity (i.e. with nondynamical  $\rho$  and  $\sigma$ ). The last term is crucial for all the subsequent properties of the flow (3.83). Comparing with Equation (3.75) one sees that it describes a *backreaction* of the scale dependent area radius  $\rho$  on the coupling flow, which is mediated by the quantum dynamics of the other fields. We shall return to this point below.

As usual fixed points of the flow correspond to zeros of the generalized beta function. The flow has two fixed points, a degenerate Gaussian one corresponding to  $h = \infty$  formally, and a non-Gaussian fixed point  $h^{\text{beta}}(\cdot)$  which is of main interest here. We postpone the discussion of the Gaussian fixed point and focus on the non-Gaussian one here. The defining relation for such a nontrivial fixed point amounts by Equation (3.85) to the differential equation

$$\frac{\lambda}{2\pi} \rho \partial_\rho h = -C(\lambda) h^2 \frac{h}{\lambda} \beta_{G/H} \left( \frac{\lambda}{h} \right), \tag{3.88}$$

where  $C(\lambda)$  is a constant, which in principle could be  $\lambda$ -dependent. The Equation (3.88) is to be interpreted with Equation (3.67) inserted and expanded in powers of  $\lambda$ . This can be solved recursively for  $h_1, h_2$ , etc. We denote the solutions by  $h_i^{\text{beta}}(\rho)$ . The unique solution adhering to the above boundary condition corresponds to a  $\lambda$ -independent  $C$ , for which we write  $p/\zeta_1$ , where  $\zeta_1$  is the first beta function coefficient. The solution then reads

$$h^{\text{beta}}(\rho, \lambda) = \rho^p - \frac{\lambda}{2\pi} \frac{2\zeta_2}{\zeta_1} - \left( \frac{\lambda}{2\pi} \right)^2 \frac{3\zeta_3}{2\zeta_1} \rho^{-p} + \dots \tag{3.89}$$

This is a nontrivial fixed point, in the sense that gravity remains self-interacting and coupled to matter.

The renormalized action (at or away from the fixed point) has no direct significance in the quantum theory. It is however instructive to note that  $L_h$  in Equation (3.56) can for any  $h(\rho) \neq \rho^p$  not be written in the form  $L_{\text{EH}}(g) = \sqrt{-g}R(g)$  for any metric with two commuting Killing vectors. In accordance with the general renormalization group picture one can of course write  $L_h \sim L_{\text{EH}}(g) + L_{\text{other}}(g)$ . Interestingly, in the 2-Killing vector reduction one can write the renormalized Lagrangian  $L_h$  as a sum of two terms which are reductions of  $L_{\text{EH}}$ . Since the parameter  $a$  in  $L_h$  can be changed at will by redefining  $\sigma \mapsto \sigma + \frac{a-\bar{a}}{2b} \rho^{-1}$ , we may assume that  $b = -p, a = -3p^2 + 2p$ , for some  $p \neq 0$ . Then

$$\lambda L_h = L_{\text{EH}} \left( \frac{h(\rho)}{\rho} g \right) + L_{\text{EH}}(g_0), \tag{3.90}$$

modulo total derivatives, where in the first term  $g$  is the metric with line element (3.8) and  $g_0$  is a metric with line element

$$\begin{aligned} dS_0^2 &= e^{-\sigma - k(\rho)} [-(dx^0)^2 + (dx^1)^2] + \left( 1 - \frac{p}{h} \int^\rho \frac{du}{u} h(u) \right)^2 (dy^1)^2 - h(\rho)^2 (dy^2)^2, \\ \partial_\rho k &= \partial_\rho \left[ \frac{3}{2} \ln h + \left( \frac{3}{2} p - 1 \right) \ln \rho \right] - \frac{2\rho \partial_\rho h}{\rho \partial_\rho h - ph} \partial_\rho \left( \frac{p}{h} \int^\rho \frac{du}{u} h(u) \right). \end{aligned} \tag{3.91}$$

The extra term  $L_{\text{EH}}(g_0) = \frac{1}{2}(\rho \partial_\rho h - ph)[(3p - 2 + 3\rho \partial_\rho \ln h)\rho^{-2} \partial^\mu \rho \partial_\mu \rho + 2\rho^{-1} \partial^\mu \rho \partial_\mu \sigma]$  involves only the area radius  $\rho$  and the conformal factor  $\sigma$ . It vanishes iff  $h(\rho)$  is proportional to  $\rho^p$ . From Equation (3.89) one sees that  $h^{\text{beta}}(\rho) \neq \rho^p$  always; by construction in the minimal subtraction scheme, but since the coefficients  $\zeta_1, \zeta_2$  are universal, the same holds in any other scheme.

The origin of this feature is the seeming violation of scale invariance on the level of the renormalized action. Recall from after Equation (3.62) that  $\partial^\mu C_\mu = \rho \partial_\rho \ln h \cdot L_h$ , so that for  $h^{\text{beta}}(\rho) \neq \rho^p$  the action is no longer scale invariant. However this is precisely the property which allows one to cancel the (otherwise) anomalous term in the trace of the would-be energy momentum tensor, as discussed before, rendering the system conformally invariant at the fixed point. Due to the lack of naive scale invariance on the level of the renormalized action the dynamics of quantum gravodynamics is different from that of quantum general relativity, in the sector considered, even at the fixed point. The moral presumably generalizes: The form of the (bare and/or renormalized) action

may have to differ from the Einstein–Hilbert action in order to incorporate the physics properties aimed at.

In the present context the most important feature of the fixed point  $h^{\text{beta}}$  is its ultraviolet stability: For  $\mu \rightarrow \infty$  all linearized perturbations around the fixed point function (3.89) are driven back to the fixed point. Since the fixed point (3.89) has the form of a power series in the loop counting parameter  $\lambda$ , the proper concept of a “linearized perturbation” has the form

$$\begin{aligned}\bar{h}(\rho, \lambda, \mu) &= h^{\text{beta}}(\rho, \lambda) + \delta\bar{h}(\rho, \lambda, \mu), \\ \delta\bar{h}(\rho, \lambda, \mu) &= \frac{\lambda}{2\pi} \bar{s}_1(\varrho, t) + \left(\frac{\lambda}{2\pi}\right)^2 \bar{s}_2(\varrho, t) + \dots\end{aligned}\tag{3.92}$$

where the  $\bar{s}_l(\varrho, t)$  are functions of  $\varrho := \rho^p$  and  $t = \frac{1}{2\pi} \ln \mu/\mu_0$ . Note that the perturbation involves infinitely many functions of two variables. The boundary condition mentioned before, which guarantees that the full  $\bar{h}$  flow is driven by the counterterms, only amounts to the requirement that all the  $\bar{s}_l(\varrho, t)$  vanish for  $\varrho \rightarrow \infty$  uniformly in  $t$ . Inserting the ansatz (3.92) into the flow equation  $\mu \frac{d}{d\mu} \bar{h} = \beta_h(\bar{h}/\lambda)$  and linearizing in  $\delta\bar{h}(\rho, \lambda, \mu)$  gives a recursive system of inhomogeneous integro-differential equations for the  $\bar{s}_l$ ,

$$\frac{d}{dt} \bar{s}_l = \zeta_1 \varrho \int_{\varrho}^{\infty} \frac{du}{u^3} \bar{s}_l(u, t) - \zeta_1 \partial_{\varrho} \bar{s}_l + R_l[\bar{s}_{l-1}, \dots, \bar{s}_1], \quad l \geq 1.\tag{3.93}$$

Here  $R_1 = 0$ , while the  $R_l$ ,  $l \geq 2$ , are complicated integro-differential operators acting (linearly) on the  $\bar{s}_1, \dots, \bar{s}_{l-1}$  (see [155] for the explicit expressions). The lowest order equation (3.93) is homogeneous and its solution is given by

$$\bar{s}_1(\varrho, t) = \varrho \int_{\varrho}^{\infty} \frac{du}{u} r_1(u - \zeta_1 t),\tag{3.94}$$

where  $r_1$  is an arbitrary smooth function of one variable satisfying  $u r_1(u) \rightarrow 0$  for  $u \rightarrow 0$ . This function can essentially be identified with the initial datum at some renormalization time  $t = 0$ , as  $r_1(\varrho) = -\varrho \partial_{\varrho} [\bar{s}_1(\varrho, t = 0)/\varrho]$ . Evidently  $\bar{s}_1(\varrho, t) \rightarrow 0$  for  $t \rightarrow \infty$ , if  $\zeta_1 < 0$ . This condition is indeed satisfied by all the systems (3.1, 3.56) considered, precisely because the coset space  $G/H$  is noncompact. Interestingly one has the simple formula [155]

$$\zeta_1 = -\frac{k+2}{2}, \quad k = \# \text{ Abelian vector fields} = \frac{1}{4}(\dim G - \dim \bar{G} - 3).\tag{3.95}$$

It follows from the value of the quadratic Casimir in the appropriate representation and is consistent with [47]. Since  $\zeta_1 < 0$  always, Equation (3.94) shows that the lowest order perturbation  $\bar{s}_1$  will always die out for  $t \rightarrow \infty$ , for arbitrary smooth initial data prescribed at  $t = 0$ . It can be shown that this continues to hold for all higher order  $\bar{s}_l$  irrespective of the signs of the coefficients  $\zeta_l$ ,  $l \geq 2$ .

**Result** (UV stability):

Given smooth initial  $\bar{s}_l(\varrho, 0)$  with  $\bar{s}_l(\infty, t) = 0$ ,  $l \geq 1$ , the solution of the linearized flow equations (3.93) is unique and satisfies

$$\bar{s}_l(\varrho, t) \longrightarrow 0 \quad \text{for } t \longrightarrow \infty \quad \text{if } \zeta_1 < 0,$$

where the convergence is uniform in  $\varrho$ .

The situation is illustrated in Figure 1. The proof of this result is somewhat technical and can be found in [155].

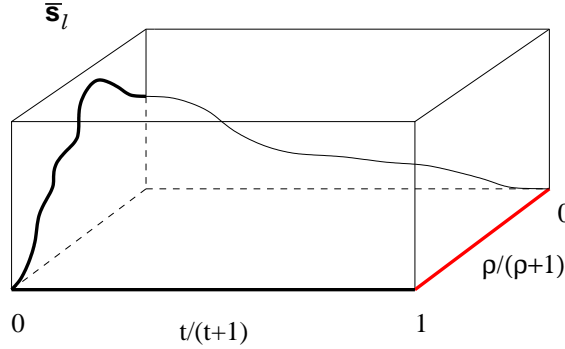


Figure 1: Schematic form of the linearized flow of the  $\bar{s}_l$ ,  $l \geq 1$ , perturbations;  $t = \frac{1}{2\pi} \ln \mu/\mu_0$ ,  $\varrho = \rho^p$ .

Often the stability properties of fixed points are not discussed by solving the linear flow equations directly, but by studying the spectral properties of the linearized perturbation operator (the “stability matrix” in Equation (A.10)). Since the generalized couplings here are functions, the linearized perturbation operator  $\Theta$  is a formal integral operator,

$$(\Theta s)(\varrho) = \int du \frac{\delta \beta_h(h/\lambda)(\varrho)}{\delta h(u)} \Big|_{h=h^{\text{beta}}} s(u), \tag{3.96}$$

with a distributional kernel  $\Theta(\varrho, u) = \sum_l (\frac{\lambda}{2\pi})^{l-1} \Theta_l(\varrho, u)$ , which can be computed from the explicit formula (3.85). For example

$$\Theta_1(\varrho, u) = \zeta_1 [\partial_u \delta(u - \varrho) + \varrho u^{-3} \theta(u - \varrho)], \tag{3.97}$$

where  $\theta(u)$  is the step function. Writing  $s(\varrho) = \sum_{l \geq 1} (\frac{\lambda}{2\pi})^{l-1} s_l(\varrho)$ ,  $\vartheta = \sum_{l \geq 1} (\frac{\lambda}{2\pi})^{l-1} \vartheta_l$ ,  $\vartheta \in \mathbb{C}$ , the spectral problem  $\Theta s = \vartheta s$  decomposes into a sequence of integro-differential equations for the  $\Theta_1 s_l - \vartheta_1 s_l = \text{rhs}$ , where the right-hand-side is determined by the solution of the lower order equations. Only the  $l = 1$  equation is a spectral problem proper,  $\Theta_1 s_1 = \vartheta_1 s_1$ ,  $\vartheta_1 \in \mathbb{C}$ . The relevant and irrelevant perturbations have spectra with negative and positive real parts, respectively. Remarkably *all* (nontrivial) eigenfunctions of  $\Theta_1$  are normalizable; the spectrum is “purely discrete” and consists of the entire halfplane  $\{\vartheta_1 \in \mathbb{C} \mid \text{Re } \vartheta_1 < 0\}$ . Indeed, the general solution to  $\Theta_1 s_1 = \vartheta_1 s_1$  is  $s_1(\varrho) = a\varrho + b\varrho \int_{\varrho}^{\infty} \frac{du}{u} e^{-\vartheta_1 u/\zeta_1}$ , with  $a, b \in \mathbb{C}$ . The first term merely corresponds to a change of normalization of  $h^{\text{beta}}(\varrho) = \varrho + O(\lambda)$  and we may set  $a = 0$ ,  $b = 1$  without loss of generality. The second term corresponds to Equation (3.94) with  $r_1(u) = e^{-\vartheta_1 u/\zeta_1}$ . This clearly confirms the above result from a different perspective. For  $l > 1$  the parameters  $\vartheta_l$  are not spectral values for  $\Theta_l$ . Moreover, since the kernels  $\Theta_l$  are distributions it is not quite clear which precise functional analytic setting one should choose for the full spectral problem  $\Theta s = \vartheta s$ . This is why above we adopted the direct strategy and determined the solutions of the linearized flow equations. Their asymptotic decay shows the ultraviolet stability of the fixed point unambiguously and independent of functional analytical subtleties.

We can put this result into the context the general discussion in Section 2 and arrive at the following conclusion:

**Conclusion:**

With respect to the non-Gaussian fixed point  $h^{\text{beta}}(\cdot)$  all couplings in the generating functional  $h(\cdot)$  are asymptotically safe. All symmetry reduced gravity systems satisfy the Criteria (PTC1)

and (PTC2) to all loop orders of sigma-model perturbation theory. As explained in Section 3.2 from the viewpoint of the graviton loop expansion the distinction between a perturbative and a non-perturbative treatment is blurred here.

It is instructive to compare these properties to that of the Gaussian fixed point. The Gaussian fixed point of the flow (3.83) is best understood in analogy to the Gaussian fixed-point of a conventional nonlinear sigma-model. For a  $G/H$  nonlinear sigma-model with Lagrangian  $L = -\frac{1}{2g_0} \mathbf{m}_{ij}(\varphi) \partial^\mu \varphi^i \partial_\mu \varphi^j$  (with  $\mathbf{m}$  satisfying Equation (3.12)) the beta function  $\beta_{G/H}(g_0) = g_0 \sum_{l \leq 1} l \zeta_l (\frac{g_0}{2\pi})^l$  has only the trivial zero  $g_0^* = 0$ . As  $g_0 \rightarrow 0$  the renormalized Lagrangian blows up, but in an expansion around  $\mathbf{m}_{ij}(\varphi) = \delta_{ij}$  one can see that for  $g_0 \rightarrow 0$  the interaction terms vanish. In this sense the fixed point  $g_0^* = 0$  is Gaussian. This holds irrespective of the sign of  $\zeta_1$ , which however determines the stability properties of the flow. The stability ‘matrix’ vanishes so that the linearized stability analysis is empty. By direct inspection of the differential equation one sees that the unstable manifold of  $g_0^*$  is one-dimensional for  $\zeta_1 > 0$  (typical for  $G/H$  compact) and empty for  $\zeta_1 < 0$  (typical for  $G/H$  noncompact). Indeed,  $-\mu \frac{d}{d\mu} \bar{g}_0 = \frac{\zeta_1}{2\pi} \bar{g}_0^2 + O(\bar{g}_0^3)$ , and if one insists on  $\bar{g}_0 \geq 0$  for positivity-of-energy reasons, the flow will be attracted to  $g_0^* = 0$  for  $\mu \rightarrow \infty$  iff  $\zeta_1 > 0$ . In particular for  $\zeta_1 < 0$  these models are, based on the Criterion (PTC2) of Section 2, not expected to have a genuine continuum limit.

The Gaussian fixed point of the symmetry reduced gravity theories can be analyzed similarly. In terms of

$$g(\rho, \lambda) := \frac{\lambda}{2\pi h(\rho, \lambda)} \quad (3.98)$$

the flow equation (3.83) reads

$$\begin{aligned} \mu \frac{d}{d\mu} \bar{g} &= \bar{g}^2 \rho \partial_\rho \left[ \frac{1}{\bar{g}} \int_\rho^\infty \frac{du}{u} \sum_{l \geq 1} l \zeta_l \bar{g}(u)^l \right] \\ &= -\zeta_1 \left[ \bar{g}^2 + \rho \partial_\rho \bar{g} \int_\rho^\infty \frac{du}{u} \bar{g}(u) \right] + O(\bar{g}^3). \end{aligned} \quad (3.99)$$

Clearly  $g_*(\rho) \equiv 0$  is a fixed point (function) and in a similar sense as before it can be interpreted as a Gaussian fixed point. In contrast to the non-Gaussian fixed point the linearized stability analysis is now empty (just as it is for the  $G/H$  sigma-model flow). One thus has to cope with the nonlinear flow equation (3.99) at least to quadratic order. This is cumbersome but the qualitative feature of interest here can readily be understood: With respect to the Gaussian fixed point  $g_*(\rho) \equiv 0$  *not* all couplings contained in the generating functional  $g(\cdot)$  are asymptotically safe. That is, there exists initial data  $g_0(\rho) = \bar{g}_0(\rho, \mu_0)$  (with  $g_0(\rho) \rightarrow 0$ , for  $\rho \rightarrow \infty$ ) for which the  $\mu \rightarrow \infty$  asymptotics does not vanish identically in  $\rho$ . To see this, it suffices to note that the right-hand-side of Equation (3.99) to quadratic order reads  $\zeta_1 \bar{g}^2 \rho \partial_\rho I(\bar{g})$ , with  $I(g) := g^{-1} \int_\rho^\infty \frac{du}{u} g(u)$ . Since  $\zeta_1 < 0$  always, initial data  $g_0$  for which  $I(g_0)$  is a strictly increasing function of  $\rho$  will give rise to solutions having the tendency to be driven towards larger values (pointwise in  $\rho$ ) as  $\mu$  increases. Conversely, only initial data  $g_0$  for which  $I(g_0)$  is strictly decreasing in  $\rho$  can be expected to give rise to a solution  $g_0$  which vanishes identically in  $\rho$  as  $\mu \rightarrow \infty$ . A rigorous theorem describing the stable and the unstable manifold is presumably hard to come by, but for our purposes it is enough to know that there exist initial data which do not give rise to solutions decaying to  $g_*(\rho) \equiv 0$  for  $\mu \rightarrow \infty$ . For example  $g_0(\rho) = \rho e^{-a\rho}$  or  $g_0(\rho) = \exp(-\sqrt{a+b} \ln \rho)$ ,  $a, b > 0$ , have this property. The upshot is that the Gaussian fixed point of the symmetry truncated gravity theories is not UV stable, the Condition (PTC2) is not satisfied, and one can presumably not use it for the construction of a genuine continuum limit.



At this point it may be instructive to contrast the quantum properties of the dimensionally reduced gravity theories with those of the same noncompact  $G/H$  sigma-model without coupling to gravity (which effectively amounts to setting  $\rho$  constant in Equation (3.16)). The qualitative differences are summarized in Table 2.

$G/H$ sigma-model	dimensionally reduced gravity with $G/H$
renormalizable	non-renormalizable
one essential coupling	$\infty$ essential couplings
$g_0$	function $h(\cdot)$
flow is formally infrared free	flow is asymptotically safe
formal trivial fixed point	non-trivial fixed point
$g_0 = 0$	$h^{\text{beta}}(\cdot)$
formally IR stable	UV stable
trace anomalous	trace anomaly vanishes

Table 2: Comparison: noncompact  $G/H$  sigma-model vs. dimensionally reduced gravity theory with  $G/H$  coset.

The comparison highlights why the above conclusion is surprising and significant. While the noncompact  $G/H$  sigma-models are renormalizable with just one relevant coupling (denoted by  $g$  in the table), at least in the known constructions they do not have a fixed point at which they are conformally invariant. Their gravitational counterparts require infinitely many relevant couplings for their UV renormalization. This infinite coupling flow has a nontrivial UV fixed point at which the theory is conformally invariant. Most importantly the stability properties of the renormalization flow are reversed (compared to the flow of  $g$ ) for all of the infinitely many relevant couplings. As there appears to be no structural reason for this surprising reversal in the reduced theory itself, we regard it as strong evidence for the existence of an UV stable fixed point for the full renormalization group dynamics.

In the table we anticipated that at the fixed point the trace anomaly of the would-be energy momentum tensor vanishes for the symmetry truncated gravity theories. This allows one to construct quantum counterparts of the constraints  $\mathcal{H}_0$  and  $\mathcal{H}_1$  as well-defined composite operators. In detail this comes about as follows. Taking the trace in Equation (3.81) gives  $\llbracket T^\mu{}_\mu(h; \phi) \rrbracket = \llbracket t^\mu{}_\mu(h; \phi) \rrbracket - \partial^2 \llbracket V \rrbracket$ , again modulo the equations of motion operator. The first term has a nonzero trace anomaly given by

$$\llbracket t^\mu{}_\mu(h; \phi) \rrbracket = \left[ \frac{\lambda}{h} \beta_h \left( \frac{h}{\lambda} \right) L \right] + \frac{1}{2\lambda} \llbracket \mathcal{L}_{W-\hat{\Xi}} \mathfrak{h}_{ij}(\phi) \partial^\mu \phi^i \partial_\mu \phi^j \rrbracket. \tag{3.100}$$

Here  $\hat{\Xi}$  is the field vector in Equation (3.76); furthermore  $W^i = (0, \dots, 0, \frac{\rho}{b h} W(\rho, \lambda))$ , where  $W(\rho, \lambda)$  is a functional of  $h$  which receives contributions only at three and higher loop orders (see Section B.3 and [154]). The improvement term is determined by the function  $f = f[h]$  in Equation (3.81), which in turn is largely determined by  $h$  and thus *cannot* be freely chosen as a function of  $\rho$ .

It is therefore a very nontrivial match that (i) upon insertion of the fixed point  $h = h^{\text{beta}}$  the function  $f^{\text{beta}}(\rho) := f[h^{\text{beta}}](\rho)$  becomes stationary ( $\mu$ -independent), (ii) the equation  $F_f[h^{\text{beta}}] f^{\text{beta}} = 0$  turns out to determine  $f^{\text{beta}}(\rho)$  completely, and (iii) the so-determined function has the property that  $\partial^2 \llbracket V \rrbracket|_{f=f^{\text{beta}}}$  precisely cancels the second term in Equation (3.100) evaluated for  $h^{\text{beta}}$ . Thus the trace anomaly vanishes precisely at the fixed point of the coupling flow

$$\llbracket T^\mu{}_\mu(h^{\text{beta}}; \phi) \rrbracket = 0. \tag{3.101}$$

As should be clear from the derivation this is a nontrivial *property* of the system, rather than one which is used to define the renormalized target space metric  $\mathfrak{h}_{ij}(\phi)$ . The latter is already determined by the warped product structure (3.59, 3.60) and the renormalization result (3.57). At the fixed point  $h^{\text{beta}}$  the first term in Equation (3.100) vanishes, but the second then is completely determined. On the other hand by Equation (3.81) one is not free to choose the improvement potential as a function of  $\rho$ -independent of  $h$ , so the cancellation is not built in. One can also verify that the *only* solution for  $f[h]$  such that  $\llbracket T^\mu{}_\mu(h; \phi) = \partial^\mu K_\mu(h) \rrbracket$  is  $h = h^{\text{beta}}$  with the above  $f^{\text{beta}}$ , such that Equation (3.101) holds. That is, scale invariance implies here conformal invariance. Since the target space metric (3.60) has indefinite signature this does not follow from general grounds [176, 194].

Due to Equation (3.101) we can now define the quantum constraints by

$$\llbracket \mathcal{H}_0 \rrbracket := \llbracket T_{00} \rrbracket, \quad \llbracket \mathcal{H}_1 \rrbracket := \llbracket T_{01} \rrbracket. \quad (3.102)$$

The linear combinations  $\llbracket \mathcal{H}_0 \pm \mathcal{H}_1 \rrbracket$  are thus expected to generate two commuting copies of a Virasoro algebra with formal central charge  $c = 2 + \dim G/H$ . This central charge is only formal because it refers to a state space with indefinite norm (see [127, 128] for an anomaly-free implementation of the Virasoro constraints in essentially noninteracting systems). In the case at hand the proper positivity requirement will be determined by the quantum observables commuting with the constraints. Their construction and the exploration of the physical state space is a major desideratum. In summary, the systems should at the fixed point be described by one whose physical states can be set into correspondence to the above quantum observables. The infrared problem has not been investigated so far, but based on results in the polarized subsector [153] one might expect it to be benign.

Despite the fact that the system is conformally invariant at the fixed point there are still scale dependent running parameters. At first sight this seems paradoxical. However, already the example of the massless continuum limit in a 3D scalar field theory exhibits this behavior. The remaining scale parameter is related to the direction of instability within the critical manifold, pointing from the Gaussian to the Fisher–Wilson fixed point, in the direction of coarse graining. The systems considered here provide an intriguing other example of this phenomenon. The critical manifold can be identified with the subset of parameter values where the system is scale (and here conformally) invariant. This fixes  $h(\cdot) = h^{\text{beta}}(\cdot)$ , but the inessential parameters contained in the renormalized fields are left unconstrained. This allows one to introduce a running parameter as follows. One evaluates the running coupling function  $\bar{h}$  at the ‘comoving’ field  $\bar{\rho}$  and sets

$$\bar{g}_h(\mu) := \frac{1}{\bar{h}(\bar{\rho}, \mu)}. \quad (3.103)$$

This quantity carries a two-fold  $\mu$ -dependence, one via the running coupling  $\mu \mapsto \bar{h}(\cdot, \mu)$  and one because now the argument at which the function is evaluated is likewise  $\mu$ -dependent. Since  $\rho$  is a field on the base manifold, the quantity  $\bar{g}_h(\mu)$  depends parametrically on the *value* of  $h(\rho(x))$  – and hence on  $x$ . Combining Equation (3.83) with Equation (3.75) one finds the following flow equation:

$$\mu \frac{d}{d\mu} \lambda \bar{g}_h = \beta_{G/H}(\lambda \bar{g}_h). \quad (3.104)$$

These are the usual flow equations for the single coupling  $G/H$  sigma-model without coupling to gravity! In other words the ‘gravitationally dressed’ functional flow for  $\bar{h}$  has been ‘undressed’ by reference to the scale dependent ‘rod field’  $\bar{\rho}$  (the term is adapted from H. Weyl’s “Maßstabsfeld”). The Equations (3.104) are not by themselves useful for renormalization purposes – which requires determination of the flow of  $\bar{h}(\cdot, \mu)$  with respect to a *fixed* set of field coordinates. Moreover in the technical sense  $\bar{g}_h$  is an “inessential” parameter. The fact that the scale dependence of  $\bar{g}_h$  is

governed by the beta function  $\beta_{G/H}$  means that for increasing  $\mu$  it will be driven away from the fixed point  $g_h = 0$ . The condition  $g_h(x, \lambda) \equiv 0$  can be traded for the specification of the Gaussian fixed point  $g(\varrho; \lambda) \equiv 0$ . Thus the parameter flow  $\bar{g}_h$  may be viewed as a coupling flow emanating (in the direction of increasing  $\mu$ ) from the Gaussian fixed point. At the non-Gaussian fixed point, on the other hand,  $\bar{g}_h$  governs the scale dependence of the ‘rod field’  $\bar{\rho}^{\text{beta}}(x, \mu) := \bar{\rho}(x, \mu)|_{h=h^{\text{beta}}}$  via

$$\mu \frac{d}{d\mu} \bar{\rho}^{\text{beta}} = \zeta_1 \frac{\lambda}{2\pi} \bar{\rho}^{\text{beta}} \bar{g}_{h^{\text{beta}}} [\bar{\rho}^{\text{beta}}]. \quad (3.105)$$

This follows from Equation (3.75) and the relation  $h\bar{\Xi}^3[h]|_{h=h^{\text{beta}}} = \rho\zeta_1 \frac{\lambda}{2\pi}$ . Here we indicated the functional dependence of  $\bar{g}_h$  on  $\bar{\rho}$ , which at the non-Gaussian fixed point gives a dependence of  $\bar{g}_{h^{\text{beta}}}$  on  $\bar{\rho}^{\text{beta}}$ . Since  $\zeta_1 < 0$ , one sees from Equation (3.105) that  $\bar{\rho}^{\text{beta}}$  is pointwise for all  $x$  a decreasing function of  $\mu$ , at least locally in  $\mu$ . In addition Equation (3.87) implies

$$\int^{\bar{\rho}^{\text{beta}}} \frac{du}{u} h^{\text{beta}}(u) = \chi^+(x^+, \mu) + \chi^-(x^-, \mu). \quad (3.106)$$

Here  $\chi^\pm(\cdot, \mu)$  are functions of one variable which by Equation (3.105) are locally decreasing in  $\mu$  and  $x^\pm = (x^0 \pm x^1)/2$  are lightcone coordinates. Since the theory is conformally invariant at the fixed point (of the couplings) one can change coordinates  $x^\pm \mapsto \chi^\pm$  to bring scaling operators into a standard form. The upshot is, as anticipated in Section 3.2, that the rod field  $\bar{\rho}^{\text{beta}}$  describes the local scale changes dynamically induced by quantum gravity and defines a resolution scale for the geometries.

For orientation we summarize here our results on the renormalization of the symmetry reduced Quantum Einstein Gravity theories (3.1):

1. The systems inherit the lack of standard perturbative renormalizability from the full theory. A cut-off independent quantum theory can be achieved at the expense of introducing infinitely many couplings combined into a generating function  $h(\cdot)$  of one variable.
2. The argument of this function is the ‘area radius’ field  $\rho$  associated with the two Killing vectors. The field  $\rho$  is (nonlinearly) renormalized but no extra renormalizations are needed to define arbitrary powers thereof.
3. A universal formula for the beta functional for  $h$  and hence for the infinitely many couplings contained in it can be given. The flow possesses a Gaussian as well as a non-Gaussian fixed point. With respect to the non-Gaussian fixed point *all* couplings in  $h$  are asymptotically safe.
4. At the fixed point the trace anomaly vanishes and the quantum constraints (well-defined as composite operators)  $[[\mathcal{H}_0]]$ ,  $[[\mathcal{H}_1]]$  can in principle be imposed. The linear combinations  $[[\mathcal{H}_0 \pm \mathcal{H}_1]]$  are expected to generate commuting copies of a centrally extended conformal algebra acting on an indefinite metric Hilbert space.
5. Despite the conformal invariance at the fixed point there is a scale dependent local parameter, whose scale dependence is governed by the beta function of the  $G/H$  sigma-model without coupling to gravity.

So far we considered the renormalization of the symmetry reduced theories in its own right, leaving the embedding into the full Quantum Gravidynamics open. The proposed relation to qualitative aspects of the Quantum Gravidynamics in the extreme UV has already been mentioned. Here we offer some tentative remarks on the embedding otherwise. The constructions presented in this section can be extended to  $2 + \epsilon$  dimensions in the spirit of an  $\epsilon$ -expansion. At the same

time this mimics quantum aspects of the 1-Killing vector reduction. One finds that the qualitative features of the renormalization flow – non-Gaussian fixed point and asymptotic safety – are still present [156]. A cosmological ‘constant’ term can likewise be included and displays a similar pattern as outlined at the end of Section 3.2. The advantage of this setting is that the UV cutoff can strictly be removed, which is hard to achieve with a nonperturbative technique. The extension of these results from a quasi-perturbative analysis to a nonperturbative one, ideally via controlled approximations, is an important open problem. The same holds for the analysis of the 1-Killing vector reduction, which holds the potential for cosmological applications. These truncations can be viewed as complementary to the ‘hierarchical’ truncations used in Section 4: A manifest truncation is initially imposed on the functional integral, but the infinite coupling renormalization flow can then be studied in great detail, often without further approximations.

## 4 Asymptotic Safety from the Effective Average Action

As surveyed in the introduction, important evidence for the asymptotic safety scenario comes from the truncated flow of the effective average action, as computed from its functional renormalization flow equation (FRGE). In this section we use the term Quantum Einstein Gravity to refer to a version of Quantum Gravidynamics where the metric is used as the dynamical variable. The key results have been outlined in the introduction. Here we present in more detail the effective average action for gravity, its flow equation, and the results obtained from its truncations.

### 4.1 The effective average action for gravity and its FRGE

The effective average action is a scale dependent variant  $\Gamma_k$  of the usual effective action  $\Gamma$ , modified by a mode-cutoff  $k$ , such that  $\Gamma_k$  can be interpreted as describing an ‘effective field theory at scale  $k$ ’. For non-gauge theories a self-contained summary of this formalism can be found in Appendix C. In the application to gauge theories and gravity two conceptual problems occur.

First the standard effective action is not a gauge invariant functional of its argument. For example in a Yang–Mills theory one gauge-fixes the functional integral with an ordinary gauge fixing condition like  $\partial^\mu A_\mu^a = 0$ , couples the Yang–Mills field  $A_\mu^a$  to a source, and constructs the ordinary effective action, the resulting functional  $\Gamma[A_\mu^a]$  is not invariant under the gauge transformations of  $A_\mu^a$ . Although physical quantities extracted from  $\Gamma[A_\mu^a]$  are expected to be gauge invariant, the noninvariance is cumbersome for renormalization purposes. The second problem is related to the fact that in a gauge theory a ‘coarse graining’ based on a naive Fourier decomposition of  $A_\mu^a(x)$  is not gauge covariant and hence not physical. In fact, if one were to gauge transform a slowly varying  $A_\mu^a(x)$  with a parameter function  $\omega(x)$  with a fast  $x$ -variation, a gauge field with a fast  $x$ -variation would arise, which however still describes the same physics.

Both problems can be overcome by using the background field formalism. The background effective action generally is a gauge invariant functional of its argument (see Appendix B). The second problem is overcome by using the spectrum of a covariant differential operator built from the *background field* configuration to discriminate between slow modes (small eigenvalues) and fast modes (large eigenvalues) [187]. This sacrifices to some extent the intuition of a spatial coarse graining, but it produces a gauge invariant separation of modes. Applied to a non-gauge theory it amounts to expanding the field in terms of eigenfunctions of the (positive) operator  $-\partial^2$  and declaring its eigenmodes ‘long’ or ‘short’ wavelength depending on whether the corresponding  $p^2$  is smaller or larger than a given  $k^2$ .

This is the strategy adopted to define the effective average action for gravity [179]. In short: The effective average action for gravity is a variant of the background effective action  $\Gamma[\langle f_{\alpha\beta} \rangle, \sigma^\alpha, \bar{\sigma}_\alpha; \bar{g}_{\alpha\beta}]$  described in Appendix B (see Equations (B.48, B.51)), where the bare action is modified by mode cutoff terms as in Appendix C, but with the mode cutoff defined via the spectrum of a covariant differential operator built from the background metric. For convenience we quickly recapitulate the main features of the background field technique here and then describe the modifications needed for the mode cutoff.

The initial bare action  $S[g]$  is assumed to be a reparameterization invariant functional of the metric  $g = (g_{\alpha\beta})_{1 \leq \alpha, \beta \leq d}$ . Infinitesimally the invariance reads  $S[g + \mathcal{L}_v g] = S[g] + O(v^2)$ , where  $\mathcal{L}_v g$  is the Lie derivative of  $g$  with respect to the vector field  $v^\gamma \partial_\gamma$ . The metric  $g$  (later the integration variable in the functional integral) is decomposed into a background  $\bar{g}$  and a fluctuation  $f$ , i.e.  $g_{\alpha\beta} = \bar{g}_{\alpha\beta} + f_{\alpha\beta}$ . The fluctuation field  $f_{\alpha\beta}$  is then taken as the dynamical variable over which the functional integral is performed; it is not assumed to be small in some sense, no expansion in powers of  $f$  is implied by the split. Note however that this linear split does not have a geometrical meaning in the space of geometries. The symmetry variation  $g_{\alpha\beta} \mapsto g_{\alpha\beta} + \mathcal{L}_v g_{\alpha\beta}$  can be decomposed in two

different ways

$$f_{\alpha\beta} \mapsto f_{\alpha\beta} + \mathcal{L}_v(\bar{g} + f)_{\alpha\beta}, \quad \bar{g}_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta}, \quad (4.1)$$

$$f_{\alpha\beta} \mapsto f_{\alpha\beta} + \mathcal{L}_v f_{\alpha\beta}, \quad \bar{g}_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} + \mathcal{L}_v \bar{g}_{\alpha\beta}. \quad (4.2)$$

We shall refer to the first one as “genuine gauge transformations” and to the second one as the “background gauge transformations”. The background effective action  $\Gamma[\langle f_{\alpha\beta} \rangle, \sigma^\alpha, \bar{\sigma}_\alpha; \bar{g}_{\alpha\beta}]$  is a functional of the expectation value  $\langle f_{\alpha\beta} \rangle$  of the fluctuation variable, the background metric  $\bar{g}_{\alpha\beta}$ , and the expectation values of the ghost fields  $\sigma^\alpha = \langle C^\alpha \rangle$ ,  $\bar{\sigma}_\alpha = \langle \bar{C}_\alpha \rangle$ . Importantly  $\Gamma[\langle f_{\alpha\beta} \rangle, \sigma^\alpha, \bar{\sigma}_\alpha; \bar{g}_{\alpha\beta}]$  is invariant under the background field transformations (4.2). So far one should think of the background geometry as being prescribed but of generic form; eventually it is adjusted self-consistently by a condition involving the full effective action (see Equation (2.48) and Appendix B).

In the next step the initial bare action should be replaced by one involving a mode cutoff term. In the background field technique the mode cutoff should be done in a way that preserves the invariance under the background gauge transformations (4.2). We now first present the steps leading to the scale dependent effective average action  $\Gamma_k[\langle g_{\alpha\beta} \rangle - \bar{g}_{\alpha\beta}, \sigma^\alpha, \bar{\sigma}_\alpha; \bar{g}_{\alpha\beta}]$  in some detail and then present the FRGE for it. The functional integrals occurring are largely formal; for definiteness we consider the Euclidean variant where the integral over Riemannian geometries is intended. The precise definition of the generating functionals is not essential here, as they mainly serve to arrive at the gravitational FRGE. The latter provides a novel tool for investigating the gravitational renormalization flow.

We begin by introducing a scale dependent variant  $W_k$  of the generating functional of the connected Greens functions. The cutoff scale is again denoted by  $k$ , it has unit mass dimension, and no physics interpretation off hand. The defining relation for  $W_k$  reads

$$W_k[J, j, \bar{j}; \bar{g}] = \int \mathcal{D}\mu_{J, \bar{j}, k}[f] \exp \left\{ -S[\bar{g} + f] - C_k^{\text{grav}}[f; \bar{g}] + \int dx \sqrt{\bar{g}} J^{\alpha\beta} f_{\alpha\beta} \right\}. \quad (4.3)$$

Here the measure  $\mathcal{D}\mu_{J, \bar{j}, k}[f]$  differs from the naive one,  $\mathcal{D}f_{\alpha\beta}$ , by gauge fixing terms and an integration over ghost fields  $C^\alpha, \bar{C}_\alpha$ , where the action for the latter is again modified by a mode-cutoff:

$$\begin{aligned} \mathcal{D}\mu_{J, \bar{j}, k}[f] = \mathcal{D}f_{\alpha\beta} \int \mathcal{D}C^\alpha \mathcal{D}\bar{C}_\alpha \exp \left\{ -\frac{1}{2\alpha} \int dx \sqrt{\bar{g}} \bar{g}^{\alpha\beta} Q_\alpha Q_\beta + \frac{1}{\kappa} \int dx \bar{C}_\alpha \bar{g}^{\alpha\beta} \frac{\partial Q_\beta}{\partial f_{\gamma\delta}} \mathcal{L}_C(\bar{g} + f)_{\gamma\delta} \right. \\ \left. - C_k^{\text{gh}}[C, \bar{C}] + \int dx \sqrt{\bar{g}} [\bar{j}_\alpha C^\alpha + j^\alpha \bar{C}_\alpha] \right\}. \end{aligned} \quad (4.4)$$

The first term in the exponent is the gauge fixing term. The gauge fixing condition  $Q_\alpha = Q_\alpha[\bar{g}; f] \approx 0$  must be invariant under Equation (4.2), for the moment we may leave it unspecified. The second term is the Faddeev–Popov action for the ghosts obtained in the usual way: One applies a genuine gauge transformation (4.1) to  $Q_\alpha$  and replaces the parameter  $v^\alpha$  by the ghost field  $C^\alpha$ . The integral over  $C^\alpha$  and  $\bar{C}_\alpha$  then exponentiates the Fadeev–Popov determinant  $\det[\delta Q_\alpha / \delta v^\beta]$ . This gauge fixing procedure has a somewhat perturbative flavor; large scale aspects of the space of geometries are not adequately taken into account. The terms  $C_k^{\text{grav}}$  and  $C_k^{\text{gh}}$  implement the mode cutoff in the gravity and the ghost sector, respectively. We shall specify them shortly. Finally we coupled in Equation (4.4) the ghosts to sources  $j^\alpha, \bar{j}_\alpha$  for later use.

The construction of the effective average action now parallels that in the scalar case. We quickly run through the relevant steps. The Legendre transform of  $W_k$  at fixed  $\bar{g}_{\alpha\beta}$  is

$$\tilde{\Gamma}_k[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] = \sup_{J, \bar{j}} \left\{ \int dx \sqrt{\bar{g}} (\bar{f}_{\alpha\beta} J^{\alpha\beta} + \sigma^\alpha \bar{j}_\alpha + \bar{\sigma}_\alpha j^\alpha) - W_k[J, j, \bar{j}; \bar{g}] \right\}. \quad (4.5)$$



As usual, if  $W_k$  is differentiable with respect to the sources, the extremizing source configurations  $J_*^{\alpha\beta}$ ,  $\bar{j}_{*,\alpha}$ ,  $J_*^\alpha$  allow one to interpret  $\bar{f}_{\alpha\beta}$ ,  $\sigma^\alpha$ ,  $\bar{\sigma}_\alpha$  as the expectation values of  $f_{\alpha\beta}$ ,  $C^\alpha$ ,  $\bar{C}_\alpha$  via

$$\bar{f}_{\alpha\beta} = \langle f_{\alpha\beta} \rangle = \frac{1}{\sqrt{\bar{g}}} \frac{\delta W_k}{\delta J_*^{\alpha\beta}}, \quad \sigma^\alpha = \langle C^\alpha \rangle = \frac{1}{\sqrt{\bar{g}}} \frac{\delta W_k}{\delta J_{*,\alpha}^\alpha}, \quad \bar{\sigma}_\alpha = \langle \bar{C}_\alpha \rangle = \frac{1}{\sqrt{\bar{g}}} \frac{\delta W_k}{\delta J_*^\alpha}. \quad (4.6)$$

Note that the expectation values defined through Equation (4.3) are in general both  $k$ -dependent and source dependent. In Equation (4.6), by construction, the  $k$ -dependence carried by  $W_k$  cancels that carried by the extremizing source. Concretely the extremizing sources are constructed by assuming that  $W_k$  has a series expansion in powers of the sources with  $\bar{g}$ -dependent coefficients; formal inversion of the series then gives a  $k$ -dependent  $J_*^{\alpha\beta}[\bar{f}, \sigma, \bar{\sigma}; \bar{g}]$  with the property that  $J_*^{\alpha\beta}[0, \sigma, \bar{\sigma}; \bar{g}] = 0$ , and similarly for the ghost sources. The formal effective field equations dual to Equation (4.6) read

$$\begin{aligned} \frac{\delta}{\delta \bar{f}_{\alpha\beta}} \tilde{\Gamma}_k[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] &= J_*^{\alpha\beta}[\bar{f}, \sigma, \bar{\sigma}; \bar{g}], \\ \frac{\delta}{\delta \bar{\sigma}^\alpha} \tilde{\Gamma}_k[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] &= J_*^\alpha[\bar{f}, \sigma, \bar{\sigma}; \bar{g}], \\ \frac{\delta}{\delta \bar{\sigma}^\alpha} \tilde{\Gamma}_k[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] &= \bar{j}_{*,\alpha}[\bar{f}, \sigma, \bar{\sigma}; \bar{g}]. \end{aligned} \quad (4.7)$$

As in the scalar case the effective average action differs from  $\tilde{\Gamma}_k$  by the cutoff action with the expectation value fields inserted,

$$\Gamma_k[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] := \tilde{\Gamma}_k[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] - C_k^{\text{grav}}[\bar{f}; \bar{g}] - C_k^{\text{gh}}[\sigma, \bar{\sigma}; \bar{g}]. \quad (4.8)$$

Sometimes it is convenient to introduce  $\langle g_{\alpha\beta} \rangle := \bar{g}_{\alpha\beta} + \bar{f}_{\alpha\beta}$ , which is the expectation value of the original ‘quantum’ metric  $g_{\alpha\beta} = \bar{g}_{\alpha\beta} + f_{\alpha\beta}$ , and to regard  $\Gamma_k$  as a functional of  $\langle g_{\alpha\beta} \rangle$  rather than  $\bar{f}_{\alpha\beta}$ , i.e.  $\Gamma_k[\langle g_{\alpha\beta} \rangle, \bar{g}_{\alpha\beta}, \sigma^\alpha, \bar{\sigma}_\alpha] := \Gamma_k[\langle g_{\alpha\beta} \rangle - \bar{g}_{\alpha\beta}, \sigma, \bar{\sigma}; \bar{g}]$ .

Usually one is not interested in correlation functions involving Faddeev–Popov ghosts and it is sufficient to know the reduced functional

$$\bar{\Gamma}_k[g] := \Gamma_k[0, 0, 0; g] \equiv \Gamma_k[g, g, 0, 0]. \quad (4.9)$$

As indicated we shall simply write  $g_{\alpha\beta}$  for its argument  $\langle g_{\alpha\beta} \rangle = \bar{g}_{\alpha\beta}$ .

The precise form of the gauge condition  $Q_\alpha[\bar{g}; f]$  is inessential, only the invariance under Equation (4.2) is important. It ensures that the associated ghost action is invariant under Equation (4.2) and  $\delta C^\alpha = \mathcal{L}_v C^\alpha$ ,  $\delta \bar{C}_\alpha = \mathcal{L}_v \bar{C}_\alpha$ . We shall ignore the problem of the global existence of gauge slices (“Gribov copies”), in accordance with the formal nature of the construction. For later reference let us briefly describe the most widely used gauge condition, the “background harmonic gauge” which reads

$$Q_\alpha[\bar{g}, f] := \sqrt{2}\kappa \bar{\nabla}^\beta \left[ f_{\alpha\beta} - \frac{1}{2} \bar{g}_{\alpha\beta} \bar{g}^{\gamma\gamma'} f_{\gamma\gamma'} \right]. \quad (4.10)$$

The covariant derivative  $\bar{\nabla}_\alpha$  involves the Christoffel symbols  $\bar{\Gamma}_{\alpha\beta}^\gamma$  of the background metric. Note that  $Q_\alpha$  of Equation (4.10) is linear in the quantum field  $f_{\alpha\beta}$ . On a flat background with  $\bar{g}_{\alpha\beta} = \eta_{\alpha\beta}$  the condition  $Q_\alpha = 0$  reduces to the familiar harmonic gauge condition,  $\partial^\beta f_{\beta\alpha} = \frac{1}{2} \partial_\alpha f_{\beta\beta}$ . In Equation (4.10)  $\kappa$  is an arbitrary constant with the dimension of a mass. We shall set  $\kappa = (2g_N)^{-1/2}$  and interpret  $g_N = 16\pi G$  as the bare Newton constant. The ghost action for the gauge condition (4.10) is

$$-\sqrt{2} \int dx \sqrt{\bar{g}} \bar{C}_\alpha \mathcal{M}[\bar{g} + f, \bar{g}]^\alpha{}_\beta C^\beta, \quad (4.11)$$

$$\text{with } \mathcal{M}[g, \bar{g}]^\alpha{}_\beta = \bar{g}^{\alpha\rho} \bar{g}^{\sigma\lambda} \bar{\nabla}_\lambda (g_{\rho\beta} \nabla_\sigma + g_{\sigma\beta} \nabla_\rho) - \bar{g}^{\rho\sigma} \bar{g}^{\alpha\lambda} \bar{\nabla}_\lambda g_{\sigma\beta} \nabla_\rho, \quad (4.12)$$

where  $\nabla_\alpha$  and  $\bar{\nabla}_\alpha$  are the covariant derivatives associated with  $g_{\alpha\beta}$  (as a short for  $\bar{g}_{\alpha\beta} + f_{\alpha\beta}$ ) and  $\bar{g}_{\alpha\beta}$ , respectively.

The last ingredient in Equations (4.3, 4.4) to be specified are the mode cutoff terms, not present in the usual background effective action. Their precise form is arbitrary to some extent. Naturally they will be taken quadratic in the respective fields, with a kernel which is covariant under background gauge transformations. These requirements are met if

$$\begin{aligned} C_k^{\text{grav}}[f; \bar{g}] &= \frac{\kappa^2}{2} \int dx \sqrt{\bar{g}} f_{\alpha\beta} \mathcal{R}_k[\bar{g}]^{\alpha\beta\gamma\delta} f_{\gamma\delta}, \\ C_k^{\text{gh}}[C, \bar{C}; \bar{g}] &= \sqrt{2} \int dx \sqrt{\bar{g}} \bar{C}_\alpha \mathcal{R}_k^{\text{gh}}[\bar{g}] C^\alpha, \end{aligned} \quad (4.13)$$

and the kernels  $\mathcal{R}_k[\bar{g}]$ ,  $\mathcal{R}_k^{\text{gh}}[\bar{g}]$  transform covariantly under  $\bar{g}_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} + \mathcal{L}_v \bar{g}_{\alpha\beta}$ . In addition they should effectively suppress covariant ‘momentum modes’ with ‘momenta’  $p^2 < k^2$ . As mentioned earlier one way of defining such a covariant scale is via the spectrum of a covariant differential operator. Concretely the following choice will be used.

Consider the Laplacian  $\Delta_{\bar{g}} := \bar{g}^{\alpha\beta} \bar{\nabla}_\alpha \bar{\nabla}_\beta$  of the Riemannian background metric  $\bar{g}$  with  $\bar{\nabla}$  being its torsionfree connection. We assume  $\bar{g}$  to be such that  $-\Delta_{\bar{g}}$  has a non-negative spectrum and a complete set of (tensorial) eigenfunctions. The spectral values  $\lambda = \lambda[\bar{g}]$  of  $-\Delta_{\bar{g}}$  will then be functionals of  $\bar{g}$  and one can choose  $\mathcal{R}_k[\bar{g}]$  and  $\mathcal{R}_k^{\text{gh}}[\bar{g}]$  such that only eigenmodes with spectral values  $\lambda[\bar{g}] \gg k^2$  (2 being the mass dimension of the operator) enter the  $f_{\alpha\beta}$  functional integral unsuppressed. Here one should think of the  $f_{\alpha\beta}$  functional integral as being replaced by one over the (complete system of) eigenfunctions of  $-\Delta_{\bar{g}}$ , for a fixed  $\bar{g}$ . Concretely, for  $\mathcal{R}_k[\bar{g}]$  and  $\mathcal{R}_k^{\text{gh}}[\bar{g}]$  we take expressions of the form

$$\mathcal{R}_k[\bar{g}]^{\alpha\beta\gamma\delta} = \mathcal{Z}_k^{\alpha\beta\gamma\delta}[\bar{g}] k^2 \mathcal{R}^{(0)}(-\Delta_{\bar{g}}/k^2), \quad \mathcal{R}_k^{\text{gh}}[\bar{g}] = \mathcal{Z}_k^{\text{gh}} k^2 \mathcal{R}^{(0)}(-\Delta_{\bar{g}}/k^2). \quad (4.14)$$

As indicated in Equation (4.14) the prefactors  $\mathcal{Z}_k$  are different for the gravitational and the ghost cutoff. For the ghosts  $\mathcal{Z}_k^{\text{gh}}$  is a pure number, whereas for the metric fluctuation  $\mathcal{Z}_k^{\alpha\beta\gamma\delta}[\bar{g}]$  is a tensor constructed from the background metric  $\bar{g}_{\alpha\beta}$ . We shall discuss the choice of these prefactors later on.

The essential ingredient in Equation (4.14) is a function  $\mathcal{R}^{(0)} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  interpolating smoothly between  $\mathcal{R}^{(0)}(0) = 1$  and  $\mathcal{R}^{(0)}(\infty) = 0$ ; for example

$$\mathcal{R}^{(0)}(u) = u[\exp(u) - 1]^{-1}. \quad (4.15)$$

Its argument  $u$  should be identified with the weighted spectral values  $\lambda[\bar{g}]/k^2$  of  $-\Delta_{\bar{g}}$ . One readily sees that then the exponentials in Equation (4.13) have the desired effect: They effectively suppress eigenmodes of  $-\Delta_{\bar{g}}$  with spectral values much smaller than  $k^2$ , while modes with  $\lambda[\bar{g}]$  large compared to  $k^2$  are unaffected. This also illustrates that a mode suppression can be defined covariantly using the background field formalism.

This concludes the definition of the effective average action and its various specializations. We now present its key properties.

#### 4.1.1 Properties of the effective average action

1. The effective average action is invariant under background field diffeomorphisms

$$\Gamma_k[\Phi + \mathcal{L}_v \Phi] = \Gamma_k[\Phi], \quad \Phi := \{g_{\alpha\beta}, \bar{g}_{\alpha\beta}, \sigma^\alpha, \bar{\sigma}_\alpha\}, \quad (4.16)$$

where all its arguments transform as tensors of the corresponding rank. This is a direct consequence of the corresponding property of  $W_k$  in Equation (4.3)

$$W_k[\mathcal{J} + \mathcal{L}_v \mathcal{J}] = W_k[\mathcal{J}], \quad \mathcal{J} := \{J^{\alpha\beta}, j^\alpha, \bar{j}_\alpha, \bar{g}_{\alpha\beta}\}, \quad (4.17)$$

where  $\mathcal{L}_v$  is the Lie derivative of the respective tensor type. Here Equation (4.16) is obtained as in Equation (B.47) of Appendix B, where the background covariance of the mode cutoff terms (4.13) is essential. Further the measure  $\mathcal{D}f_{\alpha\beta}$  is assumed to be diffeomorphism invariant.

2. It satisfies the functional integro-differential equation

$$\begin{aligned} \exp\{-\Gamma_k[\bar{f}, \sigma, \bar{\sigma}; \bar{g}]\} &= \int \mathcal{D}f \mathcal{D}C \mathcal{D}\bar{C} \exp\left\{-S_{\text{tot}}[f, C, \bar{C}; \bar{g}] - C_k[f - \bar{f}, C - \sigma, \bar{C} - \bar{\sigma}] \right. \\ &\quad \left. + \int dx \sqrt{\bar{g}} \left[ (f_{\alpha\beta} - \bar{f}_{\alpha\beta}) \frac{\delta\Gamma_k}{\delta\bar{f}_{\alpha\beta}} + (C^\alpha - \sigma^\alpha) \frac{\delta\Gamma_k}{\delta\sigma^\alpha} + (\bar{C}_\alpha - \bar{\sigma}_\alpha) \frac{\delta\Gamma_k}{\delta\bar{\sigma}_\alpha} \right] \right\}, \end{aligned} \quad (4.18)$$

where  $S_{\text{tot}} := S + S_{\text{gf}} + S_{\text{gh}}$  (with  $S_{\text{gf}}$  and  $S_{\text{gh}}$  minus the first two terms in the exponent of Equation (4.4) and  $C_k := C_k^{\text{grav}} + C_k^{\text{gh}}$ ).

3. The  $k$ -dependence of the effective average action is governed by an exact FRGE. Following the same lines as in the scalar case one arrives at [179]

$$\begin{aligned} k\partial_k\Gamma_k[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] &= \frac{1}{2} \text{Tr} \left[ \left( \Gamma_k^{(2)} + \widehat{\mathcal{R}}_k \right)_{\bar{f}\bar{f}}^{-1} \left( k\partial_k\widehat{\mathcal{R}}_k \right)_{\bar{f}\bar{f}} \right] \\ &\quad - \frac{1}{2} \text{Tr} \left[ \left\{ \left( \Gamma_k^{(2)} + \widehat{\mathcal{R}}_k \right)_{\bar{\sigma}\bar{\sigma}}^{-1} - \left( \Gamma_k^{(2)} + \widehat{\mathcal{R}}_k \right)_{\sigma\sigma}^{-1} \right\} \left( k\partial_k\widehat{\mathcal{R}}_k \right)_{\bar{\sigma}\bar{\sigma}} \right]. \end{aligned} \quad (4.19)$$

Here  $\Gamma_k^{(2)}$  denotes the Hessian of  $\Gamma_k$  with respect to the dynamical fields  $\bar{f}$ ,  $\sigma$ ,  $\bar{\sigma}$  at fixed  $\bar{g}$ . It is a block matrix labeled by the fields  $\varphi_i := \{\bar{f}_{\alpha\beta}, \sigma^\alpha, \bar{\sigma}_\alpha\}$ ,

$$\tilde{\Gamma}_k^{(2)ij}(x, y) := \frac{1}{\sqrt{\bar{g}(x)\bar{g}(y)}} \frac{\delta^2\tilde{\Gamma}_k}{\delta\varphi_i(x)\delta\varphi_j(y)}. \quad (4.20)$$

(In the ghost sector the derivatives are understood as left derivatives.) Likewise,  $\widehat{\mathcal{R}}_k$  is a block matrix with entries  $(\widehat{\mathcal{R}}_k)_{\bar{f}\bar{f}}^{\alpha\beta\gamma\delta} := \kappa^2 \mathcal{R}_k^{\text{grav}}[\bar{g}]^{\alpha\beta\gamma\delta}$  and  $\widehat{\mathcal{R}}_{\bar{\sigma}\bar{\sigma}} = \sqrt{2} \mathcal{R}_k^{\text{gh}}[\bar{g}]$ . Performing the trace in the position representation it includes an integration  $\int dx \sqrt{\bar{g}(x)}$  involving the background volume element. For any cutoff which is qualitatively similar to Equation (4.14, 4.15) the traces on the right-hand-side of Equation (4.19) are well convergent, both in the infrared and in the ultraviolet. By virtue of the factor  $k\partial_k\widehat{\mathcal{R}}_k$ , the dominant contributions come from a narrow band of generalized momenta centered around  $k$ . Large momenta are exponentially suppressed.

The conceptual status and the use of the gravitational FRGE (4.20) is the same in the scalar case discussed in Section 2.2. Its perturbative expansion should reproduce the traditional non-renormalizable cutoff dependencies starting from two-loops. In the context of the asymptotic safety scenario the hypothesis at stake is that in an exact treatment of the equation the cutoff dependencies entering through the initial data get reshuffled in a way compatible with asymptotic safety. The Criterion (FRGC1) for the existence of a genuine continuum limit discussed in Section 2.3 also applies to Equation (4.20). In brief, provided a *global* solution of the FRGE (4.20) can be found (one which exists for all  $0 \leq k \leq \infty$ ), it can reasonably be identified with a renormalized effective average action  $\lim_{\Lambda \rightarrow \infty} \Gamma_{\Lambda, k}$  constructed by other means. The intricacies of the renormalization process have been shifted to the search for fine-tuned initial functionals for which a global solution of Equation (4.20) exists. For such a global solution  $\lim_{k \rightarrow 0} \Gamma_k$  then is the full quantum effective action and  $\lim_{k \rightarrow \infty} \Gamma_k = S_*$

is the fixed point action. As already noted in Section 2.3 the appropriate positivity requirement (FRGC2) remains to be formulated; one aspect of it concerns the choice of  $\mathcal{Z}_k$  factors in Equation (4.13) and will be discussed below.

The background gauge invariance of  $\Gamma_k$  expressed in Equation (4.16) is of great practical importance. It ensures that if the initial functional does not contain non-invariant terms, the flow implied by the above FRGE will not generate such terms. In contrast locality is not preserved of course; even if the initial functional is local the flow generates all sorts of terms, both local and nonlocal, compatible with the symmetries.

For the derivation of the flow equation it is important that the cutoff functionals in Equation (4.13) are quadratic in the fluctuation fields; only then a flow equation arises which contains only second functional derivatives of  $\Gamma_k$ , but no higher ones. For example using a cutoff operator involving the Laplace operator of the full metric  $g_{\alpha\beta} = \bar{g}_{\alpha\beta} + f_{\alpha\beta}$  would result in prohibitively complicated flow equations which could hardly be used for practical computations.

For most purposes the reduced effective average action (4.9) is sufficient and it is likewise background invariant,  $\bar{\Gamma}_k[g + \mathcal{L}_v g] = \bar{\Gamma}_k[g]$ . Unfortunately  $\bar{\Gamma}_k[\bar{g}_{\alpha\beta}]$  does not satisfy an exact FRGE, basically because it contains too little information. The actual RG evolution has to be performed at the level of the functional  $\Gamma_k[\langle g \rangle, \bar{g}, \sigma, \bar{\sigma}]$ . Only *after* the evolution one may set  $\langle g \rangle = \bar{g}$ ,  $\sigma = 0$ ,  $\bar{\sigma} = 0$ . As a result, the actual theory space of Quantum Einstein Gravity in this setting consists of functionals of all four variables,  $\langle g_{\alpha\beta} \rangle, \bar{g}_{\alpha\beta}, \sigma^\alpha, \bar{\sigma}_\alpha$ , subject to the invariance condition (4.9). Since  $\Gamma_k^{(2)}$  involves derivatives with respect to  $\bar{f}_{\alpha\beta}$  at fixed  $\bar{g}_{\alpha\beta}$  it is clear that the evolution cannot be formulated in terms of  $\bar{\Gamma}_k$  alone.

4.  $\Gamma_k[\bar{f}, \sigma, \bar{\sigma}; \bar{g}]$  approaches for  $k \rightarrow 0$  the background effective action of Appendix B, since  $\mathcal{R}_k^{\text{grav}}, \mathcal{R}_k^{\text{gh}}$  vanish for  $k \rightarrow 0$ . The  $k \rightarrow \infty$  limit can be inferred from Equation (4.18) by the same reasoning as in the scalar case (see Appendix C). This gives

$$\lim_{k \rightarrow \infty} \Gamma_k[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] = S_{\text{tot}}[\bar{f}, \sigma, \bar{\sigma}; \bar{g}]. \quad (4.21)$$

Note that the bare initial functional  $\Gamma_k$  includes the gauge fixing and ghost actions. At the level of the functional  $\bar{\Gamma}_k[g]$  Equation (4.21) reduces to  $\lim_{k \rightarrow \infty} \bar{\Gamma}_k[g] = S[g]$ .

5. The effective average action satisfies a functional BRST Ward identity which reflects the invariance of  $S_{\text{tot}}$  under the BRST transformations

$$\begin{aligned} \delta_\epsilon f_{\alpha\beta} &= \epsilon \kappa^{-2} \mathcal{L}_C(\bar{g}_{\alpha\beta} + f_{\alpha\beta}), & \delta_\epsilon \bar{g}_{\alpha\beta} &= 0, \\ \delta_\epsilon C^\alpha &= \epsilon \kappa^{-2} C^\beta \partial_\beta C^\alpha, & \delta_\epsilon \bar{C}_\alpha &= \epsilon (\alpha\kappa)^{-1} Q_\alpha. \end{aligned} \quad (4.22)$$

Here  $\epsilon$  is an anti-commuting parameter. Since the mode cutoff action  $C_k$  is not BRST invariant, the Ward identity differs from the standard one by terms involving  $\mathcal{R}_k^{\text{grav}}, \mathcal{R}_k^{\text{gh}}$ . For the explicit form of the identity we refer to [179].

6. Initially the vertex (or 1-PI Greens) functions are given by multiple functional derivatives of  $\Gamma_k[\bar{f}, \sigma, \bar{\sigma}; \bar{g}]$  with respect to  $\bar{f}, \sigma, \bar{\sigma}$  at fixed  $\bar{g}$  and setting

$$\bar{f}_{\alpha\beta} = \langle f_{\alpha\beta} \rangle = 0, \quad \sigma^\alpha = \langle C^\alpha \rangle = 0, \quad \bar{\sigma}_\alpha = \langle \bar{C}_\alpha \rangle = 0 \quad (4.23)$$

after differentiation. The resulting multi-point functions  $\Gamma_k^{(n)}(x_1, \dots, x_n; g)$  are  $k$ -dependent functionals of the ( $k$ -independent)  $\langle g_{\alpha\beta} \rangle = \bar{g}_{\alpha\beta}$ . For extremizing sources obtained by formal series inversion the condition (4.23) automatically switches off the sources in Equation (4.7);

for the ghosts this is consistent with  $\Gamma_k$  having ghost number zero. Note that the one-point function  $\Gamma_k^{(1)}(x_1; g) \equiv 0$  vanishes identically.

An equivalent set of vertex functions should in analogy to the Yang–Mills case [68, 42, 67] be obtained by differentiating  $\Gamma_k[\bar{g}, \sigma, \bar{\sigma}] := \Gamma_k[0, \sigma, \bar{\sigma}; \bar{g}] \equiv \Gamma_k[\bar{g}, \bar{g}, \sigma, \bar{\sigma}]$  with respect to  $\bar{g}$ . Specifically for  $\sigma^\alpha = \bar{\sigma}_\alpha = 0$  one gets multipoint functions  $\bar{\Gamma}_k^{(n)}(x_1, \dots, x_n)$ , with the shorthand (4.9). The solutions  $(\check{g}_k)_{\alpha\beta}$  of

$$\frac{\delta \bar{\Gamma}_k}{\delta g_{\alpha\beta}}[\check{g}_k] = 0 \quad (4.24)$$

play an important role in the interpretation of the formalism (see Section 4.2).

The precise physics significance of the multipoint functions  $\Gamma_k^{(n)}$  and  $\bar{\Gamma}_k^{(n)}$  remains to be understood. One would expect them to be related to S-matrix elements on a self-consistent background but, for example, an understanding of the correct infrared degrees of freedom is missing.

This concludes our summary of the key properties of the gravitational effective average action. Before turning to applications of this formalism, we discuss the significance and the proper choice of the  $\mathcal{Z}_k$  factors in Equation (4.14), which is one aspect of the positivity issue (FRGC2) of Section 2.2. The significance of these factors is best illustrated in the scalar case. As discussed in Appendix C in scalar theories with more than one field it is important that all fields are cut off at the same  $k^2$ . This is achieved by a cutoff function of the form (C.21) where  $\mathcal{Z}_k$  is in general a matrix in field space. In the sector of modes with inverse propagator  $Z_k^{(i)}p^2 + \dots$  the matrix  $\mathcal{Z}_k$  is diagonal with entries  $\mathcal{Z}_k = Z_k^{(i)}$ . In a scalar field theory these  $\mathcal{Z}_k$  factors are automatically positive and the flow equations in the various truncations are well-defined.

In gravity the situation may be more subtle. First, consider the case where  $\phi$  is some normal mode of  $\bar{f}_{\alpha\beta}$  and that it is an eigenfunction of  $\Gamma_k^{(2)}$  with eigenvalue  $Z_k^\phi p^2$ , where  $p^2$  is a positive eigenvalue of some covariant kinetic operator, typically of the form  $-\bar{\nabla}^2 + R$ -terms. If  $Z_k^\phi > 0$  the situation is clear, and the rule discussed in the context of scalar theories applies: One chooses  $\mathcal{Z} = Z_k^\phi$  because this guarantees that for the low momentum modes the running inverse propagator  $\Gamma_k^{(2)} + \mathcal{R}_k$  becomes  $Z_k^\phi(p^2 + k^2)$ , exactly as it should be.

More tricky is the question how  $\mathcal{Z}_k$  should be chosen if  $Z_k^\phi$  is negative. If one continues to use  $\mathcal{Z}_k = Z_k^\phi$ , the evolution equation is perfectly well defined because the inverse propagator  $-|Z_k^\phi|(p^2 + k^2)$  never vanishes, and the traces of Equation (4.19) are not suffering from any infrared problems. In fact, if we write down the perturbative expansion for the functional trace, for instance, it is clear that all propagators are correctly cut off in the infrared, and that loop momenta smaller than  $k$  are suppressed. On the other hand, if we set  $\mathcal{Z}_k = -Z_k^\phi$ , then  $-|Z_k^\phi|(p^2 - k^2)$  introduces a spurious singularity at  $p^2 = k^2$ , and the cutoff fails to make the theory infrared finite. This choice of  $\mathcal{Z}_k$  is ruled out therefore. At first sight the choice  $\mathcal{Z}_k = -Z_k^\phi$  might have appeared more natural because only if  $\mathcal{Z}_k > 0$  the factor  $\exp(-C_k) \sim \exp(-\int \phi \mathcal{R}_k \phi)$  is a damped exponential which suppresses the low momentum modes in the usual way. For the other choice  $\mathcal{Z}_k = +Z_k^\phi < 0$  the factor  $\exp\{\int |\mathcal{R}_k| \phi^2\}$  is a growing exponential instead and, at least at first sight, seems to enhance rather than suppress the infrared modes. However, as suggested by the perturbative argument, this conclusion is too naive perhaps.

In all existing RG studies using this formalism the choice

$$\mathcal{Z}_k = +Z_k^\phi, \quad \text{for either sign of } Z_k^\phi, \quad (4.25)$$

has been adopted, and there is little doubt that, within those necessarily truncated calculations, this is the correct procedure. Besides the perturbative considerations above there are various arguments of a more general nature which support Equation (4.25):

- At least formally the construction of the effective average action can be repeated for Lorentzian signature metrics. Then one deals with oscillating exponentials  $e^{iS}$ , and for arguments like the one leading to Equation (4.21) one has to employ the Riemann–Lebesgue lemma. Apart from the obvious substitutions,  $\Gamma_k \rightarrow -i\Gamma_k$ ,  $\mathcal{R}_k \rightarrow -i\mathcal{R}_k$ , the evolution equation remains unaltered. For  $\mathcal{Z}_k = Z_k^\phi$  it has all the desired features, and  $Z_k^\phi < 0$  seems not to pose any special problem, since  $\exp\{\pm i \int \phi |\mathcal{R}_k| \phi\}$  for either sign leads to an IR suppression.
- For finite  $k$  the Euclidean FRGE is perfectly well-defined even if  $\mathcal{Z}_k < 0$ , while the status of the Euclidean functional integral with its growing exponential seems problematic. In principle there exists the possibility of declaring the FRGE the primary object. If a global solution to it exists this functional might define a consistent quantum theory of gravity even though the functional integral per se does not exist. As noted in Section 3.4 the inclusion of ‘other desirable’ features might then be more difficult, though.
- It might be, and there exist indications in this direction [131, 132], that for the *exact* RG flow  $\Gamma_k^{(2)}$  is always a positive operator (one with positive spectrum) along physically relevant RG trajectories. Then  $Z_k^\phi > 0$  for all modes and the problem does not arise. If this contention is correct, the  $Z_k^\phi < 0$  phenomenon which is known to occur in certain truncations would be an artifact of the approximations made.

In fact, as we shall discuss in more detail later on, the to date best truncation used for the investigation of asymptotic safety (“ $R^2$  truncation”) has only positive  $\mathcal{Z}_k$  factors in the relevant regime. On the other hand, the simpler “Einstein–Hilbert truncation” has also negative  $\mathcal{Z}_k$ ’s. If one applies the rule (4.25) to it, the Einstein–Hilbert truncation produces almost the same results as the  $R^2$  truncations [131, 132]. This is a strong argument in favor of the  $\mathcal{Z}_k = +Z_k^\phi$  rule.

In [133, 131] a slightly more general variant of the construction described here has been employed. In order to facilitate the calculation of the functional traces in the FRGE (4.19) it is helpful to employ a transverse-traceless (TT) decomposition of the metric:  $f_{\alpha\beta} = f_{\alpha\beta}^T + \bar{\nabla}_\alpha V_\beta + \bar{\nabla}_\beta V_\alpha + \bar{\nabla}_\alpha \bar{\nabla}_\beta \sigma - d^{-1} \bar{g}_{\alpha\beta} \bar{\nabla}^2 \sigma + d^{-1} \bar{g}_{\alpha\beta} \phi = \hat{f}_{\alpha\beta} + d^{-1} \bar{g}_{\alpha\beta} \phi$ . Here  $f_{\alpha\beta}^T$  is a transverse traceless tensor,  $V_\alpha$  a transverse vector, and  $\sigma$  and  $\phi$  are scalars. In this framework it is natural to formulate the cutoff in terms of the component fields appearing in the TT decomposition  $C_k \sim \int f_{\alpha\beta}^T \mathcal{R}_k f^{T\alpha\beta} + \int V_\alpha \mathcal{R}_k V^\alpha + \dots$ . This cutoff is referred to as a cutoff of “type B”, in contradistinction to the “type A” cutoff described above,  $C_k \sim \int f_{\alpha\beta} \mathcal{R}_k f^{\alpha\beta}$ . Since covariant derivatives do not commute, the two cutoffs are not exactly equal even if they contain the same shape function  $\mathcal{R}^{(0)}$ . Thus, comparing type A and type B cutoffs is an additional possibility for checking scheme (in)dependence [133, 131].

## 4.2 Geometries at different resolution scales

In this section we elaborate on the interpretation of the effective average action formalism in a gravitational context. Specifically we argue that  $\Gamma_k$  encodes information about ‘quantum geometries’ at different resolution scales.

Recall that the effective average action  $\Gamma_k$  may be regarded as the standard effective action where the bare action has been modified by the addition of the mode cutoff term  $C_k$ . Every given (exact or truncated) renormalization group (RG) trajectory can be viewed as a collection of effective field theories  $\{\Gamma_k, 0 \leq k \leq \infty\}$ . In this sense a single fundamental theory gives rise to a double infinity of effective theories – one for each trajectory and one for each value of  $k$ . As



explained in Appendix C, the motivation for this construction is that one would like to be able to ‘read off’ part of the physics contents of the theory simply by inspecting the effective action relevant to the problem under consideration. If the problem has only one scale  $k$ , the values of the running couplings and masses in  $\Gamma_k$  may be treated approximately as classical parameters.

In a gravitational context this construction has been described in Section 4.1. The above effective field theory aspect has also been used in Equation (2.62) of Section 2.4. Here we elaborate on the significance of the solutions of the effective field equations that come with it. The two distinct versions have been discussed in Item 6 of the previous Section 4.1.1; here we consider

$$\frac{\delta \bar{\Gamma}_k[\check{g}_k]}{\delta g_{\alpha\beta}} = 0. \quad (4.26)$$

Formally these stationarity equations can be viewed as a 1-parametric family of in general nonlocal generalizations of the Einstein field equations. The principle of generating or selecting physical solutions via an underlying physically acceptable state has already been described in Section 2.3. We resume this discussion below. The solutions  $\check{g}_k$  of Equation (4.26) are  $k$ -dependent simply because the equations are. The problem of identifying ‘the same’ solution for different  $k$  can in principle be addressed by introducing an evolution equation for  $\check{g}_k$ , schematically obtained by differentiating Equation (4.26) with respect to  $k$ .

We now select a state which favors geometries that are smooth and almost flat on large scales as in Section 2.3. We can think of this state as a background dependent expectation functional  $\mathcal{O} \mapsto \langle \mathcal{O} \rangle_{\bar{g}}$  where the background has been self-consistently adjusted through the condition  $\langle g_{\alpha\beta} \rangle_{\bar{g}_*[g]} = \bar{g}_{\alpha\beta}$  (see Equation (2.48)). This switches off the source, and any fixed point of the map  $g \mapsto \bar{g}_*[g]$  gives a particular solution to Equation (4.26) at  $k = 0$ , implicitly referring to the underlying state [156].

In terms of the effective average action  $\bar{\Gamma}_k$  the state should implicitly determine a family of solutions  $\{\check{g}_k, 0 \leq k \leq \infty\}$ . Structures on a scale  $k_1$  are best described by  $\Gamma_{k_1}$ . In principle one could also use  $\Gamma_{k_2}$  with  $k_2 \neq k_1$  to describe structures at scale  $k_1$  but then a further functional integration would be needed. It is natural to think of the family  $\{\check{g}_k, 0 \leq k \leq \infty\}$  as describing aspects of a ‘‘quantum spacetime’’. By a ‘‘quantum spacetime’’ we mean a manifold equipped with infinitely many metrics; in general none of them will be a solution of the Einstein field equations. One should keep in mind however that the quantum counterpart of a classical spacetime is characterized by many more data than the metric expectation values  $\check{g}_k$  alone, in particular by all the higher functional derivatives of  $\Gamma_k$  evaluated at  $\check{g}_k$ . The second derivative for instance evaluated at  $\check{g}_k$  is the inverse graviton propagator in the background  $\check{g}_k$ . Note that all these higher multi-point functions probe aspects of the underlying quantum state.

By virtue of the effective field theory properties of  $\Gamma_k$  the interpretation of the metrics  $\check{g}_k$  is as follows. Features involving a typical scale  $k_1$  are best described by  $\Gamma_{k_1}$ . Hence  $\check{g}_{k_1}$  is the average metric detected in a (hypothetical) experiment which probes aspects of the quantum spacetimes with typical momenta  $k_1$ . In more figurative terms one can think of  $\check{g}_k$  as a ‘microscope’ whose variable ‘resolving power’ is given by the energy scale  $k$ .

This picture underlies the discussions in [135, 134] where the quantum spacetimes are viewed as fractal-like and the qualitative properties of the spectral dimension (2.53) have been derived. We refer to Section 2.4 and [135, 134] for detailed expositions. The fractal aspects here refer to the generalized ‘scale’ transformations  $k \rightarrow 2k$ , say. Moreover a scale dependent metric  $\check{g}_k$  associates a resolution dependent proper length to any ( $k$ -independent) curve. The  $k$ -dependence of this proper length can be thought of as analogous to the well-known example that the length of the coast line of England depends on the size of the yardstick used.

Usually the resolving power of a microscope is characterized by a length scale  $\ell$  defined as the smallest distance of two points the microscope can distinguish. In the above analogy between the effective average action and a ‘‘microscope’’ the resolving power is implicitly given by the mass

scale  $k$  and it is not immediate how  $k$  relates to a distance. One would like to know the minimum proper distance  $\ell(k_1)$  of two points which can be distinguished in a hypothetical experiment with a probe of momentum  $k_1$ , effectively described by the action  $\Gamma_{k_1}$ . Conversely, if one wants to ‘focus’ the microscope on structures of a given proper length  $\ell$  one must know the  $k$ -value corresponding to this particular value of  $\ell$ . For non-gravitational theories in flat Euclidean space one has  $\ell(k) \approx \pi/k$ , but in quantum gravity the relation is more complicated.

Given a family of solutions  $\{\check{g}_k, 0 \leq k \leq \infty\}$  with the above interpretation the construction of a candidate for  $\ell(k)$  proceeds as follows [135, 134, 184]: One considers the spectral problem of the (tensor) Laplacian  $-\Delta_{\check{g}_k}$  associated with  $\check{g}_k$ . To avoid technicalities inessential for the discussion we assume that all geometries in the family  $\{\check{g}_k\}$  are compact and closed. The spectrum of  $-\Delta_{\check{g}_k}$  will then be discrete; we write  $-\Delta_{\check{g}_k}\phi_n(\check{g}_k) = \mathcal{E}_n(\check{g}_k)\phi_n(\check{g}_k)$ ,  $n \in \mathbb{N}$ , for the spectral problem. As indicated, both the spectral values  $\mathcal{E}_n(\check{g}_k)$  and the eigenfunctions  $\phi_n(\check{g}_k)$  will now depend on  $k$ .

The collection of eigenfunctions

$$\{\phi_n(\check{g}_k) \text{ with } n \text{ such that } \mathcal{E}_n(\check{g}_k) \approx k^2\}, \quad (4.27)$$

will be referred to as ‘‘cutoff modes at scale  $k$ ’’. The significance of these modes can be understood by returning to the original functional integral (4.3), which one can think of being performed over the (tensorial) eigenmodes of the  $k$ -independent Laplacian  $-\Delta_{\bar{g}}$  of the background metric considered. Schematically  $\bar{\Gamma}_k[\bar{g}]$  contains information about the functional integral (4.3) where all the eigenmodes with spectral values  $\mathcal{E}_n(\bar{g})$  obeying  $\mathcal{E}_n(\bar{g}) \geq k^2$  have been integrated out; the modes with  $\mathcal{E}_n(\bar{g}) \approx k^2$  are ‘the last’ to be integrated out. If one now takes for the background metric  $\bar{g}$  the itself  $k$ -dependent solution  $\check{g}_k$ , the equation  $\mathcal{E}_n(\check{g}_k) \approx k^2$  implicitly defines  $n$  and hence selects the ‘‘cutoff modes at scale  $k$ ’’. Note that at the level of  $\bar{\Gamma}_k$  the two metrics  $\langle g_{\alpha\beta} \rangle$  and  $\bar{g}_{\alpha\beta}$  are already identified.

Given a wave function in Equation (4.27) one can now ask what a typical coordinate distance  $\Delta_k x^\alpha$  is over which it varies. Converting this into a proper distance with the metric  $\check{g}_k$  defines the proposed resolving power

$$\ell(k; x) := \left[ (\check{g}(x)_k)_{\alpha\beta} \Delta_k x^\alpha \Delta_k x^\beta \right]^{1/2}. \quad (4.28)$$

The definition (4.28) is motivated by the fact that the ‘‘last set of modes integrated out’’ should set the length scale over which the (covariant version) of the averaging has been performed. In this sense  $\ell(k; x)$  is a substitute for the ‘‘blocking size’’ in the spirit of Kadanoff–Wilson. The scale  $\ell(k; x)$  may vary with the point  $x$  on the manifold considered. A case study of the relation between  $\ell$  and  $k$  in a simple situation can be found in [184].

Since  $\bar{\Gamma}_k[g]$  depends on the choice of the mode-cutoff scheme so will the solutions of Equation (4.26), and hence the resolving power  $\ell(k; x)$ . It can thus not be identified with the resolving power of an actual experimental set up, but is only meant to provide an order of magnitude estimate. The scheme independence of the resolution which can be achieved in an actual experimental set up would in this picture arise because the scheme dependence in the trajectory cancels against that in the  $\ell$  versus  $k$  relation.

This concludes our presentation of the effective average action formalism for gravity. In the next Section 4.3 we will use the FRGE for  $\Gamma_k$  as a tool to gain insight into the gravitational renormalization flow.

### 4.3 Truncated flow equations

Approximate computations of the effective average action can be done in a variety of ways: by perturbation theory, by saddle point approximations of the functional integral, or by looking for approximate solutions of the FRGE. A nonperturbative method of the latter type consists in truncating the underlying functional space. Using an ansatz for  $\Gamma_k$  where  $k$ -independent local

or nonlocal invariants are multiplied by running parameters, the FRGE (4.19) can eventually be converted into a system of ordinary differential equations for these parameters. In this section we outline how the conversion is done in principle.

A still very general truncated functional space consists of ‘all’ background invariant functionals  $\Gamma_k[g, \bar{g}, \sigma, \bar{\sigma}]$  which neglect the evolution of the ghost action. The corresponding ansatz reads [179]

$$\Gamma_k[g, \bar{g}, \sigma, \bar{\sigma}] = \bar{\Gamma}_k[g] + \hat{\Gamma}_k[g, \bar{g}] + S_{\text{gf}}[g - \bar{g}; \bar{g}] + S_{\text{gh}}[g - \bar{g}, \sigma, \bar{\sigma}; \bar{g}]. \quad (4.29)$$

To simplify the notation we wrote  $g_{\alpha\beta}$  for  $\langle g_{\alpha\beta} \rangle$  in the argument of the effective actions as before. In Equation (4.29) we extracted the classical  $S_{\text{gf}}$  and  $S_{\text{gh}}$  from  $\Gamma_k$ . The remaining functional depends on both  $g_{\alpha\beta}$  and  $\bar{g}_{\alpha\beta}$ . It is further decomposed as  $\bar{\Gamma}_k + \hat{\Gamma}_k$  where  $\bar{\Gamma}_k$  is defined in Equation (4.9) and  $\hat{\Gamma}_k$  contains the deviations for  $\bar{g} \neq g$ . Hence, by definition,  $\hat{\Gamma}_k[g, g] = 0$ , and  $\hat{\Gamma}_k$  can be viewed as a quantum correction to the gauge fixing term which vanishes for  $\bar{g} = g$ , too. The ansatz (4.29) satisfies the initial condition (4.21) if

$$\bar{\Gamma}_{k \rightarrow \infty} = S \quad \text{and} \quad \hat{\Gamma}_{k \rightarrow \infty} = 0. \quad (4.30)$$

Inserting the ansatz (4.29) into the exact form of the RG equation one obtains an evolution equation on the truncated space of  $\Gamma[g, \bar{g}]$ 's:

$$k \partial_k \Gamma_k[g, \bar{g}] = \frac{1}{2} \text{Tr} \left[ \left( \kappa^{-2} \Gamma_k^{(2)}[g, \bar{g}] + \mathcal{R}_k[\bar{g}] \right)^{-1} k \partial_k \mathcal{R}_k[\bar{g}] \right] - \text{Tr} \left[ \left( -\mathcal{M}[g, \bar{g}] + \mathcal{R}_k^{\text{gh}}[\bar{g}] \right)^{-1} k \partial_k \mathcal{R}_k^{\text{gh}}[\bar{g}] \right]. \quad (4.31)$$

This equation involves

$$\Gamma_k[g, \bar{g}] := \bar{\Gamma}_k[g] + S_{\text{gf}}[g - \bar{g}; \bar{g}] + \hat{\Gamma}_k[g, \bar{g}] \quad (4.32)$$

and  $\Gamma_k^{(2)}$ , the Hessian of  $\Gamma_k[g, \bar{g}]$  with respect to  $g_{\alpha\beta}$  at fixed  $\bar{g}_{\alpha\beta}$ .

The truncation ansatz (4.29) is still too general for practical calculations to be easily possible. Computations simplify considerably with the choice  $\hat{\Gamma}_k = (Z_{Nk} - 1)S_{\text{gf}}$  and a local curvature polynomial for  $\bar{\Gamma}_k[g]$ . The first specialization includes in  $\hat{\Gamma}_k$  only the wave function renormalization and gives

$$\Gamma_k[g, \bar{g}] = \bar{\Gamma}_k[g] + \frac{Z_{Nk}}{2\alpha} \int dx \sqrt{\bar{g}} \bar{g}^{\alpha\beta} Q_\alpha Q_\beta. \quad (4.33)$$

Here  $Q_\alpha = Q_\alpha[\bar{g}, \bar{f}]$  is given by Equation (4.10) with  $\bar{f}_{\alpha\beta} := g_{\alpha\beta} - \bar{g}_{\alpha\beta}$  replacing  $f_{\alpha\beta}$ ; it vanishes at  $g = \bar{g}$ .

For  $\bar{\Gamma}_k[g]$  two choices of local curvature polynomials have been considered in detail

$$\bar{\Gamma}_k[g] = 2\kappa^2 Z_{Nk} \int dx \sqrt{\bar{g}} [-R(g) + 2\bar{\lambda}_k], \quad (4.34)$$

$$\bar{\Gamma}_k[g] = \int dx \sqrt{\bar{g}} \{ 2\kappa^2 Z_{Nk} [-R(g) + 2\bar{\lambda}_k] + \bar{\nu}_k R^2(g) \}. \quad (4.35)$$

The truncation (4.34) will be called the *Einstein-Hilbert truncation*. It retains only the terms already present in the classical action, with  $k$ -dependent coefficients, though. It is the one where the RG flow has been originally found in [179]. The parameter  $\alpha$  in Equation (4.33) is kept constant ( $\alpha = 1$  specifically; see the discussion below for generalizations) so that in this case the truncation subspace is two-dimensional: The ansatz (4.33, 4.34) contains two free functions of the scale, the running cosmological constant  $\bar{\lambda}_k$  and  $Z_{Nk}$  or, equivalently, the running Newton constant  $g_N(k) = (2\kappa^2 Z_{Nk})^{-1}$ . Here  $\kappa^2$  is a scale independent constant related to the fixed Newton constant  $G$  in Section 1.5 by  $\kappa = (32\pi G)^{-1/2}$ .

The truncation (4.35) will be called the *R<sup>2</sup>-truncation*. It likewise keeps the gauge fixing and ghost sector classical as in Equation (4.29) but includes a local curvature squared term in  $\bar{\Gamma}_k$ . In

this case the truncated theory space is three-dimensional. Its natural (dimensionless) coordinates are  $(g, \lambda, \nu)$ , where

$$g_k := k^{d-2} \frac{g_N(k)}{16\pi}, \quad \lambda_k := k^{-2} \bar{\lambda}_k, \quad \nu_k := k^{4-d} \bar{\nu}_k. \quad (4.36)$$

Even though Equation (4.35) contains only one additional invariant, the derivation of the corresponding differential equations is far more complicated than in the Einstein–Hilbert case. We shall summarize the results obtained with Equation (4.35) [133, 131, 132] in Section 4.4.

We should mention that apart from familiarity and the retroactive justification through the results described later on, there is no structural reason to single out the truncations (4.34, 4.35). Even the truncated coarse graining flow in Equation (4.31) will generate all sorts of terms in  $\bar{\Gamma}_k[g]$ , the only constraint comes from general covariance. Both local and nonlocal terms are induced. The local invariants contain monomials built from curvature tensors and their covariant derivatives, with any number of tensors and derivatives and of all possible index structures. The form of typical nonlocal terms can be motivated from a perturbative computation of  $\Gamma_k$ ; an example is  $\int d^4x \sqrt{g} R_{\alpha\beta\gamma\delta} \ln(-\nabla^2) R^{\alpha\beta\gamma\delta}$ . Since  $\Gamma_k$  approaches the ordinary effective action  $\Gamma$  for  $k \rightarrow 0$  it is clear that such terms must be generated by the flow since they are known to be present in  $\Gamma$ . For an investigation of the non-ultraviolet properties of the theory, the inclusion of such terms is very desirable but it is still beyond the calculational state of the art (see however [180]).

The main technical complication comes from evaluating the functional trace on the right-hand-side of the flow equation (4.31) to the extent that one can match the terms against those occurring on the left-hand-side. We shall now illustrate this procedure and its difficulties in the case of the Einstein–Hilbert truncation (4.34) in more detail.

Upon inserting the ansatz (4.33) into the partially truncated flow equation (4.31) it should eventually give rise to a system of two ordinary differential equations for  $Z_{Nk}$  and  $\bar{\lambda}_k$ . Even in this simple case their derivation is rather technical, so we shall focus on matters of principle here. In order to find  $k\partial_k Z_{Nk}$  and  $k\partial_k \bar{\lambda}_k$  it is sufficient to consider (4.31) for  $g_{\alpha\beta} = \bar{g}_{\alpha\beta}$ . In this case the left-hand-side of the flow equation becomes  $2\kappa^2 \int dx \sqrt{g} [-R(g) k\partial_k Z_{Nk} + 2k\partial_k (Z_{Nk} \bar{\lambda}_k)]$ . The right-hand-side contains the functional derivatives of  $\Gamma^{(2)}$ ; in their evaluation one must keep in mind that the identification  $g_{\alpha\beta} = \bar{g}_{\alpha\beta}$  can be used only *after* the differentiation has been performed at fixed  $\bar{g}_{\alpha\beta}$ . Upon evaluation of the functional trace the right-hand-side should then admit an expansion in terms of invariants  $P_\alpha[g]$ , among them  $\int \sqrt{g}$  and  $\int \sqrt{g} R(g)$ . The projected flow equations are obtained by extracting the  $k$ -dependent coefficients of these two terms and discarding all others. Equating the result to the left-hand-side and comparing the coefficients of  $\int \sqrt{g}$  and  $\int \sqrt{g} R$ , the desired pair of coupled differential equations for  $Z_{Nk}$  and  $\bar{\lambda}_k$  is obtained.

In principle the isolation of the relevant coefficients in the functional trace on the right-hand-side can be done without ever considering any specific metric  $g_{\alpha\beta} = \bar{g}_{\alpha\beta}$ . Known techniques like the derivative expansion and heat kernel asymptotics could be used for this purpose, but their application is extremely tedious usually. For example, because the operators  $\Gamma_k^{(2)}$  and  $\mathcal{R}_k, \mathcal{R}_k^{\text{gh}}$  are typically of a complicated non-standard type so that no efficient use of the tabulated Seeley–deWitt coefficients can be made. Fortunately all that is needed to extract the desired coefficients is to get an unambiguous signal for the invariants they multiply on a suitable *subclass* of geometries  $g = \bar{g}$ . The subclass of geometries should be large enough to allow one to disentangle the invariants retained and small enough to really simplify the calculation.

For the Einstein–Hilbert truncation the most efficient choice is a family of  $d$ -spheres  $S^d(r)$ , labeled by their radius  $r$ . For those geometries  $\nabla_\rho R_{\alpha\beta\gamma\delta} = 0$ , so they give a vanishing value on all invariants constructed from  $g = \bar{g}$  containing covariant derivatives acting on curvature tensors. What remains (among the local invariants) are terms of the form  $\int \sqrt{g} P(R)$ , where  $P$  is a polynomial in the Riemann tensor with arbitrary index contractions. To linear order in the (contractions of the) Riemann tensor the two invariants relevant for the Einstein–Hilbert truncation

are discriminated by the  $S^d(r)$  metrics as they scale differently with the radius of the sphere:  $\int \sqrt{g} \sim r^d$ ,  $\int \sqrt{g}R(g) \sim r^{d-2}$ . Thus, in order to compute the beta functions of  $\bar{\lambda}_k$  and  $Z_{Nk}$  it is sufficient to insert an  $S^d(r)$  metric with arbitrary  $r$  and to compare the coefficients of  $r^d$  and  $r^{d-2}$ . If one wants to do better and include the three quadratic invariants  $\int R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta}$ ,  $\int R_{\alpha\beta}R^{\alpha\beta}$ , and  $\int R^2$ , the family  $S^d(r)$  is not general enough to separate them; all scale like  $r^{d-4}$  with the radius.

Under the trace we need the operator  $\Gamma_k^{(2)}[g, \bar{g}]$ , the Hessian of  $\Gamma_k[g, \bar{g}]$  at fixed  $\bar{g}$ . It is calculated by Taylor expanding the truncation ansatz,  $\Gamma_k[\bar{g} + \bar{f}, \bar{g}] = \Gamma_k[\bar{g}, \bar{g}] + O(\bar{f}) + \Gamma_k^{\text{quad}}[\bar{f}; \bar{g}] + O(\bar{f}^3)$ , and stripping off the two  $\bar{f}$ 's from the quadratic term,  $\Gamma_k^{\text{quad}} = \frac{1}{2} \int \bar{f} \Gamma_k^{(2)} \bar{f}$ . For  $\bar{g}_{\alpha\beta}$  a metric on  $S^d(r)$  one obtains

$$\Gamma_k^{\text{quad}}[\bar{f}; \bar{g}] = \frac{1}{2} Z_{Nk} \kappa^2 \int dx \left\{ \widehat{f}_{\alpha\beta} [-\bar{\nabla}^2 - 2\bar{\lambda}_k + C_T \bar{R}] \widehat{f}^{\alpha\beta} - \left( \frac{d-2}{2d} \right) \phi [-\bar{\nabla}^2 - 2\bar{\lambda}_k + C_S \bar{R}] \phi \right\}, \quad (4.37)$$

with

$$C_T := \frac{d(d-3)+4}{d(d-1)}, \quad C_S := \frac{d-4}{d}. \quad (4.38)$$

In order to partially diagonalize this quadratic form,  $\widehat{f}_{\alpha\beta}$  has been decomposed into a traceless part  $\widehat{f}_{\alpha\beta}$  and the trace part proportional to  $\phi$ :  $\widehat{f}_{\alpha\beta} = \widehat{f}_{\alpha\beta} + d^{-1} \bar{g}_{\alpha\beta} \phi$ ,  $\bar{g}^{\alpha\beta} \widehat{f}_{\alpha\beta} = 0$ . Further,  $\bar{\nabla}^2 = \bar{g}^{\alpha\beta} \bar{\nabla}_\alpha \bar{\nabla}_\beta$  is the Laplace operator of the background geometry, and  $\bar{R} = d(d-1)/r^2$  is the numerical value of the curvature scalar on  $S^d(r)$ .

At this point we can fix the coefficients  $\mathcal{Z}_k$  which appear in the cutoff operators  $\mathcal{R}_k$  and  $\mathcal{R}_k^{\text{gh}}$  of Equation (4.14). They should be adjusted in such a way that for every low-momentum mode the cutoff combines with the kinetic term of this mode to  $-\bar{\nabla}^2 + k^2$  times a constant. Looking at Equation (4.34, 4.35) we see that the respective kinetic terms for  $\widehat{f}_{\alpha\beta}$  and  $\phi$  differ by a factor of  $-(d-2)/2d$ . This suggests the following choice:

$$\mathcal{Z}_k^{\alpha\beta\gamma\delta} = \left[ (I - P_\phi)^{\alpha\beta\gamma\delta} - \frac{d-2}{2d} P_\phi^{\alpha\beta\gamma\delta} \right] Z_{Nk}. \quad (4.39)$$

Here  $(P_\phi)_{\alpha\beta}{}^{\gamma\delta} = d^{-1} \bar{g}_{\alpha\beta} \bar{g}^{\gamma\delta}$  is the projector on the trace part of the metric. For the traceless tensor (4.39) gives  $\mathcal{Z}_k = Z_{Nk}$ , and for  $\phi$  the different relative normalization is taken into account. Thus we obtain in the  $\widehat{f}$  and the  $\phi$ -sector, respectively:

$$\begin{aligned} \left( \kappa^{-2} \Gamma_k^{(2)}[g, g] + \mathcal{R}_k \right)_{\widehat{f}\widehat{f}} &= Z_{Nk} [-\nabla^2 + k^2 \mathcal{R}^{(0)}(-\nabla^2/k^2) - 2\bar{\lambda}_k + C_T R], \\ \left( \kappa^{-2} \Gamma_k^{(2)}[g, g] + \mathcal{R}_k \right)_{\phi\phi} &= -\frac{d-2}{2d} Z_{Nk} [-\nabla^2 + k^2 \mathcal{R}^{(0)}(-\nabla^2/k^2) - 2\bar{\lambda}_k + C_S R]. \end{aligned} \quad (4.40)$$

From now on we may set  $\bar{g} = g$  and for simplicity we have omitted the bars from the metric and the curvature. Since we did not take into account any renormalization effects in the ghost action we set  $Z_k^{\text{gh}} \equiv 1$  in  $\mathcal{R}_k^{\text{gh}}$  and obtain similarly, with  $C_V = -1/d$ ,

$$-\mathcal{M} + \mathcal{R}_k^{\text{gh}} = -\nabla^2 + k^2 \mathcal{R}^{(0)}(-\nabla^2/k^2) + C_V R. \quad (4.41)$$

Looking at Equation (4.37) we see that for  $d > 2$  the trace  $\phi$  has a ‘‘wrong sign’’ kinetic term which corresponds to a normalization factor  $Z_k^\phi < 0$ . The choice (4.39) complies with the rule (4.25) motivated earlier. As a result,  $\mathcal{Z}_k < 0$  in the  $\phi$ -sector. The negative  $Z_k^\phi$  is a reflection of the notorious conformal factor instability in the Einstein–Hilbert action.

At this point the operator under the first trace on the right-hand-side of Equation (4.31) has become block diagonal, with the  $\widehat{f}\widehat{f}$  and  $\phi\phi$  blocks given by Equation (4.40). Both block operators

are expressible in terms of the Laplacian  $\nabla^2$ , in the former case acting on traceless symmetric tensor fields, in the latter on scalars. The second trace in Equation (4.31) stems from the ghosts; it contains (4.41) with  $\nabla^2$  acting on vector fields.

It is now a matter of straightforward algebra to compute the first two terms in the derivative expansion of those traces, proportional to  $\int dx \sqrt{g} \sim r^d$  and  $\int dx \sqrt{g} R(g) \sim r^{d-2}$ . Considering the trace of an arbitrary function of the Laplacian,  $W(-\nabla^2)$ , the expansion up to second order derivatives of the metric is given by

$$\text{Tr}[W(-\nabla^2)] = (4\pi)^{-d/2} \text{tr}(I) \left\{ Q_{d/2}[W] \int d^d x \sqrt{g} + \frac{1}{6} Q_{d/2-1}[W] \int dx \sqrt{g} R(g) + O(R^2) \right\}. \quad (4.42)$$

The  $Q_n$ 's are defined as

$$Q_n[W] = \frac{1}{\Gamma(n)} \int_0^\infty dz z^{n-1} W(z) \quad (4.43)$$

for  $n > 0$ , and  $Q_0[W] = W(0)$  for  $n = 0$ . The trace  $\text{tr}(I)$  counts the number of independent field components. It equals 1,  $d$ , and  $(d-1)(d+2)/2$ , for scalars, vectors, and traceless tensors, respectively. The expansion (4.42) is derived using the heat kernel expansion

$$\text{Tr}[\exp(-is\nabla^2)] = \left( \frac{i}{4\pi s} \right)^{d/2} \text{tr}(I) \int dx \sqrt{g} \left\{ 1 - \frac{1}{6} is R(g) + O(R^2) \right\}, \quad (4.44)$$

and Mellin transform techniques [179]. Using Equation (4.42) it is easy to calculate the traces in Equation (4.31) and to obtain the RG equations in the form  $\partial_t Z_{Nk} = \dots$  and  $\partial_t (Z_{Nk} \bar{\lambda}_k) = \dots$ . We shall not display them here since it is more convenient to rewrite them in terms of the dimensionless parameters (4.36).

In terms of the dimensionless couplings  $g$  and  $\lambda$  the RG equations become a system of autonomous differential equations

$$\begin{aligned} k\partial_k g_k &= [d-2+\eta_N] g_k =: \beta_g(g_k, \lambda_k), \\ k\partial_k \lambda_k &= \beta_\lambda(g_k, \lambda_k). \end{aligned} \quad (4.45)$$

Here  $\eta_N := -k\partial_k \ln Z_{Nk}$ , the anomalous dimension of the operator  $\sqrt{g}R(g)$ , is explicitly given by

$$\eta_N = \frac{g_k B_1(\lambda_k)}{1 - g_k B_2(\lambda_k)}, \quad (4.46)$$

with the following functions of  $\lambda_k$ :

$$B_1(\lambda_k) := \frac{1}{3} (4\pi)^{1-d/2} \left[ d(d+1) \Phi_{d/2-1}^1(-2\lambda_k) - 6d(d-1) \Phi_{d/2}^2(-2\lambda_k) - 4d \Phi_{d/2-1}^1(0) - 24 \Phi_{d/2}^2(0) \right], \quad (4.47)$$

$$B_2(\lambda_k) := -\frac{1}{6} (4\pi)^{1-d/2} \left[ d(d+1) \tilde{\Phi}_{d/2-1}^1(-2\lambda_k) - 6d(d-1) \tilde{\Phi}_{d/2}^2(-2\lambda_k) \right].$$

The beta function for  $\lambda$  is given by

$$\beta_\lambda = -(2 - \eta_N) \lambda_k + \frac{1}{2} g_k (4\pi)^{1-d/2} \left[ 2d(d+1) \Phi_{d/2}^1(-2\lambda_k) - 8d \Phi_{d/2}^1(0) - d(d+1) \eta_N \tilde{\Phi}_{d/2}^1(-2\lambda_k) \right]. \quad (4.48)$$

The  $\Phi$ 's and  $\tilde{\Phi}$ 's appearing in Equations (4.47, 4.48) are certain integrals involving the normalized cutoff function  $\mathcal{R}^{(0)}$ ,

$$\begin{aligned} \Phi_n^p(w) &:= \frac{1}{\Gamma(n)} \int_0^\infty dz z^{n-1} \frac{\mathcal{R}^{(0)}(z) - z \partial_z \mathcal{R}^{(0)}(z)}{[z + \mathcal{R}^{(0)}(z) + w]^p}, \\ \tilde{\Phi}_n^p(w) &:= \frac{1}{\Gamma(n)} \int_0^\infty dz z^{n-1} \frac{\mathcal{R}^{(0)}(z)}{[z + \mathcal{R}^{(0)}(z) + w]^p}, \end{aligned} \quad (4.49)$$



for positive integers  $p$ , and  $n > 0$ .

With the derivation of the system (4.45) we managed to find an approximation to a two-dimensional projection of the FRGE flow. Its properties and the domain of applicability or reliability of the Einstein–Hilbert truncation will be discussed in Section 4.4. It will turn out that there are important qualitative features of the truncated coupling flow (4.45) which are independent of the cutoff scheme, i.e. independent of the function  $\mathcal{R}^{(0)}$ . The details of the flow pattern on the other hand depend on the choice of the function  $\mathcal{R}^{(0)}$  and hence have no intrinsic significance.

By construction the normalized cutoff function  $\mathcal{R}^{(0)}(u)$ ,  $u = p^2/k^2$ , in Equation (C.21) describes the “shape” of  $\mathcal{R}_k(p^2)$  in the transition region where it interpolates between the prescribed behavior for  $p^2 \ll k^2$  and  $k^2 \gg p^2$ , respectively. It is referred to as the *shape function* therefore.

In the literature various forms of  $\mathcal{R}^{(0)}$ ’s have been employed. Easy to handle, but disadvantageous for high precision calculations is the *sharp cutoff* [181] defined by  $\mathcal{R}_k(p^2) \sim \lim_{\hat{R} \rightarrow \infty} \hat{R} \theta(1 - p^2/k^2)$ , where the limit is to be taken after the  $p^2$  integration. This cutoff allows for an evaluation of the  $\Phi$  and  $\tilde{\Phi}$  integrals in closed form. Taking  $d = 4$  as an example, Equations (4.45) boil down to the following simple system then

$$\begin{aligned} \partial_t \lambda_k &= -(2 - \eta_N) \lambda_k - \frac{g_k}{\pi} \left[ 5 \ln(1 - 2\lambda_k) - 2\zeta(3) + \frac{5}{2} \eta_N \right], \\ \partial_t g_k &= (2 + \eta_N) g_k, \\ \eta_N &= -\frac{2g_k}{6\pi + 5g_k} \left[ \frac{18}{1 - 2\lambda_k} + 5 \ln(1 - 2\lambda_k) - \zeta(2) + 6 \right]. \end{aligned} \quad (4.50)$$

(For orientation, Equation (4.50) corresponds to the sharp cutoff with  $s = 1$  in [181]). The flow described by Equation (4.50) is restricted to the halfplane  $\{(\lambda, g), -\infty < \lambda < 1/2, -\infty < g < \infty\}$  since the beta functions are singular along the boundary line  $\lambda = 1/2$ . When a trajectory hits this line it cannot reach the infrared ( $k = 0$ ) but rather terminates at a nonzero  $k_{\text{term}}$ . Within the Einstein–Hilbert truncation this happens for all trajectories which approach a positive  $\bar{\lambda}_k$  at low  $k$ .

Also the cutoff with  $\mathcal{R}^{(0)}(u) = (1 - u)\theta(1 - u)$  allows for an analytic evaluation of the integrals; it has been used in the Einstein–Hilbert truncation in [136]. In order to check the scheme (in)dependence of certain results it is desirable to perform the calculation, in one stroke, for a whole class of  $\mathcal{R}^{(0)}$ ’s. For this purpose the following one parameter family of exponential cutoffs has been used [205, 133, 131]:

$$\mathcal{R}_s^{(0)}(u) = \frac{su}{e^{su} - 1}. \quad (4.51)$$

The precise form of the cutoff is controlled by the “shape parameter”  $s$ . For  $s = 1$ , Equation (4.51) coincides with the standard exponential cutoff (4.15). The exponential cutoffs are suitable for precision calculations, but the price to be paid is that their  $\Phi$  and  $\tilde{\Phi}$  integrals can be evaluated only numerically. The same is true for a one-parameter family of shape functions with compact support which was used in [133, 131].

The form of the expression (4.46) for the anomalous dimension illustrates the nonperturbative character of the method. For  $g_k B_2(\lambda_k) < 1$  Equation (4.46) can be expanded as

$$\eta_N = g_k B_1(\lambda_k) \sum_{n \geq 0} g_k^n B_2(\lambda_k)^n, \quad (4.52)$$

which illustrates that even a simple truncation can sum up arbitrarily high powers of the couplings. It is instructive to consider the approximation where only the lowest order is retained in Equation (4.52). In terms of the dimensionful  $G_k := g_k/k^2$  one has in  $d = 4$  and for  $\lambda_k \approx 0$ ,

$\eta_N = B_1(0)G_0k^2 + O(G_0^2k^4)$ , so that integrating  $k\partial_k G_k = \eta_N G_k$  yields

$$G_k = G_0 \left[ 1 + \frac{1}{2} B_1(0) G_0 k^2 + O(G_0^2 k^4) \right]. \quad (4.53)$$

The constant  $B_1(0) := [\Phi_1^1(0) - 24\Phi_2^2(0)]/(3\pi)$  is  $\mathcal{R}^{(0)}$ -dependent via the threshold functions defined in Equation (4.49). However, for all cutoffs one finds that  $B_1(0) < 0$ . An acceptable shape function  $\mathcal{R}^{(0)}(z)$  must interpolate between  $\mathcal{R}^{(0)}(0) = 1$  and  $\mathcal{R}^{(0)}(\infty) = 0$  in a monotonic way, the ‘drop’ occurring near  $z = 1$ . This implies that  $\Phi_1^1(0)$  and  $\Phi_2^2(0)$  are both positive. Since typically they are of order unity this suggests that  $B_1(0)$  should be negative. This has been confirmed for one-parameter families of cutoffs such as Equation (4.51) or those of [133, 131], for the family of sharp cutoffs, and for the optimized cutoff. Using also numerical methods it seems impossible to find an acceptable  $\mathcal{R}^{(0)}$  which would yield a positive  $B_1(0)$ . The approximation (4.53) is valid for  $k^2 \ll G_0^{-1} = M_{\text{Pl}}^2$ . One sees that at least in this regime  $G_k$  is a *decreasing* function of  $k$ . This corresponds to the antiscreening behavior discussed in Sections 1.1 and 1.5.

Above we illustrated the general ideas and constructions underlying truncated gravitational RG flows by means of the simplest example, the Einstein–Hilbert truncation (4.34). The flow equations for the  $R^2$  truncation are likewise known in closed form but are too complicated to be displayed here. These ordinary differential equations can now be analyzed with analytical and numerical methods. Their solution reveals important evidence for asymptotic safety. Before discussing these results in Section 4.4 we comment here on two types of possible generalizations.

Concerning generalizations of the ghost sector truncation, beyond Equation (4.29) no results are available yet, but there is a partial result concerning the gauge fixing term. Even if one makes the ansatz (4.33) for  $\Gamma_k[g, \bar{g}]$  in which the gauge fixing term has the classical (or more appropriately, bare) structure one should treat its prefactor as a running coupling:  $\alpha = \alpha_k$ . After all, the actual ‘theory space’ of functionals  $\Gamma[g, \bar{g}, \sigma, \bar{\sigma}]$  contains ‘T-type’ and ‘gauge-fixing-type’ actions on a completely symmetric footing. The beta function of  $\alpha$  has not been determined yet from the FRGE, but there is a simple argument which allows us to bypass this calculation.

In nonperturbative Yang–Mills theory and in perturbative quantum gravity  $\alpha = \alpha_k = 0$  is known to be a fixed point for the  $\alpha$  evolution. The following heuristic argument suggests that the same should be true beyond perturbation theory for the functional integral defining the effective average action for gravity. In the standard functional integral the limit  $\alpha \rightarrow 0$  corresponds to a sharp implementation of the gauge fixing condition, i.e.  $\exp(-S_{\text{gf}})$  becomes proportional to  $\delta[Q_\alpha]$ . The domain of the  $\mathcal{D}f_{\alpha\beta}$  integration consists of those  $f_{\alpha\beta}$ ’s which satisfy the gauge fixing condition exactly,  $Q_\alpha = 0$ . Adding the infrared cutoff at  $k$  amounts to suppressing some of the  $f_{\alpha\beta}$  modes while retaining the others. But since all of them satisfy  $Q_\alpha = 0$ , a variation of  $k$  cannot change the domain of the  $f_{\alpha\beta}$  integration. The delta functional  $\delta[Q_\alpha]$  continues to be present for any value of  $k$  if it was there originally. As a consequence,  $\alpha$  vanishes for all  $k$ , i.e.  $\alpha = 0$  is a fixed point of the  $\alpha$  evolution [137].

In other words we can mimic the dynamical treatment of a running  $\alpha$  by setting the gauge fixing parameter to the constant value  $\alpha = 0$ . The calculation for  $\alpha = 0$  is more complicated than at  $\alpha = 1$ , but for the Einstein–Hilbert truncation the  $\alpha$ -dependence of  $\beta_g$  and  $\beta_\lambda$ , for arbitrary constant  $\alpha$ , has been found in [133]. The  $R^2$  truncations could be analyzed only in the simple  $\alpha = 1$  gauge, but the results from the Einstein–Hilbert truncation suggest the UV quantities of interest do not change much between  $\alpha = 0$  and  $\alpha = 1$  [133, 131].

Up to now we considered pure gravity. As far as the general formalism is concerned, the inclusion of matter fields is straightforward. The structure of the flow equation remains unaltered, except that now  $\Gamma_k^{(2)}$  and  $\mathcal{R}_k$  are operators on the larger space of both gravity and matter fluctuations. In practice the derivation of the projected FRG equations can be quite formidable, the main difficulty being the decoupling of the various modes (diagonalization of  $\Gamma_k^{(2)}$ ) which in most calculational schemes is necessary for the computation of the functional traces.

Various matter systems, both interacting and non-interacting (apart from their interaction with gravity) have been studied in the literature. A rather detailed analysis of the fixed point has been performed by Percacci et al. In [72, 171, 170] arbitrary multiplets of free (massless) fields with spin 0, 1/2, 1 and 3/2 were included. In [170] a fully interacting scalar theory coupled to gravity in the Einstein–Hilbert approximation was analyzed, with a local potential approximation for the scalar self-interaction. A remarkable finding is that in a linearized stability analysis the marginality of the quartic self-coupling is lifted by the quantum gravitational corrections. The coupling becomes marginally irrelevant, which may offer a new perspective on the triviality issue and the ensued bounds on the Higgs mass. Making the number of matter fields large  $O(N)$ , the matter interactions dominate at all scales and the nontrivial fixed point of the  $1/N$  expansion [216, 217, 203] is recovered [169].

#### 4.4 Einstein–Hilbert and $R^2$ truncations

In this section we review the main results obtained in the effective average action framework via the truncated flow equations of the previous Section 4.3. To facilitate comparison with the original papers we write here  $G_k = g_N/(16\pi)$  for the running Newton constant; unless stated otherwise the results refer to  $d = 4$ .

##### 4.4.1 Phase portrait of the Einstein–Hilbert truncation

In [181] the coupling flow (4.45) implied by the Einstein–Hilbert truncation has been analyzed in detail, using both analytical and numerical methods. In particular all trajectories of this system of equations have been classified, and examples have been computed numerically. The most important classes of trajectories in the phase portrait on the  $g$ – $\lambda$  plane are shown in Figure 2. The trajectories were obtained by numerically solving the system (4.50) for a sharp cutoff; using a smooth one all qualitative features remain unchanged.

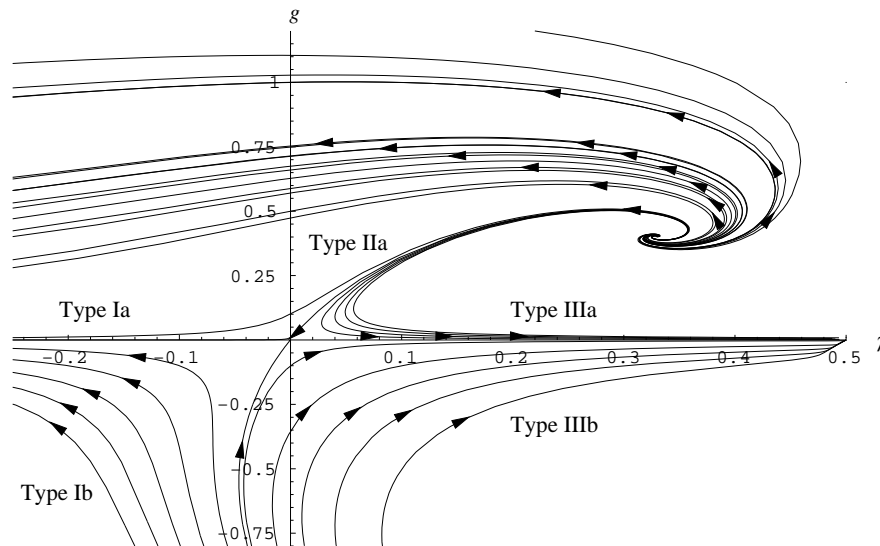


Figure 2: Flow pattern in the  $g$ – $\lambda$  plane with a sharp mode cut-off. The arrows point in the direction of the coarse graining, i.e. of decreasing  $k$ . (From [181].)

The RG flow is dominated by two fixed points  $(g_*, \lambda_*)$ : a Gaussian fixed point (GFP) at  $g_* = \lambda_* = 0$ , and a non-Gaussian fixed point (NGFP) with  $g_* > 0$  and  $\lambda_* > 0$ . There are three

classes of trajectories emanating from the NGFP: Trajectories of Type Ia and IIIa run towards negative and positive cosmological constants, respectively, and the single trajectory of Type IIa (“separatrix”) hits the GFP for  $k \rightarrow 0$ . The short-distance properties of Quantum Einstein Gravity are governed by the NGFP; for  $k \rightarrow \infty$ , in Figure 2 all RG trajectories on the half-plane  $g > 0$  run into this fixed point – its unstable manifold is two-dimensional. Note that near the NGFP the dimensionful Newton constant vanishes for  $k \rightarrow \infty$  according to  $G_k = g_k/k^2 \approx g_*/k^2 \rightarrow 0$ . The conjectured nonperturbative renormalizability of Quantum Einstein Gravity is due to this NGFP: If it was present in the untruncated RG flow it could be used to construct a microscopic quantum theory of gravity by taking the limit of infinite UV cutoff along one of the trajectories running into the NGFP, implying that the theory does not develop uncontrolled singularities at high energies [227].

The trajectories of Type IIIa cannot be continued all the way down to the infrared ( $k = 0$ ) but rather terminate at a finite scale  $k_{\text{term}} > 0$ . (This feature is not resolved in Figure 2.) At this scale the  $\beta$ -functions diverge. As a result, the flow equations cannot be integrated beyond this point. The value of  $k_{\text{term}}$  depends on the trajectory considered. The trajectory terminates when the dimensionless cosmological constant reaches the value  $\lambda = 1/2$ . This is due to the fact that the functions  $\Phi_n^p(w)$  and  $\tilde{\Phi}_n^p(w)$  – for any admissible choice of  $\mathcal{R}^{(0)}$  – have a singularity at  $w = -1$ , and because  $w = -2\lambda$  in all terms of the  $\beta$ -functions. In Equations (4.50) the divergence at  $\lambda = 1/2$  is seen explicitly. The phenomenon of terminating RG trajectories is familiar from simpler theories, such as Yang–Mills theories. It usually indicates that the truncation becomes insufficient at small  $k$ .

#### 4.4.2 Evidence for asymptotic safety – Survey

Here we collect the evidence for asymptotic safety obtained from the Einstein–Hilbert and  $R^2$  truncations, Equation (4.34) and Equation (4.35), respectively, of the flow equations in Section 4.2 [133, 131].

The details of the flow pattern depend on a number of ad-hoc choices. It is crucial that the properties of the flow which point towards the asymptotic safety scenario are robust upon alterations of these choices. This robustness of the qualitative features will be discussed in more detail below. Here let us only recapitulate the three main ingredients of the (truncated) flow equations that can be varied: The shape functions  $\mathcal{R}^{(0)}$  in Equation (4.31) can be varied, the gauge parameter  $\alpha$  in Equation (4.33) can be varied, and the vector and transversal parts in the traceless tensor modes can be treated differently (type  $A$  and  $B$  cutoffs).

Picking a specific value for the gauge parameter has a somewhat different status than the other two choices. The truncations are actually one-parameter families of truncations labelled by  $\alpha$ ; in a more refined treatment  $\alpha = \alpha_k$  would be a running parameter itself determined by the FRGE.

In practice the shape function  $\mathcal{R}^{(0)}$  was varied within the class (4.51) of exponential cutoffs and a similar one-parameter class of cutoffs with compact support [133, 131]. Changing the cutoff function  $C_k$  at fixed  $k$  may be thought of as analogous to a change of scheme in perturbation theory.

The main qualitative properties of the coupling flow can be summarized as follows:

1. *Existence of a non-Gaussian fixed point:* The NGFP exists no matter how  $\mathcal{R}^{(0)}$  and  $\alpha$  are chosen, both for type  $A$  and  $B$  cutoffs.
2. *Positive Newton constant:* While the position of the fixed point is scheme dependent (see below), all cutoffs yield *positive* values of  $g_*$  and  $\lambda_*$ . A negative  $g_*$  would have been problematic for stability reasons, but there is no mechanism in the flow equation which would exclude it on general grounds. This feature is preserved in the  $R^2$  truncation.

3. *Unstable manifold of maximal dimension:* The existence of a nontrivial unstable manifold is crucial for the asymptotic safety scenario. The fact that the unstable manifold has (for  $d = 4$ ) its maximal dimension (at least in the vicinity of the fixed point) indicates that the set of curvature invariants retained is dynamically natural. Again this is (in  $d = 4$ ) a bonus feature not built into the flow equations. It holds for both the Einstein–Hilbert and the  $R^2$  truncation.
4. *Smallness of  $R^2$  coupling:* Also with the generalized truncation the fixed point is found to exist for all admissible cutoffs. It is quite remarkable that  $\nu_*$  is always significantly smaller than  $\lambda_*$  and  $g_*$ . Within the limited precision of our calculation this means that in the three-dimensional parameter space the fixed point practically lies on the  $(\lambda, g)$ -plane with  $\nu = 0$ , i.e. on the parameter space of the pure Einstein–Hilbert truncation.

We proceeded to discuss various aspects of the evidence for asymptotic safety in more detail, namely the structure of unstable manifold and the robustness of the qualitative features of the flow. Finally we offer some comments on the full FRGE dynamics.

#### 4.4.3 Structure of the unstable manifold

This can be studied in the vicinity of the fixed point by a standard linearized stability analysis. We summarize the results for the non-Gaussian fixed point, first in the Einstein–Hilbert truncation and then in the more general  $R^2$  truncation. To set the notation recall that for a flow equation of the form  $k \partial_k g_i = \beta_i(g_1, g_2, \dots)$  the linearized flow near the fixed point is governed by the stability matrix with components  $\Theta_{ij} := \partial \beta_i / \partial g_j |_{g=g^*}$ ,

$$k \partial_k g_i(k) = \sum_j \Theta_{ij} (g_j(k) - g_j^*). \quad (4.54)$$

The general solution to this equation reads

$$g_i(k) = g_i^* + \sum_I C_I V_i^I \left( \frac{k_0}{k} \right)^{-\vartheta_I}, \quad (4.55)$$

where the  $V^I$ 's are the right-eigenvectors of  $\Theta$  with eigenvalues  $\vartheta_I$ , i.e.  $\sum_j \Theta_{ij} V_j^I = \vartheta_I V_i^I$ . The  $C_I$ 's are constants of integration,  $k_0$  is a reference scale, and decreasing  $k$  is the direction of coarse graining. Since  $\Theta$  is not symmetric in general the  $\vartheta_I$ 's are not guaranteed to be real. In principle  $\Theta$  could also be degenerate in which case the linearized analysis would only put some constraints on the structure of the unstable manifold in the vicinity of the fixed point. As a matter of fact  $\Theta$  is non-degenerate for both the Einstein–Hilbert truncation and for the  $R^2$ -truncation. In such a situation, the eigendirections with  $\text{Re } \vartheta_I > 0$  are irrelevant; they die out upon coarse graining and span the tangent space of the fixed point's stable manifold. The remaining eigendirections with  $\text{Re } \vartheta_I < 0$  are relevant perturbations which span the tangent space of the fixed point's unstable manifold. The eigenvalues  $\vartheta_I$  play a role similar to the “critical exponents” in the theory of critical phenomena. Guided by this analogy one expects them to be rather insensitive to changes in the cutoff action  $C_k$ .

As explained in Section 2.1 it is often convenient to set  $t := \ln k_0/k$  (which is to be read as  $\ln \Lambda/k - \ln k_0/\Lambda$  in the presence of an ultraviolet cutoff  $\Lambda$ ) and ask “where a coarse graining trajectory comes from” by formally sending  $t$  to  $-\infty$  (while the coarse graining flow is in the direction of increasing  $t$ ). The tangent space to the unstable manifold has its maximal dimension if *all* the essential couplings taken into account hit the fixed point as  $t$  is sent to  $-\infty$ : The fixed point is ultraviolet stable in the direction opposite to the coarse graining. This is the case for both the Einstein–Hilbert truncation and the  $R^2$  truncation, as we shall describe now in more detail.

Linearizing the flow equation (4.45) according to Equation (4.54) we obtain a pair of complex conjugate eigenvalues  $\vartheta_1 = \vartheta_2^*$  with negative real part  $\vartheta'$  and imaginary parts  $\pm\vartheta''$ . In terms of  $t = \ln(k_0/k)$  the general solution to the linearized flow equations reads

$$(\lambda_k, g_k)^T = (\lambda_*, g_*)^T + 2\{[\operatorname{Re} C \cos(\vartheta'' t) + \operatorname{Im} C \sin(\vartheta'' t)] \operatorname{Re} V + [\operatorname{Re} C \sin(\vartheta'' t) - \operatorname{Im} C \cos(\vartheta'' t)] \operatorname{Im} V\} e^{-\vartheta' t}, \quad (4.56)$$

with  $C := C_1 = (C_2)^*$  being an arbitrary complex number and  $V := V^1 = (V^2)^*$  the right-eigenvector of  $\Theta$  with eigenvalue  $\vartheta_1 = \vartheta_2^*$ . Equation (4.54) implies that, due to the positivity of  $-\vartheta'$ , all trajectories hit the fixed point as  $t$  is sent to  $-\infty$ . The nonvanishing imaginary part  $\vartheta''$  has no impact on the stability. However, it influences the shape of the trajectories which spiral into the fixed point for  $t \rightarrow -\infty$ . In summary, for any mode-cutoff employed the non-Gaussian fixed point is found to be ultraviolet attractive in both directions of the  $(\lambda, g)$ -plane.

Solving the full, nonlinear flow equations numerically [181] shows that the asymptotic scaling region where the linearization (4.56) is valid extends from  $k = \infty$  down to about  $k \approx m_{\text{Pl}}$  with the Planck mass defined as  $m_{\text{Pl}} = G_0^{-1/2}$ . Here  $m_{\text{Pl}}$  marks the lower boundary of the asymptotic scaling region. We set  $k_0 := m_{\text{Pl}}$  so that the asymptotic scaling regime extends from about  $t = 0$  to  $t = -\infty$ .

The non-Gaussian fixed point of the  $R^2$ -truncation likewise proves to be ultraviolet attractive in any of the three directions of the  $(\lambda, g, \nu)$  tangent space for all cutoffs used. The linearized flow in its vicinity is always governed by a pair of complex conjugate eigenvalues  $\vartheta_1 = \vartheta' + i\vartheta'' = \vartheta_2^*$  with  $\vartheta' < 0$ , and a real negative one  $\vartheta_3 < 0$ . The linearized solution may be expressed as

$$(\lambda_k, g_k, \nu_k)^T = (\lambda_*, g_*, \nu_*)^T + 2\{[\operatorname{Re} C \cos(\vartheta'' t) + \operatorname{Im} C \sin(\vartheta'' t)] \operatorname{Re} V + [\operatorname{Re} C \sin(\vartheta'' t) - \operatorname{Im} C \cos(\vartheta'' t)] \operatorname{Im} V\} e^{-\vartheta' t} + C_3 V^3 e^{-\vartheta_3 t}, \quad (4.57)$$

with arbitrary complex  $C := C_1 = (C_2)^*$ , arbitrary real  $C_3$ . Here  $V := V^1 = (V^2)^*$  and  $V^3$  are the right-eigenvectors of the stability matrix  $(\Theta_{ij})_{i,j \in \{\lambda, g, \nu\}}$  with eigenvalues  $\vartheta_1 = \vartheta_2^*$  and  $\vartheta_3$ , respectively. Clearly the conditions for ultraviolet stability are  $\vartheta' < 0$  and  $\vartheta_3 < 0$ . They are indeed satisfied for all cutoffs. For the exponential shape function with  $s = 1$ , for instance, we find  $-\vartheta' = 2.15$ ,  $-\vartheta'' = 3.79$ ,  $-\vartheta_3 = 28.8$ , and  $\operatorname{Re} V = (-0.164, 0.753, -0.008)^T$ ,  $\operatorname{Im} V = (0.64, 0, -0.01)^T$ ,  $V^3 = -(0.92, 0.39, 0.04)^T$ . (The vectors are normalized such that  $\|V\| = \|V^3\| = 1$ .) The trajectories (4.57) comprise three independent normal modes with amplitudes proportional to  $\operatorname{Re} C$ ,  $\operatorname{Im} C$ , and  $C_3$ , respectively. The first two are of the spiral type again, the third one is a straight line.

For any cutoff, the numerical results have several quite remarkable properties. They all indicate that, close to the non-Gaussian fixed point, the  $R^2$  flow is rather well approximated by the Einstein–Hilbert truncation:

1. The  $\nu$ -components of  $\operatorname{Re} V$  and  $\operatorname{Im} V$  are tiny. Hence these two vectors span a plane which virtually coincides with the  $(\lambda, g)$  subspace at  $\nu = 0$ , i.e. with the parameter space of the Einstein–Hilbert truncation. As a consequence, the  $\operatorname{Re} C$ - and  $\operatorname{Im} C$ -normal modes are essentially the same trajectories as the “old” normal modes already found without the  $R^2$ -term. Also the corresponding  $\vartheta'$ - and  $\vartheta''$ -values coincide within the scheme dependence.
2. The new eigenvalue  $\vartheta_3$  introduced by the  $R^2$ -term is significantly larger in modulus than  $\vartheta'$ . When a trajectory approaches the fixed point from below ( $t \rightarrow -\infty$ ), the “old” normal modes  $\propto \operatorname{Re} C, \operatorname{Im} C$  are proportional to  $\exp(-\vartheta' t)$ , but the new one is proportional to  $\exp(-\vartheta_3 t)$ , so that it decays much more quickly. For every trajectory running into the fixed point, i.e.



for every set of constants  $(\text{Re } C, \text{Im } C, C_3)$ , we find therefore that once  $-t$  is sufficiently large the trajectory lies entirely in the  $\text{Re } V$ - $\text{Im } V$  subspace, i.e. in the  $\nu = 0$ -plane practically.

Due to the large value of  $-\vartheta_3$ , the new scaling field is ‘very relevant’. However, when we start at the fixed point ( $t = -\infty$ ) and raise  $t$  it is only at the low energy(!) scale  $k \approx m_{\text{Pl}}$  ( $t \approx 0$ ) that  $\exp(-\vartheta_3 t)$  reaches unity, and only then, i.e. far away from the fixed point, the new scaling field starts growing rapidly.

3. Since the matrix  $\Theta$  is not symmetric its eigenvectors have no reason to be orthogonal. In fact, one finds that  $V^3$  lies almost in the  $\text{Re } V$ - $\text{Im } V$ -plane. For the angles between the eigenvectors given above we obtain  $\sphericalangle(\text{Re } V, \text{Im } V) = 102.3^\circ$ ,  $\sphericalangle(\text{Re } V, V^3) = 100.7^\circ$ ,  $\sphericalangle(\text{Im } V, V^3) = 156.7^\circ$ . Their sum is  $359.7^\circ$  which confirms that  $\text{Re } V$ ,  $\text{Im } V$ , and  $V^3$  are almost coplanar. Therefore as one raises  $t$  and moves away from the fixed point so that the  $V^3$  scaling field starts growing, it is again predominantly the  $\int dx \sqrt{g}$  and  $\int dx \sqrt{g} R(g)$  invariants which get excited, but not  $\int dx \sqrt{g} R(g)^2$ .

Summarizing the three points above we can say that very close to the fixed point the  $R^2$  flow seems to be essentially two-dimensional, and that this two-dimensional flow is well approximated by the coupling flow of the Einstein–Hilbert truncation. In Figure 3 we show a typical trajectory which has all three normal modes excited with equal strength ( $\text{Re } C = \text{Im } C = 1/\sqrt{2}$ ,  $C_3 = 1$ ). All its way down from  $k = \infty$  to about  $k = m_{\text{Pl}}$  it is confined to a very thin box surrounding the  $\nu = 0$ -plane.

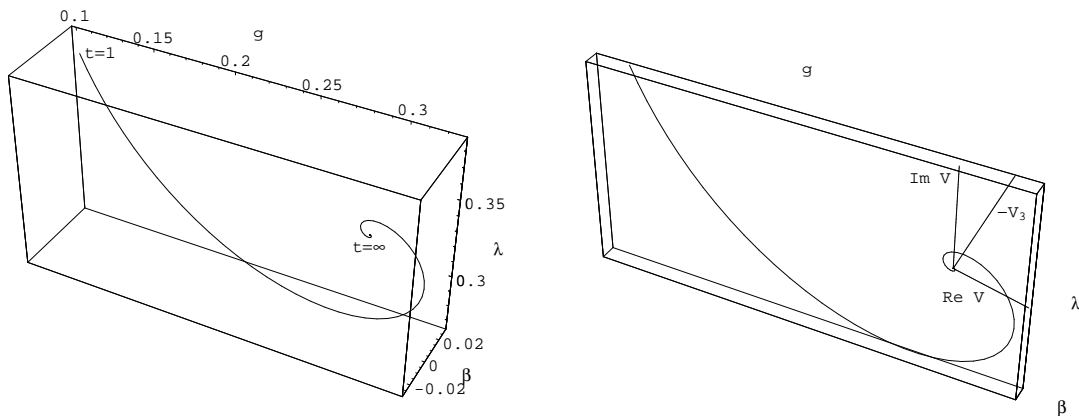


Figure 3: Trajectory of the linearized flow equation obtained from the  $R^2$ -truncation for  $-1 \geq t = \ln(k_0/k) > -\infty$ . In the right panel we depict the eigendirections and the ‘‘box’’ to which the trajectory is confined. (From [132].)

#### 4.4.4 Robustness of qualitative features

As explained before the details of the coupling flow produced by the various truncations of Equation (4.19) depend on the choice of the cutoff action ( $\mathcal{R}^{(0)}$ , type A vs. B) and the gauge parameter  $\alpha$ . Remarkably the qualitative properties of the flow, in particular those features pointing towards the asymptotic safety scenario are unchanged upon alterations of the computational scheme. Here we discuss these robustness properties in more detail. The degree of insensitivity of quantities expected to be ‘‘universal’’ can serve as a measure for the reliability of a truncation.

We begin with the very existence of a non-Gaussian fixed point. Importantly, both for type A and type B cutoffs the non-Gaussian fixed point is found to exist for all shape functions  $\mathcal{R}_s^{(0)}$ . This

generalizes earlier results in [205]. Indeed, it seems impossible to find an admissible mode-cutoff which destroys the fixed point in  $d = 4$ . This is nontrivial since in higher dimensions ( $d \gtrsim 5$ ) the fixed point exists for some but does not exist for other mode-cutoffs [181] (see however [79]).

Within the Einstein–Hilbert truncation also a RG formalism different from (and in fact much simpler than) that of the average action was used [39]. The fixed point was found to exist already in a simple RG improved 1-loop calculation with a proper time cutoff.

We take this as an indication that the fixed point seen in the Einstein–Hilbert [204, 133, 136, 39] and the  $R^2$  truncations [131] is the projection of a genuine fixed point and not just an artifact of an insufficient truncation.

Support for this interpretation comes from considering the product  $g_*\lambda_*$  of the fixed point coordinates. Recall from Section 2.3.2 that the product  $g(k)\lambda(k)$  is a dimensionless essential coupling invariant under constant rescalings of the metric [116]. One would expect that this combination is also more robust with respect to scheme changes.

In Figure 4 we show the fixed point coordinates  $(\lambda_*, g_*, \nu_*)$  for the family of shape functions (4.51) and the type B cutoff. For every shape parameter  $s$ , the values of  $\lambda_*$  and  $g_*$  are almost the same as those obtained with the Einstein–Hilbert truncation. Despite the rather strong scheme dependence of  $g_*$  and  $\lambda_*$  separately, their product has almost no visible  $s$ -dependence for not too small values of  $s$ ! For  $s = 1$ , for instance, one obtains  $(\lambda_*, g_*) = (0.348, 0.272)$  from the Einstein–Hilbert truncation and  $(\lambda_*, g_*, \nu_*) = (0.330, 0.292, 0.005)$  from the generalized truncation. One can also see that the  $R^2$  coupling  $\nu$  at the fixed point is uniformly small throughout the family of exponential shape functions (4.51).

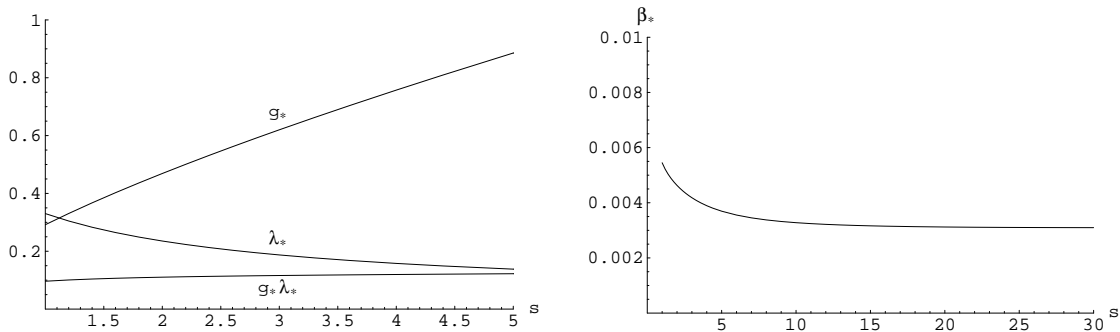


Figure 4:  $g_*$ ,  $\lambda_*$ , and  $g_*\lambda_*$  as functions of  $s$  for  $1 \leq s \leq 5$  (left panel) and  $\nu_*$  as a function of  $s$  for  $1 \leq s \leq 30$  (right panel), using the family of exponential shape functions (4.51). (From [132].)

A similar situation is found upon variation of the gauge parameter  $\alpha$ . Within the Einstein–Hilbert truncation the analysis has been performed in ref. [133] for an arbitrary constant gauge parameter  $\alpha$ , including the ‘physical’ value  $\alpha = 0$ . For example one finds

$$g_*\lambda_* \approx \begin{cases} 0.12 & \text{for } \alpha = 1, \\ 0.14 & \text{for } \alpha = 0. \end{cases} \quad (4.58)$$

The differences between the ‘physical’ (fixed point) value of the gauge parameter,  $\alpha = 0$ , and the technically more convenient  $\alpha = 1$  are at the level of about 10 to 20 per-cent. In view of this the much more involved analysis in the  $R^2$  truncation has been performed in the simpler  $\alpha = 1$  gauge only [131]. The product  $g_*\lambda_*$  with  $\alpha = 1$  is then found to differ slightly from the corresponding value in the Einstein–Hilbert truncation, however the deviation is of the same size as the difference between the  $\alpha = 0$  and the  $\alpha = 1$  results of the Einstein–Hilbert truncation. Taken together the analysis suggests the universal value  $g_*\lambda_* \approx 0.14$ .

Next we consider the  $\mathcal{R}^{(0)}$  (in)dependence of the “critical exponents”  $\vartheta'$ ,  $\vartheta''$  in Equation (4.56, 4.57). Within the Einstein–Hilbert truncation the eigenvalues are found to be reasonably constant within about a factor of 2. For  $\alpha = 1$  and  $\alpha = 0$ , for instance, they assume values in the ranges  $1.4 \lesssim -\vartheta' \lesssim 1.8$ ,  $2.3 \lesssim -\vartheta'' \lesssim 4$  and  $1.7 \lesssim -\vartheta' \lesssim 2.1$ ,  $2.5 \lesssim -\vartheta'' \lesssim 5$ , respectively. The corresponding results for the  $R^2$  truncation are shown in Figure 5. It presents the  $\mathcal{R}^{(0)}$  dependence of the critical exponents, using the family of shape functions (4.51). For the cutoffs employed  $-\vartheta'$  and  $-\vartheta''$  assume values in the ranges  $2.1 \lesssim -\vartheta' \lesssim 3.4$  and  $3.1 \lesssim -\vartheta'' \lesssim 4.3$ , respectively. While the scheme dependence of  $\vartheta''$  is weaker than in the case of the Einstein–Hilbert truncation one finds that it is slightly larger for  $\vartheta'$ . The exponent  $-\vartheta_3$  suffers from relatively strong variations as the cutoff is changed,  $8.4 \lesssim -\vartheta_3 \lesssim 28.8$ , but it is always significantly larger in modulus than  $\vartheta'$ .

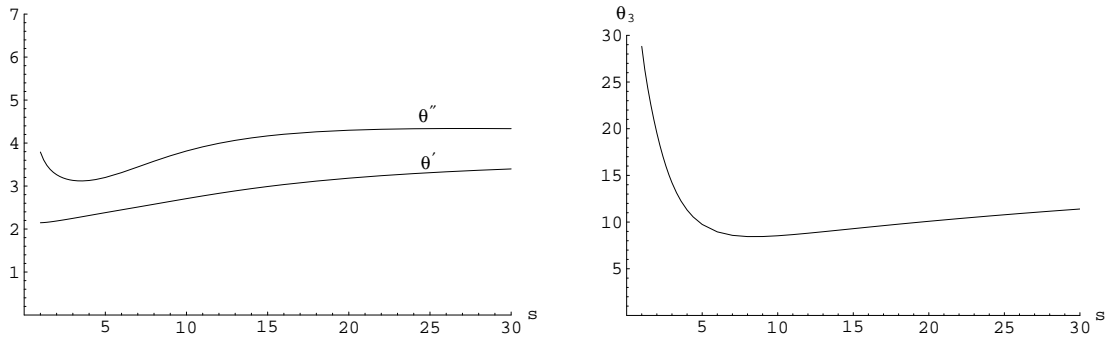


Figure 5:  $\theta' = -\text{Re } \vartheta_1$  and  $\theta'' = -\text{Im } \vartheta_1$  (left panel) and  $\theta_3 = -\vartheta_3$  (right panel) as functions of  $s$ , using the family of exponential shape functions (4.51).

In summary, the qualitative properties listed above ( $\vartheta', \vartheta_3 < 0$ ,  $-\vartheta_3 \gg -\vartheta'$ , etc.) hold for all cutoffs. The  $\vartheta$ 's have a stronger scheme dependence than  $g_* \lambda_*$ , however. This is most probably due to having neglected further relevant operators in the truncation so that the  $\Theta$  matrix we are diagonalizing is still too small.

Finally one can study the dimension dependence of these results. The beta functions produced by the truncated FRGE are continuous functions of the spacetime dimension  $d$  and it is instructive to analyze them for  $d \neq 4$ . This was done for the Einstein–Hilbert truncation in [181, 79], with the result that the coupling flow is quantitatively similar to the 4-dimensional one for not too large  $d$ . The robustness features have been explored with various cutoffs with the result that the sensitivity on the cutoff parameters increases with increasing  $d$ . In [181] a strong cutoff dependence was found for  $d$  larger than approximately 6, for two versions of the sharp cutoff (with  $s = 1, 30$ ) and for the exponential cutoff with  $s = 1$ . In [79] a number of different cutoffs were employed and no sharp increase in sensitivity to the cutoff parameters was reported for  $d \leq 10$ .

Close to  $d = 2$  the results of the  $\epsilon$ -expansion are recovered. Indeed, the fixed point of Section 1 originally found in the  $\epsilon$ -expansion is recovered in the present framework [179],

$$g_* = \frac{3}{38}\epsilon, \quad \lambda_* = -\frac{3}{38}\Phi_1^1(0)\epsilon. \quad (4.59)$$

The coefficient for  $g_*$  coincides with the one found in the  $\epsilon$ -expansion using the volume operator  $\int \sqrt{g}$  as a reference. In the expression for  $\lambda_*$ ,  $\Phi_1^1(0)$  is a scheme dependent positive constant. Of course here  $\epsilon = d - 2$  only parameterizes the dimension and does not serve double duty also as an ultraviolet regulator.

This concludes our analysis of the robustness properties of the truncated RG flow. For further details the reader is referred to Lauscher et al. [133, 131, 132]. On the basis of these robustness

properties we believe that the non-Gaussian fixed point seen in the Einstein–Hilbert and  $R^2$  truncations is very unlikely to be an artifact of the truncations. On the contrary there are good reasons to view it as the projection of a fixed point of the full FRGE dynamics. It is especially gratifying to see that within the scheme dependence the additional  $R^2$ -term has a quantitatively small impact on the location of fixed point and its unstable manifold.

In summary, we interpret the above results and their mutual consistency as quite nontrivial indications supporting the conjecture that 4-dimensional Quantum Einstein Gravity possesses a RG fixed point with precisely the properties needed for its asymptotic safety.

#### 4.4.5 Comments on the full FRGE dynamics

The generalization of the previous results to more complex truncations would be highly desirable, but for time being it is out of computational reach. We therefore add some comments on what one can reasonably expect to happen.

The key issue obviously is the dimension and the structure of the unstable manifold. For simplicity let us restrict the discussion to the ansatz (4.29, 4.33) in which the bi-metric character of the functionals and the evolution of the ghost sector are neglected. Morally speaking the following remarks should however apply equally to generic functionals  $\Gamma_k[g, \bar{g}, \sigma, \bar{\sigma}]$ . Within the restricted functional space (4.29, 4.33) only the ansatz for  $\bar{\Gamma}_k[g]$  can be successively generalized. A generic finite-dimensional truncation ansatz for  $\bar{\Gamma}_k[g]$  has the form

$$\bar{\Gamma}_k[g] = \sum_{i=0}^N g_i(k) k^{d_i} I_i[g], \quad (4.60)$$

where  $g = \langle g \rangle$  is the averaged metric and the  $I_i$  are ‘well-chosen’ local or nonlocal reparameterization invariant functionals of it.

Let us first briefly recall the scaling pattern based on the perturbative Gaussian fixed point. As described in Section 3.3 in a perturbative construction of the effective action the divergent part of the  $\ell$  loop contribution is always local and thus can be added as a counter term to a local bare action  $S[g] = \sum_i u_i P_i[g]$ , where the sum is over local curvature invariants  $P_i[g]$ . The scaling pattern of the monomials  $P_i[g]$  with respect to the perturbative Gaussian fixed point will thus reflect those of the  $I_i[g]$  in the effective action and vice versa. As explained in Section 2.3 the short-distance behavior of the perturbatively defined theory will be dominated by the  $P_i$ ’s with the largest number of derivatives acting upon  $g_{\alpha\beta}$ . In a local invariant containing the Riemann tensor to the  $p$ th power and  $q$  covariant derivatives acting on it, the number of derivatives acting on  $g_{\alpha\beta}$  is  $2p + q$ . If one starts with just a few  $P_i$ ’s and performs loop calculations one discovers that higher  $P_i$ ’s are needed as counter terms. As a consequence the high energy behavior is dominated by the bottomless chain of invariants with more and more derivatives.

As already argued in Section 2.3 in an asymptotically safe Quantum Gravidynamics the situation is different. The absence of a blow-up in the couplings is part of the defining property. The dominance of the high energy behavior by the bottomless chain of high derivative local invariants is replaced with the expectation that all invariants should be about equally important in the extreme ultraviolet.

This can be seen from the FRGE for the effective average action via the following heuristic argument. Assume that  $\bar{\Gamma}_k[g] = \sum_i u_i(k) I_i[g]$ , where the sum runs over a (dynamically determined) subset of all local and nonlocal invariants. The existence of a nontrivial fixed point means that the dimensionless couplings  $g_i(k) = k^{-d_i} u_i(k)$  approach constant values  $g_i^*$  for  $k \rightarrow \infty$ . As a consequence, the dimensionful couplings have the following  $k$ -dependence in the fixed point regime:

$$u_i(k) \approx g_i^* k^{d_i}. \quad (4.61)$$

Obviously  $u_i(k \rightarrow \infty) = 0$  for any  $d_i < 0$ . The traces on the right-hand-side of the exact flow equation (4.31) are a compact representation of the beta functions for all  $g_i$ 's. They contain the Hessian  $\bar{\Gamma}_k^{(2)}[g] = \sum_i u_i(k) I_i^{(2)}[g]$ . Let us perform the traces in the eigenbasis of  $-\nabla^2$ , denoting its eigenvalues by  $p^2$ . At least when all Hessians  $I_i^{(2)}[g]$  can be expressed in terms of  $\nabla^2$  they will be diagonal in this basis, with eigenvalues of the form  $c_i (p^2)^{-d_i/2}$ , where the constants  $c_i$ 's are of order unity. Now, as explained previously, due to the  $k\partial_k \mathcal{R}_k$  factor the functional traces receive significant contributions only from a small band of eigenvalues near  $p^2 = k^2$ . Hence  $(p^2)^{-d_i/2}$  is effectively the same as  $(k^2)^{-d_i/2}$  under the trace, and the corresponding Hessian is

$$\bar{\Gamma}_k^{(2)}(-\nabla^2 \approx k^2) \approx \sum_i u_i(k) c_i k^{-d_i}. \quad (4.62)$$

If the coefficients  $u_i(k)$  were constant, the high energy limit  $k \rightarrow \infty$  of Equation (4.62) would indeed be dominated by the higher derivative invariants, their importance growing as  $k^{-d_i}$ . However the couplings  $u_i(k)$  are  $k$ -dependent by themselves and near a NGFP they run according to Equation (4.61). As a result the growing factor  $k^{-d_i}$  is compensated by the  $k^{d_i}$  which stems from the fixed point running. Therefore  $\bar{\Gamma}_k^{(2)}$  is essentially a sum of the form  $\sum_i g_i^* c_i$  in which the higher order invariants are merely equally significant as the lower ones, and the same is true for the beta functions. The bottomless chain of higher derivative invariants is replaced on both sides of the FRGE (4.31) by quantities which could have a well-defined limit as  $k \rightarrow \infty$ .

Clearly the above argument can be generalized to action functionals depending on all the fields  $g_{\alpha\beta}, \bar{g}_{\alpha\beta}, \sigma^\alpha, \bar{\sigma}_\alpha$ . Also the choice of field variables is inessential and the argument should carry over to other types of flow equations. It suggests that there could indeed be a fixed point action  $S_*[g] = \lim_{k \rightarrow \infty} \bar{\Gamma}_k[g] = \lim_{k \rightarrow \infty} \Gamma_k[g, g, 0, 0]$  which is well-defined when expressed in terms of dimensionless quantities and which describes the extreme ultraviolet dynamics of Quantum Einstein Gravity. By construction the unstable manifold of this fixed point action would be nontrivial.

Without further insight unfortunately little can be said about its dimension. Among the *local* invariants in Equation (4.57) arguably only a finite number should be relevant. This is because the power counting dimensions  $d_i < 0$  of the infinite set of irrelevant local invariants may receive large positive corrections which makes them relevant with respect to the NGFP. An example for this phenomenon is provided by the  $I_2[g] := \int d^4x \sqrt{g} R(g)^2$  invariant. It is power counting marginal ( $d_2 = 0$ ) but with respect to the NGFP the scaling dimension of the associated dimensionless coupling is shifted to a large positive value  $d_2^{\text{NGFP}} = -\vartheta_3$  of  $O(10)$ . Nevertheless it seems implausible that this will happen to couplings with arbitrarily large negative power counting dimension as correspondingly large corrections would be required.

On the other hand this reasoning does not apply to nonlocal invariants. For example arbitrary functions  $F(I_2[g])$  of the above  $I_2[g]$  are likewise power counting marginal, and on account of a similar positive shift they too would become relevant with respect to the NGFP. While such terms would not occur in the perturbative evaluation of the effective action, in the present framework they are admissible and their importance has to be estimated by computation. As a similar an ansatz of the form

$$\bar{\Gamma}_k[g] = -\frac{1}{16\pi G} \int dx \sqrt{g} R(g) + \frac{1}{8\pi G} F_k(V[g]) \quad (4.63)$$

was considered in [180, 182]. Here  $V[g] = \int dx \sqrt{g}$  is the volume of the Riemannian manifold,  $F_k$  is an arbitrary scale dependent function, and  $G$  is the ordinary Newton constant whose evolution is neglected here. The results obtained point towards a 'quenching' of the cosmological constant similar to but more pronounced as in the mechanism of [211].

Also scalar modes like the conformal factor have vanishing power counting dimension. The way how such dimensionless scalars enter the effective action is then not constrained by the above 'implausibility' argument. An unconstrained functional occurrence however opens the door to a potentially infinite-dimensional unstable manifold.

Another core issue are of course the positivity properties of the Quantum Einstein Gravity defined through the FRGE. As already explained in Section 1.5 the notorious problems with positivity and causality which arise within standard perturbation theory around flat space in higher-derivative theories of Lorentzian gravity are not an issue in the FRG approach. For example if  $\Gamma_k$  is of the  $R + R^2$  type, the running inverse propagator  $\Gamma_k^{(2)}$  when expanded around flat space has ghosts similar to those in perturbation theory. For the FRG flow this is irrelevant, however, since in the derivation of the beta functions no background needs to be specified explicitly. All one needs is that the RG trajectories are well defined down to  $k = 0$ . This requires only that  $\Gamma_k^{(2)} + \mathcal{R}_k$  is a positive operator for all  $k$ . In the exact theory this is believed to be the case.

A rather encouraging first result in this direction comes from the  $R^2$  truncation [131]. In the FRG formalism the problem of the higher derivative ghosts is to some extent related to the negative  $Z_k^\phi$  factors discussed in Section 2.1. It was found that, contrary to the Einstein–Hilbert truncation, the  $R^2$  truncation has only positive  $Z_k^\phi$  factors in the fixed point regime  $k \geq m_{\text{Pl}}$ . Hence in this truncation the existence of ‘safe’ couplings appears to be compatible with the absence of unphysical propagating modes, as required by the scenario.



## 5 Conclusions

The goal of this review would be reached if a reasonably convincing case for an unorthodox scenario has been made. Future work will have to focus on four areas:

1. Consolidating the existence of a non-Gaussian fixed point and that of an asymptotically safe coupling flow. This may be done in various formalisms, field variables, and approximations.
2. Clarifying the microstructure of the geometries, identification of the antiscreening degrees of freedom, and the role of the ultraviolet cutoff.
3. Clarifying the physically adequate notion of unitarity and its interplay with Areas 1 and 2.
4. Characterization of generic observables and working out sound consequences for the macro-physics.

## 6 Acknowledgements

We wish to thank E. Seiler for a critical reading of the manuscript and for suggesting various improvements on the field theoretical parts. M.N. would like to thank C. Bervellier, D. Litim, R. Loll, and A. Niemi for comments and correspondence. M.R. wishes to thank A. Bonnano, O. Lauscher, F. Saueressig, J. Schwindt and H. Weyer for fruitful collaborations on the work summarized here, and R. Percacci and C. Wetterich for many discussions.

## A Reminder on Kadanoff–Wilson Renormalization

The modern view of renormalization has been shaped by Kadanoff and Wilson. See [114] and [234, 231, 232, 233] for first hand accounts and a guide to the original articles. In the present context the relevance of a Kadanoff–Wilson view on renormalization is two-fold: First it allows one to formulate the notion of renormalizability without reference to perturbation theory, and second it allows one to treat at least in principle renormalizable and non-renormalizable theories on the same footing. For convenience we briefly summarize the main principles of the Kadanoff–Wilson approach to renormalization here:

### Kadanoff–Wilson view on renormalization – Main principles:

1. A theory is not defined in terms of a given action, but in terms of a field content and the Steps 2–6 below.
2. The functional integral is performed in piecemeal, integrating out fast modes, retaining slow modes, while keeping the values of observables fixed. This “coarse graining” process results in a flow in the space of actions which depends on the chosen coarse graining operation.
3. Starting from a retroactively justified initial action ideally *all* interaction monomials generated by the flow are included in a typical action; in any case many more than just the power-counting renormalizable ones. Then one classifies the coefficients of the monomials into essential (couplings) and inessential (field redefinitions).
4. A fixed point (FP) in the flow of couplings is searched for. The position of the FP depends on the chosen coarse graining operation, but the rates of approach to it typically do not (“universality”).
5. The flow itself decides which monomials are relevant in the vicinity of a FP and hence defines the dynamics. The scaling dimensions with respect to a non-Gaussian FP may be different from (corrected) power-counting dimensions referring to the Gaussian FP.
6. The dimension of the unstable manifold and hence the “degree” of renormalizability depends on the FP!

We first add some general remarks and then elaborate on the Points 1–6.

The more familiar perturbative notion of renormalizability is neither sufficient (e.g.  $\Phi_4$  theory in  $d = 4$ ) nor necessary (e.g. Gross–Neveu model in  $d = 3$ ) for renormalizability in the above sense.

As summarized here, these principles describe the construction of a so-called massive continuum limit of a statistical field theory initially formulated on a lattice, say. A brief reminder: In a lattice field theory there is typically a dynamically generated scale, the correlation length  $\xi$ , which allows one to convert lattice distances into a physical length scale, such that say,  $\xi$  lattice spacings equal 1 fm. The lattice points  $n = (n_1, \dots, n_d) \in \mathbb{Z}^d$  are then traded for dimensionful distances  $x_i = (n_i/\xi)$  fm. Taking the lattice spacing to zero amounts to sending  $\xi$  to infinity while keeping  $x_i$  fixed. If the correlation functions of some lattice fields are rescaled accordingly (including a ‘wave function’ renormalization factor) and the limit exists, this defines a massive continuum limit of the lattice theory.

Let us now elaborate on the various points. The comments are of a generic nature, whenever a formula is needed to make the point, we consider the case of a scalar field theory on a  $d$ -dimensional Euclidean lattice with lattice spacing  $1/\Lambda$  and  $d \geq 2$ . Then  $\chi_x$  denotes the scalar field multiplet at point  $x$  and  $\chi_p$  is its Fourier transform. We freely combine results and viewpoints from the following reviews [102, 21, 77, 146, 111].

1. This will become clear from Point 5 below.

2. The rationale for the piecemeal performance of the functional integral is that in statistical mechanics language a critical problem is decomposed into a sequence of subcritical ones. Here a critical problem is one where fluctuations of the dynamical variables over vastly different length scales have to be taken into account; for a subcritical problem the opposite is true. In more detail, let  $\mathcal{O}$  be a function of the fields  $\chi$  whose functional average is meant to be a macroscopic observable, but whose statistical average is sensitive to fluctuations of the microscopic fields  $\chi$  on very different length scales. The replacement by a sequence of subcritical problems is done by specifying a “blocking” kernel  $K : \text{Configurations} \times \text{Configurations} \rightarrow \mathbb{R}$ ,  $K(\chi', \chi) = K(\{\chi'_p\}, \{\chi_p\})$ , such that

- (a)  $K(\chi', \chi) = K_{l,\delta l}(\chi', \chi)$  has support mostly on configurations  $\{\chi_p\}$  with  $l - \delta l \leq p \leq l$ .
- (b)  $\int \prod_{p \leq \Lambda} d\chi_p K(\chi', \chi) = 1$ .

Then

$$\langle \mathcal{O} \rangle = \int \prod_{p \leq \Lambda} d\chi_p \mathcal{O}(\chi) e^{-S[\chi]} = \int \prod_{p \leq \Lambda - \delta l} d\chi'_p \mathcal{O}'(\chi') e^{-S'[\chi']}, \quad (\text{A.1})$$

with

$$\mathcal{O}'(\chi') e^{-S'[\chi']} = \int \prod_{p \leq \Lambda} d\chi_p K_{\Lambda, \delta l}(\chi', \chi) \mathcal{O}(\chi) e^{-S[\chi]}. \quad (\text{A.2})$$

Taking  $\mathcal{O} = 1$  defines the coarse grained action functional  $S'$ , after which Equation (A.2) can be used to define the coarse grained observables  $\mathcal{O}'$ . Property 1 entails that only field configurations with a similar ‘degree of roughness’ have to be considered in evaluating the functional integral in Equation (A.2). It should thus be much more amenable to (numerical or analytical) approximation techniques than the original functional integral (A.1).

Once Equation (A.2) has been evaluated one can iterate the procedure. The formulas (A.1, A.2) remain valid with the basic kernel  $K$  replaced by its  $n$ -fold convolution product, for which we write  $K_{\Lambda, n\delta l}(\chi', \chi)$ . For most choices a kernel  $K_{l,\delta l}$  will not be reproducing, i.e.  $\int \prod_{p \leq \Lambda} d\chi'_p K_{l-\delta l, \delta l}(\chi'', \chi') \times K_{l,\delta l}(\chi', \chi) =: K_{l-2\delta l, 2\delta l}(\chi'', \chi)$  will not (despite the suggestive notation) coincide with the original kernel  $K_{l,\delta l}$ , just with modified parameters. Technically it is thus easier to specify the iterated kernel directly, which is of course still normalized. The  $n$ -fold iterated kernel will have support mostly on configurations with  $e^{-t} := \frac{l}{\Lambda} \leq \frac{p}{\Lambda} \leq 1$ , if  $l = n\delta l$ , and  $e^{-t}$  is the fraction of the momentum modes over which the functional integral has been performed after  $n$  iterations. In the above terminology the critical problem (A.1) has been replaced by the sequence of subcritical problems (A.2). In each iteration, referred to as a *coarse graining* step defined by the kernel  $K$ , only a small fraction of the degrees of freedom is integrated out. The action  $S = S_\Lambda$  at the cutoff scale  $p = \Lambda$  is called the *microscopic (or bare) action*, the  $S' = S_l$  reached after integrating out the ‘fast’ modes in the range  $l/\Lambda \leq p/\Lambda \leq 1$  is called the *coarse grained action* at scale  $l$ , and similarly for the fields  $\chi' = \chi_l$ . Note that the action  $S_l[\cdot]$  as a functional is defined for all field configurations, though for the evaluation of Equation (A.1) only  $S_l[\chi_l]$  is needed.

Throughout we shall follow the sloppy field theory convention that the coarse graining operates on the action. Of course what really gets updated is the functional measure

$$d\mu_l[\chi] = \prod_p d\chi_p e^{-S_l[\chi]}. \quad (\text{A.3})$$

In the (lattice) regularized theory the decomposition of the measure into a flat reference measure  $\prod_p d\chi_p$  and a Boltzmann factor parameterized by the action is unproblematic. The flow in the measures can thus be traded for a flow in the actions (as long as the Jacobian is

taken into account that comes from the reference measure upon a change of field variables  $\chi \mapsto \chi'(\chi)$ . The Wilsonian “space of actions” refers to a cone of positive measures (A.3) which is preserved under the coarse graining operation considered.

Using Equation (A.2) for the kernel  $K_{l,\delta l}(\chi, \chi') = \delta(\chi, \chi') - \delta l \partial_{\delta l} K_{l,0}(\chi, \chi') + O((\delta l)^2)$ , one readily gets the flow equation

$$\partial_l e^{-S_l[\chi]} = \int \prod_{p \leq \Lambda} d\chi'_p \partial_{\delta l} K_{l,0}(\chi, \chi') e^{-S_l[\chi']}, \quad \text{with} \quad \int \prod_{p \leq \Lambda} d\chi'_p \partial_{\delta l} K_{l,0}(\chi', \chi) = 0. \quad (\text{A.4})$$

For a given coarse graining kernel and a given initial action  $S_\Lambda[\chi]$  this flow equation in principle determines the flow of actions  $l \mapsto S_l[\chi]$ . Of course the usefulness of such a flow equation will largely depend on a good choice for the kernel. In particular the kernel should be ‘almost diagonal’ in the field configurations  $\chi, \chi'$ , so that the multiple integrals are replaced by something simpler, typically a single remaining momentum integral. Specific action-dependent choices for such kernels lead to the various Wilson type (Wegner–Houghton, Polchinski) flow equations for  $S_l$ , which have been employed in the literature (see [130] for the kernel giving Polchinski’s version and [192, 21, 130] for the relation to field redefinitions). Flow equations of this type are known as (exact) *functional renormalization group equations* (FRGEs) of Wilsonian type.

For the sake of contradistinction let us already mention here another type of flow equations, which is formulated in terms of the generating functional for the vertex (1-PI) functions and which uses a mode suppression scheme rather than a coarse graining procedure, namely the effective average action. To set up a functional renormalization flow one does not specify the coarse graining flow by iteration of a 1-step kernel, but rather starts from a functional integral of the form

$$\exp W_k[J] = \int \mathcal{D}\chi \exp \left\{ -S[\chi] - C_k[\chi] + \int dx \chi(x) J(x) \right\}. \quad (\text{A.5})$$

It differs from the standard one (formally) defining the generating functional for the connected correlation functions only by the presence of the  $e^{-C_k[\chi]}$  factor, where  $C_k$  can be thought of as a  $k$ -dependent modification of the bare action. The  $e^{-C_k[\chi]}$  factor later is chosen such that it suppresses the momentum modes  $\chi_p$  with  $p^2 \ll k^2$ , while the modes with  $p^2 \gg k^2$  are integrated out unsuppressed. The response to a variation in the extra scale  $k$  allows one to write down flow equations for

$$\Gamma_k[\phi] = -C_k[\phi] + \Gamma[\phi] \Big|_{S \mapsto S+C_k}. \quad (\text{A.6})$$

Here  $\Gamma[\phi]$  is the usual effective action, which computed for  $S + C_k$  is just the Legendre transform of  $W_k[J]$  in Equation (A.5). This flow equation contains to a certain extent the same information as the original functional integral. This framework will be described in more detail in Appendix C.

3. In general the functional form of the action will change drastically in each coarse graining step  $S_l := S \mapsto S' := S_{l+\delta l}$ . A way to keep track of the change is to consider all (or sufficiently many) interaction monomials  $P_\alpha[\chi]$ ,  $\alpha \in \mathbb{N}$ , compatible with the symmetries of the theory. As an organizing principle one can take the number of derivatives (derivative expansion) or the power of the field (see Equation (A.7)). The only constraints on the possible terms come from symmetry requirements, here e.g. Euclidean invariance or evenness under  $\chi \mapsto -\chi$ , as well as locality requirements for the putative fixed point action. ‘Sufficiently many’  $P_\alpha[\chi]$

means that the action  $S_l$  can be replaced with a good degree of accuracy (as far the evaluation of observables is concerned) by its expansion in  $S_l[\chi] \simeq \sum_{\alpha} u_{\alpha}(l) P_{\alpha}[\chi]$ . In this sense, we shall refer to the set  $P_{\alpha}[\chi]$  loosely as a ‘basis’ in the space of interaction monomials. Importantly it will typically include many more monomials than just the power-counting renormalizable ones. Different bases may be related nonlinearly and nonlocally. For example often an explicit parameterization through running vertex functions  $v_n(x_1, \dots, x_n; l)$  is used, which upon separation of the  $l$  dependence,  $v_n(x_1, \dots, x_n; l) = \sum_{\alpha} v_{n,\alpha}(x_1, \dots, x_n) u_{\alpha}(l)$ , is formally equivalent to a parameterization in terms of monomials  $P_{\alpha}[\chi]$ :

$$\begin{aligned} S_l[\chi] &= \sum_n \frac{1}{n!} \int dx_1 \dots dx_n v_n(x_1, \dots, x_n; l) \chi(x_1) \dots \chi(x_n) \\ &= \sum_{\alpha} u_{\alpha}(l) P_{\alpha}[\chi]. \end{aligned} \quad (\text{A.7})$$

The upshot is that the flow in the action functionals  $S_l[\chi]$  is replaced by a flow  $l \mapsto u_{\alpha}(l, \Lambda, u^{\text{initial}})$  in the parameters. The parameters may be dimensionful, in which case a dependence on both  $l$  and  $\Lambda$  (not just on the ratio) enters. For later reference we also indicated the dependence on the initial values  $u_{\alpha}^{\text{initial}} = u_{\alpha}(l = \Lambda, \Lambda, u^{\text{initial}})$  at scale  $\Lambda$ . Our conventions will be such that  $u_{\alpha}(l)$  carries mass dimension  $d_{\alpha}$  and  $P_{\alpha}[\chi]$  mass dimension  $-d_{\alpha}$ . The kinetic term  $\int dx (\partial\chi)^2$  and its coefficient are taken to be dimensionless which fixes  $\chi$  to have mass dimension  $(d-2)/2$  in  $d$  Euclidean dimensions.

The parameters  $u_{\alpha}$  can be classified into essential and inessential ones. A parameter combination is called *inessential* if the response of the bare Lagrangian to a change in it can be absorbed by a field reparameterization. Explicitly the existence of an inessential parameter combination is signaled by the fact that there exists (locally) a vector field  $u \mapsto \sum_{\alpha} z_{\alpha}(u) \frac{\partial}{\partial u_{\alpha}} = z(u)$  on coupling space such that

$$\sum_{\alpha} z_{\alpha}(u) \frac{\partial L}{\partial u_{\alpha}} = -\frac{\delta S}{\delta \chi} \xi[\chi] + \frac{\delta \xi[\chi]}{\delta \chi} \quad (\text{A.8})$$

for some possibly nonlocal functional  $\xi[\chi](x)$  of the fields  $\chi$ . The right-hand-side is the response of the measure (A.3) with respect to a field reparameterization  $\chi(x) \mapsto \chi'(x) = \chi(x) + \epsilon \xi[\chi](x) + O(\epsilon^2)$ . The concept of field reparameterizations is familiar from power counting renormalizable field theories where linear field parameterizations  $\xi[\chi] \sim \chi$  give rise to wave function renormalizations, e.g.  $z(u) = 2u_0 \partial / \partial u_0 + 2u_2 \partial / \partial u_2 + 4\partial / \partial u_4$  for  $L = u_0 (\partial\chi)^2 + u_2 \chi^2 + u_4 \chi^4$ . In the Kadanoff–Wilson setting the bare Lagrangian contains arbitrary interaction monomials and as a consequence also nonlinear and nonlocal field reparameterizations are allowed (see [192, 21] for further discussion). The first term in Equation (A.8) comes from expanding  $e^{-S[\chi']}$ , with  $S[\chi'] = S[\chi] + \epsilon \int dx \frac{\delta S[\chi]}{\delta \chi(x)} \xi[\chi](x) + O(\epsilon^2)$ , the second term from expanding the Jacobian  $\det \frac{\delta \chi'(x)}{\delta \chi(y)} = \exp \text{Tr} \log \frac{\delta \chi'(x)}{\delta \chi(y)} = 1 + \epsilon \int dx \frac{\delta \xi[\chi](x)}{\delta \chi(x)} + O(\epsilon^2)$ . For the mode suppressed generating functional (A.5) one obtains a field reparameterization (pre-) Ward identity ([237], see Section 5.4.3)

$$\begin{aligned} \int dx \mathcal{D}_k \left[ \frac{\delta}{\delta J} \right] (x) e^{-W_k[J]} &= 0, \\ \mathcal{D}_k[\chi] &= -\xi[\chi] \frac{\delta(S + C_k)[\chi]}{\delta \chi} + \frac{\delta \xi[\chi]}{\delta \chi} + J \xi[\chi], \end{aligned}$$

where in  $\mathcal{D}_k[\chi]$  all terms are evaluated at the same point.



The number  $|I|$  of commuting linearly independent vector fields with the property (A.8) is a characteristic of the Lagrangian, and as in [227] we shall assume that one can choose adapted coordinates such that  $z_i(u) = \partial/\partial z_i$ ,  $i \in I$ . The remaining parameters  $g_i$ ,  $i \in E := \mathbb{N} \setminus I$ , are called *essential parameters or coupling constants*. By definition they are such that  $(z, g) \mapsto u_\alpha(z, g)$  is a diffeomorphism. For convenience we also assume that they have been made dimensionless (with respect to mass dimensions) by a redefinition  $g_i = l^{-d_i} u_i$ , if  $d_i$  is the mass dimension of  $u_i$ . The dimensionless couplings then depend only on the ratio  $l/\Lambda$  and we write  $g_i = g_i(t)$ ,  $t = \ln \Lambda/l - t_0$ ,  $t \geq -t_0$ . Note that this affects the flow direction: Decreasing  $l$  (increasing number of coarse graining steps) corresponds to increasing  $t$ . The variable offset  $t_0$  is useful because by making it large it formally allows one to ask “where a coarse graining trajectory comes from”. For large  $\Lambda$  the offset  $t_0$  can roughly be identified with the (logarithm of the) renormalization point  $\mu$  used in perturbative quantum field theory. Under the above conditions the parameter flow of the  $u_\alpha$  will typically decompose into an autonomous flow equation for the couplings and a non-autonomous flow equation for the inessential parameters. That is  $l \frac{du_\alpha}{dl} = b_\alpha(u(l); l)$  decomposes into

$$-\frac{dg_i}{dt} = \beta_i(g), \quad \text{with } i \in E, \quad l \frac{dz_i}{dl} = b_i(z, g; l), \quad \text{with } i \in I. \quad (\text{A.9})$$

The beta functions  $\beta_i$  (which carry no explicit  $l$  dependence as the  $g_i$  are dimensionless) define a vector field  $\sum_{i \in E} \beta_i(g) \partial/\partial g_i$ ; its integral curves  $t \mapsto g(t)$  are the *renormalization group trajectories*. The flow equations (A.9) of course depend on the chosen coarse graining operation.

4. A fixed point is a zero of the beta functions,  $\beta_i(g^*) = 0$ . The position of  $g_i^*$  thus also depends on the chosen coarse graining operation. A fixed point is called a *UV fixed point* for a given trajectory if this trajectory emanates from the fixed point, i.e.  $g_i(t = -t_0) = g_i^*$ . Formally this can be described by viewing the flow equation for the couplings as one in the offset parameter  $t_0$ , viz.  $dg_i/dt_0 = \beta_i(g(t_0))$ . The condition for a UV fixed point then translates into  $\lim_{t_0 \rightarrow \infty} g_i(t_0) = g_i^*$ . A fixed point is intrinsic to a given coarse graining flow to the same extent that  $\sum_i \beta_i(g) \partial/\partial g_i$  defines a (coordinate independent) vector field.

Certain inessential parameters are still allowed to ‘run’ at the fixed point of the couplings. As a consequence the action with  $g_i = g_i^*$  is not unique, rather the fixed point in the couplings corresponds to a submanifold  $\mathcal{M}_{\text{FP}}$  of fixed point actions  $S_*[\chi]$ . More precisely the class of field redefinitions which commute with the given coarse graining operation will give rise to marginal perturbations (see below) of the fixed point and typically the vector space spanned by these marginal perturbations coincides with the tangent space of  $\mathcal{M}_{\text{FP}}$  at the fixed point [77]. In this case  $S_*[\chi]$  is unique modulo reparameterization terms (like the ones on the right-hand-side of Equation (A.3)) and we shall refer to it as ‘the’ fixed point action.

Most statistical field theories have at least one fixed point, the so-called *Gaussian fixed point*. This means there exists a choice of field variables for which the fixed point action  $S_*[\chi]$  is quadratic in the fields, i.e. only the  $n \leq 2$  terms in Equation (A.4) are nonzero. In a local field theory the fixed point action will typically be local, here  $v_2(x_1, x_2) \sim (u_0 \partial^2 + u_2) \delta(x_1 - x_2)$ , but more generally one could allow for nonlocal ones, here e.g. with  $v_2(x_1, x_2)$  smooth.

Given a fixed point and a coarse graining operation one can (under suitable regularity conditions) decompose the space of actions (the cone of measures) into a stable manifold and an unstable manifold. All actions in the *stable manifold* are driven into the fixed point. The set of points reached on a trajectory emanating from the fixed point is called the *unstable manifold*; any individual emanating trajectory is called a *renormalized trajectory*. The stable manifold is typically infinite-dimensional; this corresponds to the infinitely many interaction

monomials that die out under the successive coarse graining. The dimension of the unstable manifold is of crucial importance because it determines the “degree of renormalizability”.

So far the entire discussion was for a fixed coarse graining operation,  $R$ , say. All concepts (fixed point, stable and unstable manifold, etc.) referred to a given  $R$ . If one now changes  $R$ , the location of the fixed point will change in the given coordinate system provided by the essential couplings. The set of points reached belongs to the *critical manifold* [77]. One aspect of *universality* is that the rates of approach to the fixed point are typically independent of the choice of  $R$ . More generally all quantities defined through a scaling limit are expected to be independent of the choice of  $R$  (within a certain class). Limitations may arise as follows. One parametric families of coarse graining operations  $R_\delta$  may have ‘bifurcation points’ where the dimension of  $\mathcal{M}_{\text{FP}}(R_\delta)$  equals the number of independent marginal perturbations for  $\delta > \delta_0$  and is smaller for  $\delta = \delta_0$ . One expects the emergence of new fixed points (or periodic cycles) at such bifurcation points. The physics interpretation of  $\delta$  may be the (analytically continued) dimension of the system or the range of the interaction.

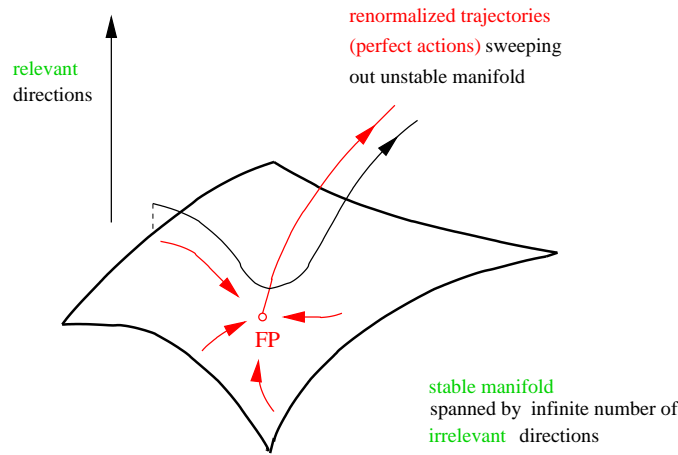


Figure 6: Schematic view of the action flow in the vicinity of a (UV) fixed point. Shown is a typical flow line (black) and a renormalized trajectory (red) emanating from the fixed point. The arrows on the flow lines give the direction of the coarse graining, i.e. of decreasing  $l$  or increasing  $t$ .

5. A lower bound on the dimension of the unstable manifold is obtained from a linearized analysis. The tangent spaces to the stable and unstable manifold at the fixed point are called the spaces of relevant and irrelevant perturbations, respectively. In terms of the couplings  $g_i(t)$  this amounts to the familiar criterion in terms of the right eigenvectors  $V_i^I$  and eigenvalues  $\vartheta_I$  of the stability matrix

$$\Theta_{ij} = \left. \frac{\beta_i(\mathbf{g})}{\partial g_j} \right|_{\mathbf{g}=\mathbf{g}^*}. \quad (\text{A.10})$$

The solution of the linearized flow equation in Equation (A.5) reads  $g_i(t) - g_i^* = \sum_I C_I V_i^I e^{-\vartheta_I t}$ , for constants  $C_I$ . One should add that  $\Theta_{ij}$  is often degenerate and that it is not necessarily symmetric. The eigenvectors are thus not assured to span the tangent space at the fixed point and the eigenvalues are not assured to be real. Whatever eigenvectors  $\Theta_{ij}$  has, however, they are of interest. The space of *irrelevant perturbations* is spanned by the eigenvectors with  $\text{Re } \vartheta_I > 0$ ; the linearized coupling perturbations  $|g_i(t) - g_i^*|$  then decay exponentially in  $t$ . Similarly the space of *relevant perturbations* is spanned by the eigenvectors with  $\text{Re } \vartheta_I < 0$ ;

for them  $|g_i(t) - g_i^*|$  grows exponentially in  $t$ . The borderline case are *marginal perturbations* spanned by the eigenvectors with  $\text{Re } \vartheta_I = 0$ . For them a linearized analysis is insufficient and a refined analysis is needed to decide whether  $g_i(t)$  is driven towards (respectively away from) the fixed point for increasing  $t$ , in which case  $g_i$  is sometimes said to be marginally irrelevant (respectively marginally relevant).

The significance of the stability matrix can be illustrated at the Gaussian fixed point. If the fixed point action  $S_*[\chi]$  is not just quadratic in the fields but also local,  $v_2(x_1, x_2) = \delta(x_1 - x_2)(-u_0 \partial^2 + u_2)$  in Equation (A.4), say, the eigenvalues and eigenvectors of the stability matrix reproduce the structure based on mass (or power counting) scaling dimensions. If one adopts a parameterization where the local Gaussian fixed point is described by  $g_i^* = 0$ , the matrix (A.10) has a set of right eigenvectors  $V^I$  whose eigenvalues are  $-d_i$ , where  $d_i$  are the mass dimensions of the dimensionful couplings  $u_i = k^{d_i} g_i$ . In the setting of Equation (A.1, A.2) this amounts to the following. Consider a coarse graining transformation where momenta in the range  $e^{-t} \leq p/\Lambda \leq 1$  are integrated out. To every monomial  $P_i[\chi]$  with mass dimension  $-d_i$  there corresponds an eigenoperator  $P^I[\chi] = \sum_i V_i^I P_i[\chi]$  whose highest dimensional element is  $P_i[\chi]$  and the corresponding eigenvalue is  $e^{td_i}$ . Here one can see directly that the monomials which are irrelevant with respect to a local Gaussian fixed point (those which ‘die out’ under successive coarse graining operations) are the ones with mass dimension  $-d_i > 0$ . For example with the conventions set after Equation (A.4) an  $\int dx \chi^4$  term has mass dimension  $-d_i = d - 4$  in  $d$  Euclidean dimensions.

On the other hand the amount of information that can be extracted from the stability matrix is often limited by the fact that it is degenerate. For illustration let us consider some simple examples. It is convenient to consider the flow as a function of the off-set  $t_0$  so that  $dg_i/dt_0 = \beta_i(g(t_0))$  is the appropriate flow equation. Let us assume that (for reasons of positivity of energy, say) the couplings are required to be non-negative. In the case of a single coupling, then  $\beta(g) = \mp g^2$  has the fixed point  $g^* = 0$ , but with the upper sign the unstable manifold is one-dimensional, while with the lower sign the unstable manifold is empty. Indeed the solution  $g(t_0) = 1/(g_0^{-1} \pm t_0)$  is in the first case positive for all  $t_0 > 0$  and approaches the fixed point for  $t_0 \rightarrow \infty$ , while with the lower sign the fixed point cannot be reached with positive values of the coupling. Both are ‘paradigmatic’ situations mimicking the perturbative behavior of a Yang–Mills coupling and a  $\phi_4^4$  coupling, respectively. Note that in both cases the stability matrix (A.10) vanishes identically, so the attempt to gain insight into the unstable manifold via the linearized analysis already fails in this trivial example. A multi-dimensional generalization is  $\beta_i(g) = -\sum_{j=1}^N (g_j - g_j^*)^2$ , where the unstable manifold of the fixed point  $g_i^* > 0$  consists only of the halfline  $\{g_i^* + s, s > 0\}$  (of co-dimension  $N - 1$ ).

An important case when a linearized analysis is insufficient is when the number of independent marginal perturbations is larger than the dimension of the target space to  $\mathcal{M}_{\text{FP}}(R)$  (see the remark at the end of Comment 4). One may then be able to enlarge the stable or the unstable manifold (or both) by submanifolds of points which are driven towards or away from it with ‘vanishing speed’.

Returning to the general discussion, a schematic pattern of a coarse graining flow in the vicinity of a fixed point is shown in Figure 6. Individual flow lines  $t \mapsto g_i(t)$  starting outside the stable manifold will in general first approach the fixed point, without touching down, and then shoot away from it. In order to (almost) touch down at the fixed point the initial values  $g_i^{\text{initial}}$  have to be carefully fine tuned. With ideal fine tuning the trajectory then splits into two parts. One part that moves within the stable manifold into the fixed point and another part that emerges from it. The latter has the fixed point couplings as its initial values,  $g_i^{\text{initial}} = g_i^*$ , and is called a *renormalized trajectory*; with respect to it the fixed point is an ultraviolet one. Its physics significance is that the actions associated with points on the

renormalized trajectory are *perfect*, in the sense that the effect of the cutoff on observables is completely erased, even when the couplings are not close to their fixed point values. More realistic than the construction of perfect actions is that of improved actions, designed such that for given values of the couplings the cutoff effects are systematically diminished (see [102] for further discussion). In order to identify a renormalized trajectory the initial values  $g_i^{\text{initial}}$  of the relevant couplings have to be fine tuned. A statistical field theory for which this amounts to a manifestly finite-dimensional problem is called strictly renormalizable in the ultraviolet, otherwise we suggest to call it weakly renormalizable.

**UV strict (weak) renormalizability:**

- (a) There exists an UV fixed point, i.e.  $\beta_i(g_i^*) = 0$  and  $g_i(t = -t_0) = g_i^*$ .
- (b) Its unstable manifold is finite (infinite)-dimensional.

Importantly this is a nonperturbative definition of renormalizability. The averages defined by backtracing a perfect action along the renormalized trajectory into the fixed point have all desirable properties: They are independent of the cutoff scale  $\Lambda$  and independent of all the irrelevant couplings, in the sense that they become computable functions of the relevant couplings. Whenever the unstable manifold is well defined as a geometric object, its dimension and structure is of course a coordinate independent notion. In practice the existence of (patches of) the unstable manifold is only established retroactively once a good ‘basis’ of interaction monomials  $P_\alpha[\chi]$  has been found. Then a large class of other choices will be on an equal footing giving the concept a geometric flavor. The condition of renormalizability as discussed so far only singles out a subclass of regularized (discretized) statistical field theories. We shall outline below why and how this subclass can be used to construct continuum quantum field systems with certain robustness properties. Before turning to this we complete our list of comments on the Kadanoff–Wilson renormalization.

- 6. This is a direct consequence of Point 4. The ‘degree’ of renormalizability is the dimension of the unstable manifold, which depends on the fixed point considered.

In the framework of statistical lattice field theories the importance of the construction principle 1–6 lies in the fact that it allows one to construct a scaling limit without introducing uncontrolled approximations. Usually this employs the concept of a correlation length and of the critical manifold as the locus of all points in the space of measures with infinite correlation length. In a gravitational context it is not obvious how to adapt this notion adequately, which is why we tried to avoid direct use of it, see the discussion at the end of Point 4. In statistical field theories a scaling limit is usually constructed in terms of multipoint functions of the basic fields. The lattice spacing is sent to zero while simultaneously the couplings are moved back into the fixed point along a renormalized trajectory. As mentioned before we are interested here in massive scaling limits, meaning that among the systems on the renormalized trajectory used only the one ‘at’ the fixed point is scale invariant. In the case of quantum gravity the multipoint functions of the basic fields might not be physical quantities, so the appropriate requirement is that a scaling limit exists for generic physical quantities (see Section 1.1).

Initially all concepts in the construction 1–6 refer to a choice of coarse graining operation. One aspect of universality is that all statistical field theories based on fixed points related by a change of coarse graining operation have the same scaling limit. One can thus refer to an equivalence class of scaling limits as a *continuum limit*. The construction then entails that whenever a continuum limit exists physical quantities become independent of the UV cutoff *and* of the choice of the coarse graining operation (within a certain class).

So far we did not presuppose unitarity/positivity of the resulting quantum field system. For a Euclidean statistical field theory, however, reflection positivity of the class of actions used provides

an easily verifiable sufficient condition for the existence of a positive definite inner product on the physical state space of the theory. This is why relativistic unitary continuum quantum field theories (QFTs) can be constructed along the above lines in a way that does not require uncontrolled approximations from the outset. Of course the rigorous construction of a relativistic unitary QFT along these lines remains an extremely challenging problem and has only been achieved in a few cases. However numerical techniques often allow one to verify the renormalizability properties to good accuracy. The resulting QFT then is renormalizable in the above sense “for all practical purposes”; an example are Yang-Mills theories. As a consequence the extracted continuum physics will have the desired universal properties for all practical purposes likewise.

This concludes our ‘birds eye’ summary of basic renormalization group concepts. The concrete implementation in a given field theory quickly becomes fairly technical. A key problem is to obtain mathematical control over the scale dependencies and these are often extremely hard to come by (see e.g. [195]). Outside a dedicated group of specialists one usually resorts to uncontrolled approximations. In this review we do so likewise. First, because of the early stage in which the investigation of the ‘gravitational renormalization group’ is, and second, because also the controlled approximations – where they exist – often draw from experience gained from uncontrolled approximations.

## B The Background Effective Action

The effective action summarizes the content of a field theory in a way which is technically convenient and physically instructive. It is familiar from perturbation theory as the generating functional of the “1-particle irreducible” (1-PI) Green’s functions. It can however be given a meaning independent of perturbation theory and its functional derivatives can be used to reconstruct all the correlation functions of the field theory under consideration. Since the latter can be viewed as the moments of the functional measure, this replaces the “cone of measures” (or the space of “Wilsonian effective actions  $S_l$ ”) in the Wilsonian setting by the “space of effective action functionals” as the arena on which the renormalization group acts.

The standard effective action admits two fruitful generalizations, discussed in Appendices B and C, respectively. The first one is adapted to theories with symmetries, which can be field reparameterization symmetries, gauge symmetries, or both. These effective actions are known as “background effective actions”. In Appendix B.2 we provide a concise summary of these constructions. The background reparameterization and/or gauge invariance also provides an, often crucial, simplification of the renormalization. In Appendix B.3 we describe this for the case of field reparameterization symmetries in Riemannian sigma-models, which also provides some of the renormalization prerequisites for Section 3. In a gravitational context the background field formalism also provides a setting which, despite the name, can reasonably be regarded as “background independent”.

The second generalization of the standard effective action is one where, roughly, the bare action  $S_\Lambda$  is replaced by  $S_\Lambda + C_{\Lambda,k}$ , where  $C_{\Lambda,k}$  effectively suppresses field modes with momenta less than  $k$ . This leads to the effective average action reviewed in Appendix C.

For definiteness we assume a Euclidean setting throughout these Appendices.

### B.1 Standard effective action and its perturbative construction

We begin with a quick reminder on the standard effective action: After coupling  $\chi(x)$  to a source  $J(x)$  one has an (initially formal) functional integral representation for the Euclidean generating functional of the connected Schwinger functions:  $W[J] = \ln \int \mathcal{D}\chi \exp\{-S[\chi] + \int dx \chi(x)J(x)\}$ . Source dependent normalized expectation values of some (smooth) observable  $\mathcal{O}(\chi)$  are defined by

$$\langle \mathcal{O} \rangle_J = e^{-W[J]} \mathcal{O} \left( \frac{\delta}{\delta J} \right) e^{W[J]}, \quad (\text{B.1})$$

and it is assumed that (at least for vanishing source) the map  $\mathcal{O} \mapsto \langle \mathcal{O} \rangle$  is a positive functional in the sense that positive functions  $\mathcal{O}(\chi)$  have positive expectation values. In order to make the functional integral well-defined a UV cutoff  $\Lambda$  is needed; for example one could replace  $\mathbb{R}^d$  by a  $d$ -dimensional lattice  $\mathbb{Z}^d$  with lattice spacing  $\Lambda^{-1}$ . The flat reference functional measure  $\mathcal{D}\chi$  would then be proportional to  $\prod_{x \in \mathbb{Z}^d} d\chi(x)$ . In the following we implicitly assume such a UV regularization but leave the details unspecified and use a continuum notation for the fields and their Fourier transforms. The dependence on the cutoff  $\Lambda$  will be specified only when needed. We will also omit overall normalizations in the functional integrals. Whenever  $\mathcal{D}\chi e^{-S[\chi]}$  in addition to being positive is also a normalized measure (on a suitable space of functions  $\chi$ ) it follows from Hölder’s inequality that  $W[J]$  is a convex functional of the source, i.e.  $W[\alpha J_1 + \beta J_2] \leq \alpha W[J_1] + \beta W[J_2]$ , for  $\alpha + \beta = 1$ ,  $\alpha, \beta \geq 0$ . Taking  $J_1 = J$ ,  $J_2 = J + f$  and expanding in powers of  $f$  gives

$$0 \leq \int dx dy \frac{\delta^2 W[J]}{\delta J(x) \delta J(y)} f(x) f(y) = \left\langle \left( \int dx (\chi(x) - \langle \chi(x) \rangle_J f(x)) \right)^2 \right\rangle_J. \quad (\text{B.2})$$

This means the second functional derivative  $W^{(2)}(x, y)$  is a kernel of positive type; under suitable falloff conditions it defines a positive bounded integral operator on the space of the functions



*f.* Kernels of positive type allow one to (re-)construct a Hilbert space such that schematically  $W^{(2)}(x, y)$  is recovered as an inner product (“a two-point function”). A fully fledged reconstruction of the operator picture requires knowledge of all multipoint functions and is roughly the content of the Osterwalder–Schrader reconstruction theorem.

Since  $W[J]$  is convex, the *effective action* can be introduced as the Legendre transform  $\Gamma[\phi] := \sup_J \{ \int \phi J - W[J] \}$ , which is a convex functional of  $\phi$ . Although  $W[J]$  is always convex it may not be differentiable everywhere. In fact,  $W[J]$  has ‘cusps’ in the case of spontaneous symmetry breaking. Even on the subspace of homogeneous solutions the supremum in the definition of  $\Gamma[\phi]$  may then be reached for several configurations  $J_* = J_*[\phi]$  and  $\Gamma[\phi]$  is ‘flat’ in these directions. If  $W[J]$  admits a series expansion in powers of  $J$ , a formal inversion of the series  $\delta W[J]/\delta J = \phi$  defines a unique  $J_*[\phi]$  with the property  $J_*[\phi = 0] = 0$  [237]. Often this extra assumption isn’t needed and we shall write  $J_* = J_*[\phi]$  for any configuration on which the supremum in the Legendre transform is reached; functional derivatives with respect to  $J$  evaluated at  $J_*[\phi]$  will be denoted by  $\delta/\delta J_*$ . The defining properties of the Legendre transform then read

$$\Gamma[\phi] = \int dx \phi J_*[\phi] - W[J_*[\phi]], \tag{B.3}$$

$$\frac{\delta W[J_*]}{\delta J_*} := \langle \chi \rangle_{J_*} = \phi, \tag{B.4}$$

$$\frac{\delta \Gamma[\phi]}{\delta \phi} = J_*[\phi], \tag{B.5}$$

$$\frac{\delta^2 \Gamma[\phi]}{\delta \phi \delta \phi} \frac{\delta^2 W[J_*]}{\delta J_* \delta J_*} = 1, \tag{B.6}$$

where Equation (B.6) is short for the fact that the  $\phi$ -dependent integral operators with kernels  $\Gamma^{(2)}(x, y) := \delta^2 \Gamma[\phi]/\delta \phi(x) \delta \phi(y)$  and  $G(x, y) := \delta^2 W[J_*]/\delta J_*(x) \delta J_*(y)$  are inverse to each other. Since  $G(x, y)$  is a kernel of positive type, so is  $\Gamma^{(2)}(x, y)$ .

To switch off the source one selects configurations  $\phi_*$  such that  $\delta \Gamma[\phi]/\delta \phi|_{\phi=\phi_*} = J_*[\phi_*] = 0$ . As  $J_*[\phi]$  defined by series inversion one can directly take  $\phi_* = 0$  as the source-free condition. This is typical in the absence of spontaneous symmetry breaking, otherwise one should use the defining relation  $0 = \delta \Gamma[\phi]/\delta \phi|_{\phi=\phi_*}$  for  $\phi_*$  to switch off the source. The vertex functions are defined for  $n \geq 2$  by

$$\Gamma^{(n)}(x_1, \dots, x_n) := \frac{\delta}{\delta \phi(x_1)} \dots \frac{\delta}{\delta \phi(x_n)} \Gamma[\phi] \Big|_{\phi=\phi_*}. \tag{B.7}$$

In a situation without spontaneous symmetry breaking the  $\Gamma^{(n)}$  are independent of the choice of  $\phi_*$ . The original connected Greens functions can be reconstructed from the  $\Gamma^{(n)}$  and  $[\Gamma^{(2)}]^{-1} = G|_{\phi=\phi_*}$  by purely algebraic means, as can be seen by repeated differentiation of Equation (B.4). Finally we remark that both the ‘connectedness property’ of the multipoint functions and the ‘irreducibility property’ of the vertex functions can be characterized intrinsically [55], i.e. without going through the above construction.

Inserting Equation (B.3) into the definition of  $W[J]$  one sees that the effective action is characterized by the following functional integro-differential equation:

$$\exp\{-\Gamma[\phi]\} = \int \mathcal{D}\chi \exp \left\{ -S[\chi] + \int dx (\chi - \phi)(x) \frac{\delta \Gamma[\phi]}{\delta \phi(x)} \right\} \tag{B.8}$$

(see e.g. [223]). In itself of course Equation (B.8) is useless because one still has to perform a functional integral. It can be made computationally useful, however, in two ways. First as a tool to generate a recursive algorithm to compute  $\Gamma[\phi]$  perturbatively, and second as a starting point to derive a functional differential equation for  $\Gamma[\phi]$ . For the latter we refer to Appendix C.2, here we briefly recap the perturbative construction.

It is helpful to restore the implicit dependence on the UV cutoff  $\Lambda$  and we write  $\Gamma_\Lambda[\phi]$  from now on. In outline, the perturbative algorithm based on Equation (B.8) involves the following steps. One introduces the loop counting parameter  $\lambda$  as follows:  $\Gamma_\Lambda \mapsto \frac{1}{\lambda}\Gamma_\Lambda$ ,  $S_\Lambda \mapsto \frac{1}{\lambda}S_\Lambda$ ,  $\chi - \phi = \sqrt{\lambda}f$ , where  $S_\Lambda$  is the bare action depending on  $\Lambda$ . After the rescaling one expands the exponent on the right-hand-side in powers of  $\sqrt{\lambda}$ . Schematically the result is

$$S_\Lambda[\phi + \sqrt{\lambda}f] - \sqrt{\lambda}f^i \partial_i \Gamma_\Lambda[\phi] = S_\Lambda[\phi] + \frac{\lambda}{2} f^i f^j \partial_i \partial_j S_\Lambda[\phi] + \lambda \sqrt{\lambda} \left( \frac{1}{3!} \partial_i \partial_j \partial_k S_\Lambda[\phi] f^i f^j f^k - f^i \partial_i \Gamma_{\Lambda,1}[\phi] \right) + O(\lambda^2), \quad (\text{B.9})$$

where we momentarily use DeWitt's 'condensed index notation' [68], that is, functional differentiation  $\delta/\delta\phi(x)$  is denoted by  $\partial_i$  and the index contraction is short for a  $x$ -integration. Further we used an ansatz for  $\Gamma_\Lambda[\phi]$  of the form

$$\Gamma_\Lambda[\phi] = S_\Lambda[\phi] + \sum_{\ell \geq 1} \Gamma_{\Lambda,\ell}[\phi] \lambda^\ell. \quad (\text{B.10})$$

Note that the term linear in  $f$  drops out without assuming that  $\phi$  is an extremizing configuration. Equation (B.9) is now re-inserted into Equation (B.8) and the exponentials involving positive powers of  $\sqrt{\lambda}$  are expanded. This reduces the evaluation of the functional integral on the right-hand-side to the evaluation of correlators  $\langle f^{i_1} f^{i_2} \dots f^{i_n} \rangle_\Lambda$  with respect to the Gaussian measure with covariance  $\partial_i \partial_j S_\Lambda[\phi]$ . By the source-free condition the ones with an odd number of  $f$ 's vanish, so that also the right-hand-side gives an expansion in integer powers of  $\lambda$ . Matching both sides of Equation (B.9) then gives a recursive algorithm for the computation of the  $\Gamma_{\Lambda,\ell}[\phi]$ ,  $\ell \geq 1$ . The first two equations are

$$\begin{aligned} \Gamma_{\Lambda,1}[\phi] &= \frac{1}{2} \text{Tr}_\Lambda \ln \partial^i \partial_i S_\Lambda[\phi], \\ \Gamma_{\Lambda,2}[\phi] &= \frac{1}{24} \partial_{i_1} \partial_{i_2} \partial_{i_3} \partial_{i_4} S_\Lambda[\phi] \langle f^{i_1} f^{i_2} f^{i_3} f^{i_4} \rangle_\Lambda \\ &\quad - \frac{1}{72} \partial_{i_1} \partial_{i_2} \partial_{i_3} S_\Lambda[\phi] \partial_{j_1} \partial_{j_2} \partial_{j_3} S_\Lambda[\phi] \langle f^{i_1} f^{i_2} f^{i_3} f^{j_1} f^{j_2} f^{j_3} \rangle_\Lambda \\ &\quad + \frac{1}{6} \partial_{i_1} \partial_{i_2} \partial_{i_3} S_\Lambda[\phi] \partial_j \Gamma_{\Lambda,1}[\phi] \langle f^{i_1} f^{i_2} f^{i_3} f^j \rangle_\Lambda - \frac{1}{2} \partial_i \Gamma_{\Lambda,1}[\phi] \partial_j \Gamma_{\Lambda,1}[\phi] \langle f^i f^j \rangle_\Lambda. \end{aligned} \quad (\text{B.11})$$

The expression for  $\Gamma_{\Lambda,1}[\phi]$  is the familiar regularized one-loop determinant. Inserting this into the second equation the reducible parts cancel and one can verify the equivalence to the two-loop result in [237] (Equation (A6.12)). The presence of the UV cutoff in  $S_\Lambda$  renders the expressions for  $\Gamma_{\Lambda,\ell}[\phi]$ ,  $\ell \geq 1$ , well-defined.

The removal of the cutoff is done by a recursive procedure which is based on the following crucial fact: In a perturbatively (quasi) renormalizable QFT the divergent (as  $\Lambda \rightarrow \infty$ ) part of the  $\ell$  loop contribution to the effective action  $\Gamma_{\Lambda,\ell}[\phi]$  is *local*, i.e. it equals a single  $dx$  integral over a local function in the fields  $\phi$  and their derivatives. Moreover this divergent part has the *same* structure as the bare action (2.2) with specific parameter functions  $u_{\alpha,\ell}(u(\mu), \Lambda, \mu)$ . This can be used to compute the parameter functions in the bare action (2.2) recursively in the number of loops.

Schematically one proceeds as follows. Let  $S_{\Lambda,\leq L}[\chi_\Lambda]$  denote the bare action (2.2) at  $L$  loop order, with the parameter functions  $u_{\alpha,\ell}(u(\mu), \Lambda, \mu)$ ,  $\ell \leq L$ , known. In particular the limits  $\lim_{\Lambda \rightarrow \infty} S_{\Lambda,\leq L}[\chi_\Lambda] = S_{\mu,\leq L}[\chi_\mu]$  is finite and defines the  $L$ -loop renormalized action. Further the parameter functions are such that  $\sum_{\ell \leq L} \Gamma_{\Lambda,\ell}[\phi] \lambda^\ell$  has a finite limit as  $\Lambda \rightarrow \infty$ . Using this bare

action  $S_{\Lambda, \leq L}$  one computes the effective action  $\Gamma_{\Lambda, L+1}$  at the next order. According to the above result it can be decomposed as

$$\Gamma_{\Lambda, L+1}[\phi] = \Gamma_{\infty, \ell}[\phi] - S_{\Lambda, L+1}^{c.t.}[\phi], \quad (\text{B.12})$$

where  $S_{\Lambda, L+1}^{c.t.}[\phi] = \sum_{\alpha} u_{\alpha, L+1}(u(\mu), \Lambda, \mu) P_{\alpha}[\phi]$ , determines the parameter functions at the next order. The ‘‘counter term’’ action  $\lambda^{L+1} S_{\Lambda, L+1}^{c.t.}$  is then added to the bare action  $S_{\Lambda, \leq L}$  to produce  $S_{\Lambda, \leq L+1}$ . Re-computing  $\Gamma_{\Lambda, L+1}$  with this new bare action, the limit  $\Lambda \rightarrow \infty$  turns out to be finite, that is subdivergencies cancel as well. The traditional proof of this result (e.g. for scalar field theories) involves the analysis of Feynman diagrams and their (sub-)divergencies. More elegant is the use of flow equations (see [123] for a recent review). Once the renormalized  $\Gamma_{\infty}[\phi] = \sum_{\ell \geq 0} \Gamma_{\infty, \ell}[\phi] \lambda^{\ell}$  is known (to exist), the vertex functions (B.7) and the S-matrix elements are in principle likewise known to all loop orders. The latter can then be shown to be invariant under local redefinitions of the fields.

## B.2 Survey of background field formalisms

In the above brief recap on the standard effective action we considered for simplicity only theories without gauge or field reparameterization symmetries. The standard effective action formalism can also be used for theories with symmetries, but then has several drawbacks. Most importantly the standard effective action is not a (gauge or field reparameterization) invariant functional of its argument. The background field method is designed to produce a modified ‘‘background’’ effective action which is an invariant functional of its argument. This feature is instrumental for the renormalization process. We first illustrate the role of the background field configuration in the scalar case and then discuss the two variants needed to cope with gauge and field reparameterization symmetries consecutively. The variant used to address gauge and field reparameterization symmetries in combination (Vilkowisky–deWitt) will not be needed here, but we shall mention it briefly for the sake of comparison.

Before turning to the formalism let us briefly comment on the status of the background field configuration. First one should stress that the term ‘‘background field’’ in this context does *not* refer to a solution of the classical field equations. It is an auxiliary device used to formulate covariance properties that constrain the renormalization flow. The dependence on the background configuration is controlled by splitting Ward identities that, roughly, ensure that nothing depends on the way how the integration variable in the functional integral is decomposed into a reference configuration and a fluctuation field. Once a source free condition is imposed the ‘‘background’’ gets related to the expectation value of the fluctuation field through a consistency condition involving the full quantum effective action (see Equations (2.51) and (B.53) below).

When applied to quantum gravity, the background field formalism can be viewed as giving rise to a formulation with a state-dependent dynamically adjusted reference metric. The bare manifold is initially equipped with a reference geometry, but rather than being external it eventually gets related to the average of the quantum metric in a self-consistent manner. As an incomplete classical analogy one may take the variational principles for general relativity which in addition to the dynamical metric invoke a fiducial background metric (see e.g. [74] and references therein). The latter only provides the desired covariance properties and in the presence of generic boundary conditions serves to define conserved quantities relative to the background.

As a further clarification one should add that in the case of gauge theories the on-shell projections of the Green’s functions (e.g. S-matrix elements) computed from the background effective action are not assumed to be physical quantities from the outset. In particular in the case of gravity the precise role of the on-shell projections remains to be understood. The main reason for focusing on the background effective action here are its ‘benign’ properties under UV renormalization. The

background covariance constrains the renormalization flow, both on the level of a Wilsonian action and on the level of an effective average action.

The background field formalism comes in two main variants: one based on a *linear* background-fluctuation split and the other based on a *geodesic* background-fluctuation split. The latter is used to cope with field reparameterization symmetries. Both variants can be applied to non-gauge theories (regular Lagrangians) and gauge theories (singular Lagrangians). We briefly discuss these variants consecutively, later on the variants in Sections B.2.2 and B.2.3 will be used.

### B.2.1 Linear split, non-gauge theories

In this simple case the use of a background field formalism is ‘overkill’. However it is a convenient way to introduce the relevant structures. The starting point is again the generating functional for the connected Green’s functions, but with a modified coupling to the source  $J$ . Specifically, one linearly decomposes the quantum field, i.e. the integration variable  $\chi$ , according to  $\chi(x) = \bar{\chi}(x) + f(x)$ , where  $\bar{\chi}$  is a background configuration to be adjusted later and  $f$  is the fluctuation field, i.e. the new integration variable. In the corresponding generating functional only  $f$  not the complete field  $\chi$  is coupled to the source. For the sake of illustration we allow the action to depend explicitly on the background field  $\chi$ , i.e.  $S = S[\chi; \bar{\chi}]$ . One introduces

$$\exp W[J; \bar{\chi}] = \int \mathcal{D}f \exp \left\{ -S[\bar{\chi} + f; \bar{\chi}] + \int dx f(x) J(x) \right\}, \quad (\text{B.13})$$

and for a given  $\phi$  its Legendre transform

$$\Gamma[\phi; \bar{\chi}] := \sup_J \left\{ \int dx \phi J - W[J; \bar{\chi}] \right\}. \quad (\text{B.14})$$

Assuming for simplicity that  $W[J; \bar{\chi}]$  is differentiable everywhere, for an extremizing configuration  $J_* = J_*[\phi; \bar{\chi}]$  one has

$$\phi = \langle f \rangle_{J_*} = \frac{\delta W[J_*; \bar{\chi}]}{\delta J_*}, \quad \frac{\delta \Gamma[\phi; \bar{\chi}]}{\delta \phi} = J_*[\phi, \bar{\chi}], \quad (\text{B.15})$$

i.e.  $\phi$  can be interpreted as the normalized expectation value  $\langle f \rangle_{J_*}$  of the fluctuation field. The counterpart of Equation (B.8) is

$$\exp\{-\Gamma[\phi; \bar{\chi}]\} = \int \mathcal{D}\chi \exp \left\{ -S[\chi; \bar{\chi}] + \int dx (\chi - \bar{\chi} - \phi)(x) \frac{\delta \Gamma[\phi; \bar{\chi}]}{\delta \phi(x)} \right\}. \quad (\text{B.16})$$

Differentiating with respect to  $\bar{\chi}$  gives

$$\frac{\delta \Gamma}{\delta \bar{\chi}} - \frac{\delta \Gamma}{\delta \phi} - \left\langle \frac{\delta S}{\delta \bar{\chi}} \right\rangle_{J_*}. \quad (\text{B.17})$$

We shall refer to identities of this form as “splitting Ward identities”. Here Equation (B.17) expresses the fact that that the “linear splitting” symmetry  $\delta \langle f \rangle = \eta$ ,  $\delta \bar{\chi} = -\eta$  is violated only by the explicit background dependence of the action.

The condition that the source  $J_*$  in Equation (B.15) vanishes is usually solved by formal power series inversion and then gives  $\phi = 0$  as the only solution,  $J_*[\phi = 0; \bar{\chi}] = 0$ . More generally the vanishing of  $J_*[\phi; \bar{\chi}]$  defines locally  $\bar{\chi}$  as a function of  $\phi$ , say,  $\bar{\chi} = \bar{\chi}_*(\phi)$ . For such configurations  $\Gamma$  becomes a functional of a single field

$$\bar{\Gamma}[\phi] := \Gamma[\phi; \bar{\chi}_*(\phi)], \quad (\text{B.18})$$

which obeys

$$\frac{\delta\bar{\Gamma}[\phi]}{\delta\phi} = \left\langle \frac{\delta S}{\delta\bar{\chi}} \right\rangle_{J^*} \frac{\delta\chi_*}{\delta\phi}. \quad (\text{B.19})$$

We briefly mention two applications of the relations (B.65), both for actions  $S[\chi]$  without explicit background dependence. The relation (B.17) is often used for constant background field configurations  $\bar{\chi}$ . It then implies that vertex functions with vanishing external momenta can be written as  $\bar{\chi}$  derivatives of vertex functions with a smaller number of legs

$$\begin{aligned} \frac{\partial}{\partial\bar{\chi}}\Gamma^{(n)}(x_1, \dots, x_n; \bar{\chi}) &= \int dy \Gamma^{(n)}(x_1, \dots, x_n, y; \bar{\chi}), \\ \text{with } \Gamma^{(n)}(x_1, \dots, x_n; \bar{\chi}) &= \left. \frac{\delta\Gamma[\phi; \bar{\chi}]}{\delta\phi(x_1) \dots \delta\phi(x_n)} \right|_{\phi=0}. \end{aligned} \quad (\text{B.20})$$

In the context of renormalization theory the following discussion (taken from Howe et al. [110]) illustrates the point of the identity (B.17). Expanding an interaction monomial like  $\chi^4$  one gets

$$\chi^4 = \bar{\chi}^4 + 4f^3\bar{\chi} + 6f^2\bar{\chi}^2 + 4f\bar{\chi}^3 + f^4, \quad (\text{B.21})$$

where in principle all interaction vertices could be renormalized differently. However, because of Equation (B.17) this does *not* happen, the wave function renormalization constants for  $\bar{\chi}$  and  $f$  are actually the same:

$$Z_{\bar{\chi}} = Z_f. \quad (\text{B.22})$$

This entails that the counterterms are functionals of the *full* field  $\chi = \bar{\chi} + f$ , not of  $\bar{\chi}$  and  $f$  separately. Relations like Equations (B.17) – referred to as “linear splitting Ward identities” in [110] – thus provide a crucial simplification in the description of the renormalization flow, once the background-fluctuation split has been adopted for other reasons.

Good reasons to adopt such a split exist in theories with symmetries, which can be local gauge symmetries, or field reparameterization symmetries, or both. In all situations the background field method offers key advantages in that it can produce an effective action which is an *invariant functional of its argument*. Via the above splitting principle this then greatly restricts the form of the (Wilsonian) renormalized action. In a non-background field formalism these symmetries in contrast have to be imposed by relating possibly noninvariant terms or pieces of the renormalized action via conventional Ward identities, like Equation (A.9) in the case of field reparameterizations.

We first describe the background field technique for a non-gauge theory where reparameterization invariance in field space is aimed at, and then for a diffeomorphism gauge theory (the case of Yang–Mills theories runs completely parallel). Finally we mention the setting where both gauge and field reparameterization invariance is aimed at. General references are [68, 49, 223, 224, 129, 178].

## B.2.2 Geodesic split, non-gauge theories

Here reparameterization invariance in field space is aimed at; the original construction is due to Honerkamp et al. [107, 109]. Since invariance under local field redefinitions is a hallmark of physical quantities this field reparameterization invariant effective action is an object much more intrinsic to the field theory under consideration. Morally speaking in this technique the field reparameterization Ward identity (A.9) is built in, and does not have to be imposed along with the solution/definition of the functional integral. For the construction of the covariant effective action the field configuration space is equipped with a metric and the associated metric connection. The Lagrangian is assumed to be reparameterization invariant in the sense that

$$L_{\mathfrak{h}}(\chi + v) = L_{\mathfrak{h} + \mathcal{L}_v \mathfrak{h}}(\chi) + O(v^2), \quad (\text{B.23})$$

$$-\partial^\mu \mathcal{C}_\mu + (\mathcal{L}_v \mathfrak{h})_{ij} \frac{\partial L_{\mathfrak{h}}}{\partial \mathfrak{h}_{ij}} = v^i \frac{\delta S_{\mathfrak{h}}}{\delta \chi^i}. \quad (\text{B.24})$$

Here  $v = v^i(\chi)\partial/\partial\chi^i$  is a vector field on the configuration space  $\mathcal{M}$  and  $S_{\mathfrak{h}} = \int dx \sqrt{\mathfrak{h}} L_{\mathfrak{h}}$  is the action associated with  $L_{\mathfrak{h}}$ . The second equation is a consequence of the first; the quantity  $\mathcal{C}_{\mu}$  can be viewed as a “diffeomorphism current”. For example  $\mathcal{C}_{\mu} = v^i \partial L_{\mathfrak{h}} / (\partial(\partial^{\mu}\chi^i))$ , if  $L_{\mathfrak{h}}$  depends algebraically on  $\chi^i$  and  $\partial\chi^i$ , and with obvious generalizations to higher derivative theories. Formally such a reparameterization invariance can always be achieved. Naturally it arises for a scalar field theory when the metric in field space enters via the kinetic term in the Lagrangian

$$L_{\mathfrak{h}}(\chi) = \frac{1}{2} \mathfrak{h}_{ij}(\chi) \partial^{\mu}\chi^i \partial_{\mu}\chi^j + \dots, \quad (\text{B.25})$$

where the other terms are supposed to preserve Equations (B.23, B.24). The prototype example of such systems are Riemannian sigma-models, for which the resulting “covariant background field method” is summarized in Appendix B.3. From a Wilsonian perspective it is natural to allow for higher derivative interactions and to use only the invariance (B.23, B.24) for the formulation of the framework.

One now describes the configurations  $\chi$  in terms of an arbitrary (off-shell) background configuration and geodesic normal coordinates  $\xi$ , which are new dynamical fields. That is, a nonlinear background-fluctuation split  $\chi^i = \chi_{\mathfrak{h}}^i(\bar{\chi}, \xi)$  is used. Here  $\chi_{\mathfrak{h}}$  is the function on  $\mathcal{M} \times T\mathcal{M}$  such that its value  $\chi_{\mathfrak{h}}^i(\bar{\chi}, \xi)$  gives the endpoint of the (locally unique) geodesic in  $(\mathcal{M}, \mathfrak{h})$  connecting  $\bar{\chi}$  to  $\chi$  and having  $\xi = \xi^i \partial/\partial\chi^i$  as the tangent vector at  $\bar{\chi}$ . The normal coordinate expansion of  $\chi_{\mathfrak{h}}$  is a power series in  $\xi$  with coefficients built from the Christoffel symbols  $\Gamma_{jk}^i(\bar{\chi})$  of  $\mathfrak{h}_{ij}$  evaluated at  $\bar{\chi}$ . We shall also need the inverse series  $\xi_{\mathfrak{h}}^i(\bar{\chi}, \chi)$ , defined by  $\chi^i = \chi_{\mathfrak{h}}^i(\bar{\chi}, \xi_{\mathfrak{h}}(\bar{\chi}, \chi))$ . To quadratic order one has

$$\begin{aligned} \chi^i &= \chi_{\mathfrak{h}}^i(\bar{\chi}, \xi) = \bar{\chi}^i + \xi^i - \frac{1}{2} \Gamma_{jk}^i(\bar{\chi}) \xi^j \xi^k + \dots \\ \xi^i &= \xi_{\mathfrak{h}}^i(\bar{\chi}, \chi) = -(\bar{\chi} - \chi)^i + \frac{1}{2} \Gamma_{jk}^i(\bar{\chi}) (\bar{\chi} - \chi)^j (\bar{\chi} - \chi)^k + \dots \end{aligned} \quad (\text{B.26})$$

One sees that  $\xi_{\mathfrak{h}}$  transforms like a vector with respect to local reparameterizations of the background field  $\bar{\chi}$  and like a scalar with respect to reparameterizations of  $\chi$ . Later on we shall need to relate functional derivatives with respect to the different fields. The relevant relations are

$$\begin{aligned} \frac{\delta}{\delta\chi^i} &= D^j{}_i(\bar{\chi}, \xi) \frac{\delta}{\delta\xi^j}, & D^j{}_i(\bar{\chi}, \xi) &:= \left( \frac{\delta\xi_{\mathfrak{h}}^j}{\delta\chi^i} + \Gamma_{ik}^j(\bar{\chi}) \xi_{\mathfrak{h}}^k \right) (\bar{\chi}, \chi_{\mathfrak{h}}(\bar{\chi}, \xi)), \\ \frac{\delta}{\delta\bar{\chi}^i} &= C^j{}_i(\bar{\chi}, \xi) \frac{\delta}{\delta\xi^j}, & C^j{}_i(\bar{\chi}, \xi) &:= \frac{\delta\xi_{\mathfrak{h}}^j}{\delta\bar{\chi}^i}(\bar{\chi}, \chi_{\mathfrak{h}}(\bar{\chi}, \xi)). \end{aligned} \quad (\text{B.27})$$

One can also verify that for a function  $F(\bar{\chi}, \xi)$  an  $n$ -th ordinary derivative with respect to  $\xi$ 's becomes a symmetrized covariant derivative with respect to  $\chi$ 's with  $F(\bar{\chi}, \xi) = F(\bar{\chi}, \xi_{\mathfrak{h}}(\bar{\chi}, \chi))$  viewed as a function of  $\bar{\chi}, \chi$ . Finally the invariant measures are related by

$$\mathcal{D}\chi := \prod_x d\chi^i \sqrt{\det \mathfrak{h}(\chi)} = \prod_x d\xi^i \sqrt{\det \mathfrak{h}(\bar{\chi})} \left( 1 - \frac{1}{6} R_{ij}(\bar{\chi}) \xi^i \xi^j + \dots \right) =: \mathcal{D}\xi. \quad (\text{B.28})$$

With these geometrical properties at hand it is now straightforward to construct generating functionals which are reparameterization invariant functionals of their arguments. Starting with  $W[J, \bar{\chi}]$  in Equation (B.13) (with  $S$  being not explicitly background dependent) one simply replaces the linear but noncovariant source coupling  $(\chi - \bar{\chi})J$  by the covariant but nonlinear one  $\xi_{\mathfrak{h}}^i(\bar{\chi}, \chi) J_i(\bar{\chi})$ , where  $J_i(\bar{\chi})$  is a co-vector with respect to background field reparameterizations. This gives

$$\exp W[J; \bar{\chi}] = \int \mathcal{D}\chi \exp \left\{ -S_{\mathfrak{h}}[\chi] + \int dx \xi_{\mathfrak{h}}^i(\bar{\chi}, \chi) J_i(\bar{\chi}) \right\} = \int \mathcal{D}\xi \exp \left\{ -S_{\mathfrak{h}}[\xi, \bar{\chi}] + \int dx \xi^i J_i(\bar{\chi}) \right\}. \quad (\text{B.29})$$



The first expression shows that the dependence on the background configuration  $\bar{\chi}$  enters only through the source term; in the second expression we changed variables according to  $\chi = \chi_{\mathfrak{b}}(\bar{\chi}, \xi)$ . Correspondingly there are two ways to introduce a mean fluctuation field

$$\langle \xi^i \rangle_J = \frac{\delta W[J; \bar{\chi}]}{\delta J_i} = \langle \xi_{\mathfrak{b}}^i(\bar{\chi}, \cdot) \rangle_J =: \xi_{\mathfrak{b}}^i(\bar{\chi}, \langle \chi \rangle_J), \quad (\text{B.30})$$

which treat  $\xi$  and  $\chi$  as the dynamical field, respectively, with mean fields  $\langle \xi \rangle_J$  and  $\langle \chi \rangle_J$ . Both are functions of  $\bar{\chi}$ , and  $\langle \chi \rangle_J$  is implicitly defined by Equation (B.30). The Legendre transforms are

$$\begin{aligned} \Gamma[\langle \chi \rangle; \bar{\chi}] &= \sup_J \left\{ \int dx \xi_{\mathfrak{b}}^i(\bar{\chi}, \langle \chi \rangle) J_i(\bar{\chi}) - W[J, \bar{\chi}] \right\}, \\ \Gamma[\langle \xi \rangle; \bar{\chi}] &= \sup_J \left\{ \int dx \xi^i J_i(\bar{\chi}) - W[J, \bar{\chi}] \right\}, \end{aligned} \quad (\text{B.31})$$

respectively, which gives for the extremizing configurations

$$J_{*,i}(\langle \chi \rangle, \bar{\chi}) = (D^{-1})_i{}^j(\bar{\chi}, \langle \chi \rangle) \frac{\delta \Gamma[\langle \chi \rangle; \bar{\chi}]}{\delta \langle \chi^j \rangle}, \quad J_{*,i}(\langle \xi \rangle, \bar{\chi}) = \frac{\delta \Gamma[\langle \xi \rangle; \bar{\chi}]}{\delta \langle \xi^i \rangle}. \quad (\text{B.32})$$

The Legendre transforms satisfy the functional integro-differential equations

$$\exp -\Gamma[\langle \chi \rangle; \bar{\chi}] = \int \mathcal{D}\chi \exp \left\{ -S_{\mathfrak{b}}[\chi] + \int dx [\xi_{\mathfrak{b}}^i(\bar{\chi}, \chi) - \xi_{\mathfrak{b}}^i(\bar{\chi}, \langle \chi \rangle)] (D^{-1})_i{}^j \frac{\delta \Gamma}{\delta \langle \chi^j \rangle} \right\}, \quad (\text{B.33})$$

$$\exp -\Gamma[\langle \xi \rangle; \bar{\chi}] = \int \mathcal{D}\xi \exp \left\{ -S_{\mathfrak{b}}[\xi, \bar{\chi}] + \int dx [\xi^i - \langle \xi^i \rangle] \frac{\delta \Gamma}{\delta \langle \xi^i \rangle} \right\}, \quad (\text{B.34})$$

with the arguments of the source terms as in Equation (B.32). The reason for carrying both variants along is to highlight that the effective action based on a geodesic split is at least formally a coordinate independent concept, where patches in field space around a reference configuration are invariantly described. Technically the use of the normal coordinate field  $\xi$  as a dynamical variable is simpler but presumably not indispensable. The  $n$ -point functions are related by [51]

$$\Gamma_{;(i_1 \dots i_n)}[\langle \chi \rangle, \bar{\chi}] = \langle \xi^{i_1} \dots \xi^{i_n} \rangle, \quad (\text{B.35})$$

where the left-hand-side is the symmetrized covariant derivative with respect to  $\langle \chi \rangle$  (as in  $D$  in Equation (B.27) for  $n = 1$ ) and the right-hand-side is the vertex (“1-particle irreducible”) function of the normal coordinate fields.

For later reference let us also note that once in  $\Gamma[\langle \xi \rangle; \bar{\chi}]$  the source-free condition  $J_*[\langle \xi \rangle; \bar{\chi}] = 0$  is imposed, with  $\langle \xi \rangle_J = 0$  as the only solution within formal power series inversion, one has

$$\bar{\Gamma}[\bar{\chi}] := \Gamma[0, \bar{\chi}] = -W[0, \bar{\chi}], \quad (\text{B.36})$$

which is the counterpart of Equation (B.17). Comparing with Equation (B.29), one sees that the functional integral for  $\bar{\Gamma}[\bar{\chi}]$  reads

$$\exp -\bar{\Gamma}[\bar{\chi}] = \int \mathcal{D}\xi \exp \left\{ -S_{\mathfrak{b}}[\xi, \bar{\chi}] + \int dx \xi^i(x) J_i(\bar{\chi})(x) \right\}, \quad (\text{B.37})$$

where the source  $J_i(\bar{\chi})$  is constrained by the requirement that  $\langle \xi^i \rangle_J$  (given by the same functional integral with a  $\xi^i$  insertion) vanishes. For generalized Riemannian sigma-models the perturbative renormalization of this generating functional is summarized in Appendix B.3.

We now describe two types of Ward identities for these systems: diffeomorphism type Ward identities and the nonlinear splitting Ward identities mentioned earlier. The former are really generalized Ward identities relating different theories in the sense that a compensating change in the metric tensor  $\mathfrak{h}_{ij}$  is needed. On the classical level an example is Equation (B.24); only if  $\mathfrak{h}_{ij}$  admits Killing vectors and one takes for  $v^i$  one of the Killing vectors does Equation (B.24) reduce to a conservation equation proper, with  $\mathcal{C}_\mu$  being the associated Noether current. In fact, Equation (B.24) can be promoted to a ‘‘Diffeomorphism Ward identity’’ in the quantum theory, at least perturbatively [201, 162] and presumably also in a non-perturbative formulation. In perturbation theory the ‘‘equations of motion operator’’ appearing on the right hand side of Equation (B.24) is a finite operator, i.e. it is the same when viewed as a functional of the bare fields and couplings, and when viewed as a function of the renormalized fields and couplings. Once the second term on the left-hand-side has been defined in terms of normal products, the diffeomorphism current  $\mathcal{C}_\mu$  must be finite as well.

For the background effective action  $\Gamma[\langle\xi\rangle, \bar{\chi}]$  a similar diffeomorphism Ward identity arises as follows. Since a geodesic is a coordinate independent concept, transforming all of the ingredients in the definition of  $\chi^i = \chi_{\mathfrak{h}}^i(\bar{\chi}, \xi)$  by an infinitesimal diffeomorphism  $V : \mathcal{M} \rightarrow \mathcal{M}$ ,  $V = \text{id} + v + O(v^2)$ , gives

$$\chi^i + v^i(\chi) = \chi_{\mathfrak{h} - \mathcal{L}_v \mathfrak{h}}^i(\bar{\chi}^j + v^j(\bar{\chi}), \xi^j + \partial_k v^j(\bar{\chi})\xi^k). \quad (\text{B.38})$$

Assuming that the functional integral in Equation (B.29) has been invariantly regularized, and noting that the quantum field in the symmetry variation  $\delta_v \bar{\chi}^i = v^i(\bar{\chi})$ ,  $\delta_v \mathfrak{h}_{ij} = -(\mathcal{L}_v \mathfrak{h})_{ij}$ ,  $\delta \xi^i = \partial_j v^i \xi^j$  enters linearly, one is lead to the following diffeomorphism Ward identity for  $\Gamma[\langle\xi\rangle, \bar{\chi}]$  [110]:

$$(\mathcal{L}_v \mathfrak{h})_{ij} \frac{\partial \Gamma}{\partial \mathfrak{h}_{ij}} - v^i \frac{\delta \Gamma}{\delta \bar{\chi}^i} - \partial_j v^i \langle \xi^j \rangle \frac{\delta \Gamma}{\delta \langle \xi^i \rangle} = 0. \quad (\text{B.39})$$

The ‘price’ for the linearity here is that the metric tensor changes as well.

This is different in the nonlinear splitting Ward identities, which control the dependence on the background field configuration. If the metric in Equation (B.38) is kept fixed the splitting symmetry  $\delta_v \bar{\chi}^i = v^i$ ,  $\delta_v \xi^i = C^i_j(\bar{\chi}, \xi)v^j$  becomes nonlinear (see Equation (B.27)). The corresponding ‘‘nonlinear splitting Ward identity’’ reads as follows:

$$\frac{\delta \Gamma[\langle\xi\rangle; \bar{\chi}]}{\delta \bar{\chi}^i} + \langle C^j_i(\bar{\chi}, \cdot) \rangle \frac{\delta \Gamma[\langle\xi\rangle; \bar{\chi}]}{\delta \langle \xi^j \rangle} = 0. \quad (\text{B.40})$$

Here  $\langle C^j_i(\bar{\chi}, \xi) \rangle$  is the average of the matrix field  $C(\bar{\chi}, \xi)$  in Equation (B.27) defined as a renormalized composite operator and the variations are at fixed  $\langle \chi \rangle$ . Conceptually Equation (B.40) expresses the fact that in Equation (B.33) the dependence on the background  $\bar{\chi}$  enters only through the source term. Differentiating this equation with respect to the background at fixed  $\langle \chi \rangle$  and switching to the normal coordinate field as dynamical variable one finds Equation (B.40), if somewhat formally. A technically cleaner way of arriving at Equation (B.40) is by working with the normal coordinate field throughout, promoting the splitting invariance to a nilpotent operation and defining  $\langle C^j_i(\bar{\chi}, \cdot) \rangle$  by variation with respect to a source. For details be refer to Howe et al. [110]. Further nonlinear renormalizations of  $\xi$  itself have to be taken into account (see Appendix B.3 for a summary). Note that for a flat geometry on field space the equation reduces to Equation (B.17). In a perturbative construction the nonlinear splitting Ward identity (B.40) has the important consequence that the counterterms depend on the full field  $\chi = \chi_{\mathfrak{h}}(\bar{\chi}, \xi)$  only and not on  $\bar{\chi}$ ,  $\xi$  individually. The use of normal coordinates in Equation (B.40) has technical advantages, in principle however a counterpart for the  $\Gamma[\langle\chi\rangle, \bar{\chi}]$  effective action exists as well [51].

### B.2.3 Linear split, gauge theories

Since we are interested here in diffeomorphism invariant theories we consider this type of gauge invariance. The case of Yang–Mills fields is largely parallel (see [108, 1, 25] for the latter). Let  $S[g]$  be any diffeomorphism invariant action of a Riemannian metric  $g = (g_{\alpha\beta})_{1 \leq \alpha, \beta \leq d}$ . Infinitesimally the invariance reads  $S[g + \mathcal{L}_v g] = S[g] + O(v^2)$ , where

$$\mathcal{L}_v g_{\alpha\beta} = v^\gamma \partial_\gamma g_{\alpha\beta} + \partial_\alpha v^\gamma g_{\gamma\beta} + \partial_\beta v^\gamma g_{\gamma\alpha} \quad (\text{B.41})$$

is the Lie derivative of  $g$  with respect to the vector field  $v^\gamma \partial_\gamma$ . As before we decompose the metric  $g$  (later the integration variable in the functional integral) into a background  $\bar{g}$  and a fluctuation  $f$ , i.e.  $g_{\alpha\beta} = \bar{g}_{\alpha\beta} + f_{\alpha\beta}$ . Note that  $f_{\alpha\beta}$  is *not* assumed to be small in some sense, no expansion in powers of  $f$  is implied by the split. Note however that this linear split does not have a geometrical meaning in the space of geometries.

The symmetry variation  $g_{\alpha\beta} \mapsto g_{\alpha\beta} + \mathcal{L}_v g_{\alpha\beta}$  can be decomposed in two different ways,

$$f_{\alpha\beta} \mapsto f_{\alpha\beta} + \mathcal{L}_v (\bar{g} + f)_{\alpha\beta}, \quad \bar{g}_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta}, \quad (\text{B.42})$$

$$f_{\alpha\beta} \mapsto f_{\alpha\beta} + \mathcal{L}_v f_{\alpha\beta}, \quad \bar{g}_{\alpha\beta} \mapsto \bar{g}_{\alpha\beta} + \mathcal{L}_v \bar{g}_{\alpha\beta}. \quad (\text{B.43})$$

We shall refer to the first one as “genuine gauge transformations” and to the second one as the “background gauge transformations”. The background effective action  $\Gamma[\langle g_{\alpha\beta} \rangle, \bar{g}_{\alpha\beta}, \text{ghosts evs.}]$  to be constructed will be a functional of the averaged metric  $\langle g_{\alpha\beta} \rangle$ , the background metric  $\bar{g}_{\alpha\beta}$ , and sources, which is invariant under the background field transformations (B.43). We quickly run through the relevant steps.

The background generating functional  $W[J, j, \bar{j}; \bar{g}]$  is formally defined by

$$W[J, j, \bar{j}; \bar{g}] = \int \mathcal{D}\mu_{j, \bar{j}}[f] \exp \left\{ -S[\bar{g} + f] + \int dx \sqrt{\bar{g}} J^{\alpha\beta} f_{\alpha\beta} \right\}. \quad (\text{B.44})$$

The measure  $\mathcal{D}\mu_{j, \bar{j}}[f]$  differs from the naive measure  $\mathcal{D}f_{\alpha\beta}$  by gauge-fixing and ghost contributions, as well as sources  $j^\alpha, \bar{j}_\alpha$  for the ghosts:

$$\begin{aligned} \mathcal{D}\mu_{j, \bar{j}}[f] = \mathcal{D}f_{\alpha\beta} \int \mathcal{D}C^\alpha \mathcal{D}\bar{C}_\alpha \exp \left\{ -\frac{1}{2\alpha} \int dx \sqrt{\bar{g}} \bar{g}^{\alpha\beta} Q_\alpha Q_\beta + \frac{1}{\kappa} \int dx \bar{C}_\alpha \bar{g}^{\alpha\beta} \frac{\partial Q_\beta}{\partial f_{\gamma\delta}} \mathcal{L}_C(\bar{g} + f)_{\gamma\delta} \right. \\ \left. + \int dx \sqrt{\bar{g}} [\bar{j}_\alpha C^\alpha + j^\alpha \bar{C}_\alpha] \right\}. \end{aligned} \quad (\text{B.45})$$

Here  $Q_\alpha = Q_\alpha[\bar{g}; f] \approx 0$  is the gauge fixing condition, which must be invariant under Equation (B.43), but which we may leave unspecified here. A widely used gauge condition is the background harmonic gauge (see Section 4.1). We shall ignore the problem of global existence of gauge slices in accordance with the formal nature of the construction. The second term in the exponent is the ghost part; it is obtained along the same as lines in Yang–Mills theory: One applies a genuine gauge transformation (B.42) to  $Q_\alpha$  and replaces the parameters  $v^\alpha$  by the ghost field  $C^\alpha$ . The integral over the ghosts  $C^\alpha$  and  $\bar{C}_\alpha$  exponentiates the Faddeev–Popov determinant  $\det[\delta Q_\alpha / \delta v^\beta]$  then. It is crucial that the ghost and gauge fixing terms are invariant under the background field transformations (B.43) together with

$$\delta C^\alpha = \mathcal{L}_v C^\alpha, \quad \delta \bar{C}_\alpha = \mathcal{L}_v \bar{C}_\alpha. \quad (\text{B.46})$$

Finally we coupled in Equation (B.45) the ghosts to sources  $j^\alpha, \bar{j}_\alpha$ , for later use. The key fact now is the invariance

$$W[\mathcal{J} + \mathcal{L}_v \mathcal{J}] = W[\mathcal{J}], \quad \mathcal{J} := \{J^{\alpha\beta}, j^\alpha, \bar{j}_\alpha, \bar{g}_{\alpha\beta}\}, \quad (\text{B.47})$$

where  $\mathcal{L}_v$  is the Lie derivative of the respective tensor type. This invariance property follows from Equation (B.45) if one performs a compensating transformation (B.43, B.46) on the integration variables  $f_{\alpha\beta}$ ,  $C^\alpha$ ,  $\bar{C}_\alpha$ , and uses the invariance of all but the source terms. Importantly, at this point one must assume that the formal measure  $\mathcal{D}f_{\alpha\beta}$  is diffeomorphism invariant.

The background effective action now is defined by

$$\Gamma[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] = \sup_{J, \bar{J}, \bar{j}} \left\{ \int dx \sqrt{\bar{g}} (\bar{f}_{\alpha\beta} J^{\alpha\beta} + \sigma^\alpha \bar{J}_\alpha + \bar{\sigma}_\alpha J^\alpha) - W[J, \bar{J}, \bar{j}; \bar{g}] \right\}. \quad (\text{B.48})$$

The extremizing source configurations  $J_*^{\alpha\beta}$ ,  $\bar{J}_{*,\alpha}$ ,  $J_*^\alpha$  are characterized by

$$\begin{aligned} \frac{\delta}{\delta \bar{f}_{\alpha\beta}} \Gamma[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] &= J_*^{\alpha\beta}[\bar{f}, \sigma, \bar{\sigma}; \bar{g}], \\ \frac{\delta}{\delta \sigma^\alpha} \Gamma[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] &= J_*^\alpha[\bar{f}, \sigma, \bar{\sigma}; \bar{g}], \\ \frac{\delta}{\delta \bar{\sigma}_\alpha} \Gamma[\bar{f}, \sigma, \bar{\sigma}; \bar{g}] &= \bar{J}_{*,\alpha}[\bar{f}, \sigma, \bar{\sigma}; \bar{g}]. \end{aligned} \quad (\text{B.49})$$

As usual, if  $W$  is differentiable everywhere with respect to the sources one can interpret  $\bar{f}_{\alpha\beta}$ ,  $\sigma^\alpha$ ,  $\bar{\sigma}_\alpha$  as the source dependent expectation values of  $f_{\alpha\beta}$ ,  $C^\alpha$ ,  $\bar{C}_\alpha$  via

$$\bar{f}_{\alpha\beta} = \langle f_{\alpha\beta} \rangle = \frac{1}{\sqrt{\bar{g}}} \frac{\delta W}{\delta J_*^{\alpha\beta}}, \quad \sigma^\alpha = \langle C^\alpha \rangle = \frac{1}{\sqrt{\bar{g}}} \frac{\delta W}{\delta J_{*,\alpha}}, \quad \bar{\sigma}_\alpha = \langle \bar{C}_\alpha \rangle = \frac{1}{\sqrt{\bar{g}}} \frac{\delta W}{\delta J_*^\alpha}, \quad (\text{B.50})$$

and Equations (B.49, B.50) are related by formal inversion. Indeed, the extremizing sources are usually constructed by assuming that  $W_k$  has a series expansion in powers of the sources with  $\bar{g}$ -dependent coefficients; formal inversion of the series then gives a  $J_*^{\alpha\beta}[\bar{f}, \sigma, \bar{\sigma}; \bar{g}]$  with the property that  $J_*^{\alpha\beta}[0, \sigma, \bar{\sigma}; \bar{g}] = 0$ , and similarly for the ghost sources.

It is convenient to regard  $\Gamma$  as a functional of  $\bar{g}_{\alpha\beta}$  and  $g_{\alpha\beta} := \bar{g}_{\alpha\beta} + \bar{f}_{\alpha\beta}$  instead of  $\bar{g}_{\alpha\beta}$  and  $\bar{f}_{\alpha\beta}$ . We thus set

$$\Gamma[g_{\alpha\beta}, \bar{g}_{\alpha\beta}, \sigma^\alpha, \bar{\sigma}_\alpha] := \Gamma[g_{\alpha\beta} - \bar{g}_{\alpha\beta}, \sigma^\alpha, \bar{\sigma}_\alpha; \bar{g}_{\alpha\beta}]. \quad (\text{B.51})$$

The crucial invariance property can then be summarized as

$$\Gamma[\Phi + \mathcal{L}_v \Phi] = \Gamma[\Phi], \quad \Phi := \{g_{\alpha\beta}, \bar{g}_{\alpha\beta}, \sigma^\alpha, \bar{\sigma}_\alpha\}, \quad (\text{B.52})$$

where all its arguments transform as tensors of the corresponding rank. Equation (B.52) is a direct consequence of Equation (B.47).

Finally we have to switch off the sources. Since  $\Gamma$  has ghost number zero  $\sigma^\alpha = 0$ ,  $\bar{\sigma}_\alpha = 0$  is always a solution of Equation (B.49). For the metric source  $J_*^{\alpha\beta}$  this is different. Within the realm of formal power series inversions  $\bar{f}_{\alpha\beta} \equiv 0$  is always a solution of  $J_*^{\alpha\beta} \equiv 0$ . Combined with the usual uniqueness assumption it is the only solution, and the ‘‘self-consistent background determination’’ at which the background field method aims at degenerates. Indeed, note that in  $\langle f_{\alpha\beta} \rangle_{J_*, J_*, \bar{J}_*}$  the  $\bar{g}$  dependence drops out, as  $\bar{f}_{\alpha\beta}$  is prescribed. The expectation value of the full quantum metric,  $q_{\alpha\beta} := f_{\alpha\beta} + \bar{g}_{\alpha\beta}$ , say, just gives back the prescribed background  $\langle q_{\alpha\beta} \rangle_{J_*, J_*, \bar{J}_*} = \bar{g}_{\alpha\beta}$ . This evidently has a somewhat perturbative flavor, although no direct reference to perturbation theory is made.

To go beyond that, we directly impose

$$\frac{\delta}{\delta g_{\alpha\beta}} \Gamma[g, \bar{g}, 0, 0] = 0 \quad \text{iff} \quad \bar{g}_{\alpha\beta} = \bar{g}_*[g]_{\alpha\beta}, \quad (\text{B.53})$$

as the source-free condition. It adjusts the background  $\bar{g}_{\alpha\beta}$  used as a reference in the functional integral self-consistently to the dynamical  $g_{\alpha\beta}$ , which is the average  $\langle q_{\alpha\beta} \rangle$  of the quantum metric.

Equation (B.53) can be viewed as the vanishing of the one-point function for all  $g$ . The multi-point functions are defined by

$$\Gamma^{(n)}(x_1, \dots, x_n; g) = \frac{\delta}{\delta g(x_1)} \cdots \frac{\delta}{\delta g(x_n)} \Gamma[g, \bar{g}, 0, 0] \Big|_{\bar{g}=\bar{g}_*[g]}. \quad (\text{B.54})$$

They should contain the complete information about the system, including the state (see Section 2.3). Their precise physics meaning however remains to be understood. In the case of gauge theories which allow for a perturbative definition of scattering states, the background effective action is known to produce the same perturbative S-matrix as the standard effective action [1, 25].

In view of Equation (B.53, B.54) one would like to introduce a functional  $\bar{\Gamma}$  of a single metric only, whose one-point functions have stationary points. This can be done as follows. We define the final “background effective action” [156] by

$$\bar{\Gamma}[g] := \Gamma[g, \bar{g}_*[g], \sigma = 0, \bar{\sigma} = 0]. \quad (\text{B.55})$$

For the solutions of Equation (B.53) obtained by formal power series inversion one has  $\bar{g}_*[g]_{\alpha\beta} = g_{\alpha\beta}$  and thus recovers the usual definition. Generally Equation (B.55) is invariant under  $g_{\alpha\beta} \mapsto g_{\alpha\beta} + \mathcal{L}_v g_{\alpha\beta}$ , as desired, and obeys

$$\frac{\delta \bar{\Gamma}[g]}{\delta g_{\alpha\beta}} = 0. \quad (\text{B.56})$$

#### B.2.4 Geodesic split, gauge theories

The background gauge invariant effective action in Section B.2.3 depend parametrically on the choice of the background gauge condition  $Q^a$ . This dependence is a consequence of the field parameterization dependence of the effective actions based on a linear fluctuation background split. The geometrical approach of Vilkovisky and deWitt [68, 223] is designed to overcome this drawback and at least formally it produces an off-shell effective action with all the desirable properties: It is gauge invariant with respect to the background field, gauge invariant with respect to the mean of the fluctuation field, and independent of the choice of the gauge fixing surface. In brief, the strategy is to project locally onto the gauge invariant subspace and then apply the techniques of Section B.2.2. Since we shall not use this formulation here it may suffice to refer to [129] for a brief survey and to [178, 68, 223] for detailed expositions.

With these remarks we conclude our brief introduction to the background field formalism. In the context of the asymptotic safety scenario the variant from Section B.2.2 has been used in [154, 155] (see Section 3) and the variant from Section B.2.4 in [179, 133, 131] (see Section 4).

### B.3 Renormalization of Riemannian sigma-models

Here we summarize the results on the UV renormalization of Riemannian sigma-models needed for Section 3. We largely follow the thorough treatment of Osborn [162] and use the results of [84, 201, 57, 110, 220, 80].

The goal will be to construct the renormalized effective action based on a geodesic background-fluctuation split and as discussed in Section 3.2.2. The variant where the normal coordinate field is the dynamical variable is used as well as dimensional regularization. The latter has the advantage that the additional curvature dependent terms in the  $\mathcal{D}\xi$  measure in Equation (B.28) do not contribute. The explicit counterterms have been computed in minimal subtraction and all scheme dependent quantities will refer to this scheme. The dynamical scalar fields will be denoted by  $\phi^i$ , the background fields by  $\bar{\phi}^i$  and the normal coordinate fields by  $\xi^i$ .

For the purposes of renormalization it is useful to consider an extended Lagrangian of the form

$$\lambda L(\phi; \mathcal{G}) = \frac{1}{2} \bar{\gamma}^{\mu\nu} \mathfrak{h}_{ij} \partial_\mu \phi^i \partial_\nu \phi^j + \bar{\gamma}^{\mu\nu} \partial_\mu \phi^i V_{\nu i} + \frac{1}{2} \bar{R}^{(2)}(\bar{\gamma}) D + F. \quad (\text{B.57})$$

Here  $\mathcal{G} = \{\mathfrak{h}, V, D, F\}$ ,  $\mathcal{G} = \mathcal{G}(\phi; x)$ , is a collection of generalized couplings/sources (of the tensor type indicated by the index structure) that depend both on the fields  $\phi^j$  and explicitly on the point  $x$  in the “base space”. The latter is a two-dimensional Riemannian space with a classical background metric  $\bar{\gamma}_{\mu\nu}(x)$ , extended to  $d$  dimensions in the sense of dimensional regularization, and  $\bar{R}^{(2)}(\bar{\gamma}) = R^{(2)}(\bar{\gamma})/(d-1)$ . The action functional is  $S[\phi; \mathcal{G}] = \int d^d x \sqrt{\bar{\gamma}} L(\phi; \mathcal{G})$ . The explicit  $x$ -dependence of the sources  $\mathcal{G}$  allows one to define local composite operators via functional differentiation after renormalization. In addition the scalar source  $F$  provides an elegant way to compute the nonlinear renormalizations of the quantum fields in the background expansion [110].

Recall that the geodesic background-fluctuation split involves decomposing the fields  $\phi^j$  into a background field configuration  $\bar{\phi}^j$  and a power series in the fluctuation fields  $\xi^j$  whose coefficients are functions of  $\bar{\phi}^j$ . The series is defined in terms of the unique geodesic  $s \mapsto \Sigma(\bar{\phi}, \xi; s)$  from the point  $\bar{\phi}$  (with  $s = 0$ , say) in the target manifold to the (nearby) point  $\phi$  (with  $s = 1$ ), where  $\xi^j$  is the tangent to  $\Sigma^j$  vector at  $\bar{\phi}$ . We shall write  $\phi^j = \Sigma^j(\bar{\phi}; \xi, 1)$  for this series, and refer to  $\phi$ ,  $\bar{\phi}$ , and  $\xi$  as the full field, the background field, and the quantum field, respectively. On the bare level one starts with fields  $\phi_B^j := \Sigma^j(\bar{\phi}_B, \xi_B; 1)$  which upon renormalization are replaced by  $\phi^j = \Sigma^j(\bar{\phi}, \xi; 1)$ . The expansion of the Lagrangian (B.57) can be computed from

$$L(\phi; \mathcal{G}) = \sum_{n \geq 0} \frac{1}{n!} \frac{d^n}{ds^n} L(\Sigma(\bar{\phi}, \xi; s); \mathcal{G}) \Big|_{s=0}. \quad (\text{B.58})$$

The  $s$  derivatives can be reduced to background covariant derivatives  $\bar{D}_\mu \xi^i = \partial_\mu \xi^i + \Gamma_{jk}^i(\bar{\phi}) \partial_\mu \bar{\phi}^j \xi^k$ , where  $\Gamma_{jk}^i(\bar{\phi})$  are the Christoffel symbols of  $\mathfrak{h}$  evaluated at  $\bar{\phi}$ . One easily verifies that along with  $\xi^i$  and  $\partial_\mu \bar{\chi}^i$  also  $\bar{D}_\mu \xi^i := \partial_\mu \xi^i + \Gamma_{jk}^i(\bar{\chi}) \partial_\mu \bar{\chi}^j \xi^k$  transform as vectors under background field transformations. Hence all monomials built from a covariant tensor  $T_{ij}(\mathfrak{h})$  in  $\mathfrak{h}$ ,

$$T_{ij}(V_* \mathfrak{h})(\hat{\phi}) = \frac{\partial \phi^k}{\partial \hat{\phi}^i} \frac{\partial \phi^l}{\partial \hat{\phi}^j} T_{kl}(\mathfrak{h})(\phi), \quad \hat{\phi} = V(\phi), \quad (\text{B.59})$$

by contracting it with combinations of  $\xi^i$ ,  $\partial_\mu \bar{\phi}^i$ ,  $D_\mu \xi^i$ , will be invariant under background field transformations. These are precisely the monomials entering the expansion of a reparameterization invariant Lagrangian like  $L(\phi; \mathcal{G})$ . For example for the metric part  $L_{\mathfrak{h}}(\chi) = \frac{1}{2} \mathfrak{h}_{ij}(\chi) \partial_\mu \chi^i \partial^\mu \chi^j$  one finds from Equation (B.58)

$$\begin{aligned} L_{\mathfrak{h}}(\phi) &= L_{\mathfrak{h}}(\bar{\phi}) + \sum_{n \geq 1} L_n(\bar{\phi}, \xi), \\ L_1(\bar{\phi}, \xi) &= \mathfrak{h}_{ij}(\bar{\phi}) \bar{\gamma}^{\mu\nu} \partial_\mu \bar{\phi}^i \bar{D}_\nu \xi^j, \\ L_2(\bar{\phi}, \xi) &= \frac{1}{2} \bar{\gamma}^{\mu\nu} [\mathfrak{h}_{ij}(\bar{\phi}) \bar{D}_\mu \xi^i \bar{D}_\nu \xi^j + R_{ijkl}(\bar{\phi}) \partial_\mu \bar{\phi}^i \partial_\nu \bar{\phi}^j \xi^k \xi^l], \end{aligned} \quad (\text{B.60})$$

etc., with  $L_n$  of order  $n$  in  $\xi$ . Here  $R_{ijkl}$  is the Riemann tensor of  $\mathfrak{h}$  evaluated at  $\bar{\phi}$ .

Both  $\xi^i$  and  $\bar{D}_\mu \xi^i$  transform as vectors under reparameterizations of the background fields  $\bar{\phi} \mapsto V(\bar{\phi})$ . The  $n$ -th order term  $L_n$  in Equation (B.58) is of order  $n$  in  $\xi$  and at most quadratic in  $\bar{D}_\mu \xi$ . The generalized couplings/sources  $\mathcal{G}$  are evaluated at  $\bar{\phi}$  and transform according to their respective tensor type under  $\bar{\phi} \mapsto V(\bar{\phi})$ . In total this renders each term in Equation (B.58) individually invariant under these background field transformations. The term  $L_2$  quadratic in  $\xi$  is used to define the propagators, all other terms are treated as interactions. Due to the expansion around a nontrivial background field configuration  $L_2$  does not have a standard kinetic term,

e.g. the first term in Equation (B.57) gives rise to  $\frac{1}{2}\bar{\gamma}^{\mu\nu}\mathfrak{h}_{ij}(\bar{\phi})\bar{D}_\mu\xi^i\bar{D}_\nu\xi^j + \dots$ . To get standard kinetic terms one introduces the components  $\xi^a := E_i^a(\bar{\phi})\xi^i$ , with respect to a frame field satisfying  $\mathfrak{h}_{ij}(\bar{\phi}) = E_i^a(\bar{\phi})E_j^b(\bar{\phi})\eta_{ab}$ . Rewriting  $L_2$  in terms of the frame fields  $\xi^a$  a potential is generated which combines with the other terms in  $L_2$ , but the  $\xi^a$  have a standard kinetic term and are formally massless. These fields are given a mass  $\mu$  which sets the renormalization scale. The development of perturbation theory then by-and-large follows the familiar lines, the main complication comes from the complexity of the interaction Lagrangians  $L_n$ ,  $n \geq 3$ . In addition nonlinear field renormalizations are required; the transition function  $\xi \mapsto \xi_B(\xi)$  can be computed from the differential operator  $Z - 1$  in Equation (B.57) below. For our purposes we in addition have to allow for a renormalization  $\bar{\phi}_B(\bar{\phi})$  of the background fields. As usual we adopt the convention that the fields  $\phi_B^j$  remain dimensionless for base space dimension  $d \neq 2$ .

With this setting the bare couplings/sources  $\mathcal{G}^B(\phi_B; x)$  have dimension  $[\mu]^{d-2}$  and are expressed as a dimensionless sum of the renormalized  $\mathcal{G}(\phi; x)$  and covariant counter tensors built from  $\mathcal{G}(\phi; x)$ . A suitable parameterization is

$$\begin{aligned} \mathfrak{h}_{ij}^B &= \mu^{d-2} [\mathfrak{h}_{ij} + T_{ij}(\mathfrak{h})], \\ V_{\mu i}^B &= \mu^{d-2} \left[ Z^V(\mathfrak{h})^j{}_i V_{\mu j} + N_i{}^{jk}(\mathfrak{h})\bar{\partial}_\mu \mathfrak{h}_{jk} \right], \\ D^B &= \mu^{d-2} [Z(\mathfrak{h})D + U(\mathfrak{h})], \\ F^B &= \mu^{d-2} [Z(\mathfrak{h})F + Y]. \end{aligned} \tag{B.61}$$

Here  $\bar{\partial}_\mu$  denotes differentiation with respect to  $x$  at fixed  $\phi$ . The quantities  $T_{ij}, N_i{}^{jk}, U, Y$  and  $Z^V - 1, Z - 1$  contain poles in  $(2 - d)$  (but no other type of singularities) whose coefficients are defined by minimal subtraction. Except for  $Y$  they depend on  $\mathfrak{h}_{ij}$  only;  $Y$  in addition depends quadratically on  $V_{\mu i}$  and  $\bar{\partial}_\mu \mathfrak{h}_{jk}$ , but the quadratic forms with which they are contracted again only depend on  $\mathfrak{h}_{ij}$ . All purely  $\mathfrak{h}$ -dependent counter tensors are algebraic functions of  $\mathfrak{h}_{ij}$ , its covariant derivatives, and its curvature tensors.  $Z - 1$  and  $Z^V - 1$  specifically are linear differential operators acting on scalars and vectors on the target manifold, respectively. The combined pole and loop expansion takes the form

$$\mathcal{O} = \sum_{\nu \geq 1} \sum_{l \geq \nu} \frac{1}{(2-d)^\nu} \left( \frac{1}{2\pi} \right)^l \mathcal{O}^{(\nu, l)} \tag{B.62}$$

for any of the quantities  $T_{ij}, N_i{}^{jk}, U, Z^V - 1, Z - 1, Y$ . The residue of the simple pole is denoted by  $\mathcal{O}^{(1)}$ . We do not include explicitly powers of the loop counting parameter  $\lambda$  in Equation (B.62). For the purely  $\mathfrak{h}$ -dependent counter terms of interest here they are easily restored by inserting  $\mathfrak{h}/\lambda$  and utilizing the scaling properties listed below. However once  $\mathfrak{h}$  is ‘deformed’ into a nontrivial function of  $\lambda$  the ‘scaling decomposition’ (B.62) no longer coincides with the expansion in powers of  $\lambda$  and the former is the fundamental one. Under a constant rescaling of the metric the purely  $\mathfrak{h}$ -dependent counter term coefficients transform homogeneously as follows

$$\begin{aligned} \mathcal{O}^{(\nu, l)}(\Lambda^{-1}\mathfrak{h}) &= \Lambda^{l-1} \mathcal{O}^{(\nu, l)}(\mathfrak{h}) && \text{for } \mathcal{O} = T_{ij}, U, \\ \mathcal{O}^{(\nu, l)}(\Lambda^{-1}\mathfrak{h}) &= \Lambda^l \mathcal{O}^{(\nu, l)}(\mathfrak{h}) && \text{for } \mathcal{O} = Z, Z^V, N. \end{aligned} \tag{B.63}$$

In principle the higher order pole terms  $\mathcal{O}^{(\nu, l)}$ ,  $l \geq \nu \geq 2$ , are determined recursively by the residues  $\mathcal{O}^{(1, l)}$  of the first order poles via ‘generalized pole equations’. The latter can be worked out in analogy to the quantum field theoretical case (see [57, 162]). Taking the consistency of the cancellations for granted one can focus on the residues of the first order poles, which we shall do throughout.



Explicit results for them are typically available up to and including two loops [84, 220, 57, 110, 162]. For the metric  $\mathfrak{h}$  and the dilaton  $D$  beta functions also the three-loop results are known\*:

$$\begin{aligned} T_{ij}^{(1,1)}(\mathfrak{h}) &= R_{ij}, \\ T_{ij}^{(1,2)}(\mathfrak{h}) &= \frac{1}{4} R_{iklm} R_j{}^{klm}, \\ T_{ij}^{(1,3)}(\mathfrak{h}) &= \frac{1}{6} R_{imn}{}^k R_{jpqk} R^{pnmq} - \frac{1}{8} R_{iklj} R^k{}_{mnp} R^{lmnp} - \frac{1}{12} \nabla_n R_{iklm} \nabla^k R_j{}^{lmn}, \end{aligned} \quad (\text{B.64})$$

where the three-loop term has been computed independently in [80] and [96]. For  $D$  the results are [162, 57, 220]

$$U^{(1,1)}(\mathfrak{h}) = \frac{c_T}{6}, \quad U^{(1,2)}(\mathfrak{h}) = 0, \quad U^{(1,3)}(\mathfrak{h}) = \frac{1}{48} R_{ijkl} R^{ijkl}, \quad (\text{B.65})$$

where  $c_T$  is the dimension of the target manifold. For the other quantities one has [162, 110, 220]

$$\begin{aligned} [Z^V(\mathfrak{h})_i^j]^{(1,1)} &= \frac{1}{2} [-\nabla^2 \delta_i^j + R_i{}^j], \\ [Z^V(\mathfrak{h})_i^j]^{(1,2)} &= \frac{1}{4} R_i{}^{klj} \nabla_k \nabla_l, \\ Z(\mathfrak{h})^{(1,1)} &= -\frac{1}{2} \nabla^2, \\ Z(\mathfrak{h})^{(1,2)} &= 0, \\ Z(\mathfrak{h})^{(1,3)} &= -\frac{3}{16} R^{iklm} R^j{}_{klm} \nabla_i \nabla_j, \\ [N_i{}^{jk}(\mathfrak{h})]^{(1,1)} &= \frac{1}{2} \delta_i^j \nabla^k - \frac{1}{4} \mathfrak{h}^{jk} \nabla_i, \\ [N_i{}^{jk}(\mathfrak{h})]^{(1,2)} &= \frac{1}{2} R_i{}^{jkl} \nabla_l. \end{aligned} \quad (\text{B.66})$$

The expressions for  $Y^{(1,1)}$  and  $Y^{(1,2)}$  are likewise known [162] but are not needed here.

Some explanatory comments should be added. First, in addition to the minimal subtraction scheme the above form of the counter tensors refers to the background field expansion in terms of Riemannian normal coordinates. If a different covariant expansion is used the counter tensors change. Likewise the standard form of the higher pole equations [57, 162] is only valid in a preferred scheme. For instance for the metric counter terms in this scheme additive contributions to  $T_{ij}(\mathfrak{h})$  of the form  $\mathcal{L}_V \mathfrak{h}_{ij}$  are absent [110]. Note that adding such a term for  $\nu = 1$  leaves the metric beta function in Equation (B.76) below unaffected, provided  $V^j$  is functionally independent of  $\mathfrak{h}_{ij}$ .

So far only the full fields entered,  $\phi_B^j$  on the bare and  $\phi^j$  on the renormalized level. Their split into background and quantum contributions is however likewise subject to renormalization. A convenient way to determine the transition function  $\xi_B^j(\xi)$  from the bare to the renormalized quantum fields was found by Howe, Papadopolous, and Stelle [110]. In effect one considers the inversion  $\xi^j(\bar{\phi}; \phi - \bar{\phi})$  of the normal coordinate expansion  $\phi^j = \phi^j(\bar{\phi}; \xi)$  of the renormalized fields. If  $Z$  in Equation (B.61, B.66) is regarded as a differential operator acting on the second argument of this function, i.e. on  $\phi$ ,

$$\xi^j(\xi_B) = Z \xi^j(\bar{\phi}; \phi - \bar{\phi}), \quad (\text{B.67})$$

---

\*Our conventions are:  $\nabla_i v^k = \partial_i v^k + \Gamma^k{}_{ij} v^j$ , with  $\Gamma^k{}_{ij} = \frac{1}{2} \mathfrak{h}^{kl} [\partial_j \mathfrak{h}_{il} + \partial_i \mathfrak{h}_{jl} - \partial_l \mathfrak{h}_{ij}]$ . The Riemann tensor is defined by  $(\nabla_i \nabla_j - \nabla_j \nabla_i) v^k = R^k{}_{lij} v^l$ , so that  $R^k{}_{lij} = \partial_i \Gamma^k{}_{lj} - \partial_j \Gamma^k{}_{li} + \Gamma^k{}_{im} \Gamma^m{}_{lj} - \Gamma^k{}_{jm} \Gamma^m{}_{li}$ . The Ricci tensor is  $R_{ij} = R^m{}_{imj}$ .

one obtains the desired  $\xi_B^j(\xi)$  relation by inversion. To lowest order  $Z^{(1,1)} = -\frac{1}{2}\nabla^2$  yields

$$\xi_B^i = \xi^i + \frac{1}{2-d} \frac{\lambda}{2\pi} \left[ \frac{1}{3} R^i_j \xi^j + \frac{1}{4} \nabla_k R^i_j \xi^k \xi^j - \frac{1}{24} \nabla^i R_{kj} \xi^k \xi^j + O(\xi^3) \right]. \quad (\text{B.68})$$

At each loop order the coefficient is a power series in  $\xi$  whose coefficients are covariant expressions built from the metric  $h_{ij}(\bar{\phi})$  at the background point.

With all these renormalizations performed the result can be summarized in the proposition [110, 162] that the source-extended background functional

$$\exp -\Gamma[\bar{\phi}; \mathcal{G}] = \int [\mathcal{D}\xi] \exp \left\{ -S[\mathcal{G}_B, \phi_B] + \frac{1}{\lambda} \int d^d x J_i(\bar{\phi}) \xi^i \right\} \quad (\text{B.69})$$

defines a finite perturbative measure to all orders of the loop expansion. The additional source  $J_i(\bar{\phi})$  here is constrained by the requirement that  $\langle \xi^j \rangle = 0$ . The key properties of  $\Gamma(\bar{\phi}; \mathcal{G})$  are:

- It is invariant under reparameterizations of the background fields  $\bar{\phi}$ .
- It obeys a simple renormalization group equation (which would not be true without the F-source).
- A generalized action principle holds that allows one to construct local composite operators of dimension 0, 1, 2, by variation with respect to the renormalized sources.

Let  $V(\phi)$ ,  $V_i(\phi)$ ,  $V_{ij}(\phi)$  be a scalar, a vector, and a symmetric tensor on the target manifold, respectively. ‘Pull-back’ composite operators of dimension 0,1,2 are defined by [162]

$$\begin{aligned} \llbracket V(\phi) \rrbracket &= \lambda V \cdot \frac{\partial}{\partial F} L_B = \mu^{d-2} Z(\mathfrak{h}) V, \\ \llbracket V_i(\phi) \partial^\mu \phi^i \rrbracket &= \lambda V_i \cdot \frac{\partial}{\partial V_{\mu i}} L_B = \mu^{d-2} \left[ \partial^\mu \phi^i Z^V(\mathfrak{h})^j_i V_j + V_i \cdot \frac{\partial}{\partial V_{\mu i}} Y \right], \\ \llbracket \frac{1}{2} V_{ij}(\phi) \partial^\mu \phi^i \partial_\mu \phi^j \rrbracket &= \lambda V \cdot \frac{\partial}{\partial \mathfrak{h}} L_B - \frac{\mu^{d-2}}{\sqrt{\bar{\gamma}}} \partial_\mu \left[ \sqrt{\bar{\gamma}} \partial^\mu \phi^i N_i^{jk}(\mathfrak{h}) V_{jk} + \sqrt{\bar{\gamma}} V_{ij} \cdot \frac{\partial}{\partial (\bar{\partial}_\mu h_{ij})} Y \right]. \end{aligned} \quad (\text{B.70})$$

The functional derivatives here act on functionals on the target manifold at fixed  $x$ , e.g.  $V \cdot \frac{\partial}{\partial F} = \int d^D \phi \sqrt{h} V(\phi; x) \frac{\partial}{\partial F(\phi; x)}$ . For  $h_{ij}$  in addition the dependence of the counter terms on  $\bar{\partial}_\mu h_{ij}$  has to be taken into account, so that  $V \cdot \frac{\partial}{\partial \mathfrak{h}} := V_{ij} \cdot \frac{\partial}{\partial h_{ij}} + \bar{\partial}_\mu V_{ij} \cdot \frac{\partial}{\partial (\bar{\partial}_\mu h_{ij})}$ . Further  $L_B = L(\mathcal{G}_B, \phi_B)$  is the bare Lagrangian regarded as a function of the renormalized quantities. The contractions on the base space are with respect to the background metric  $\bar{\gamma}_{\mu\nu}$ . The additional total divergence in the last relation of Equation (B.70) reflects the effect of operator mixing. The normal products as given in Equation (B.70) still refer to the functional measure as defined by the source-extended Lagrangian. After all differentiations have been performed the sources should be set to zero or rendered  $x$ -independent again to get the composite operators e.g. for the purely metric sigma-model.

The definition (B.70) of the normal products is consistent with redefinitions of the couplings/sources that change the Lagrangian only by a total divergence. The operative identities are

$$(Z^V)_i^j \partial_j V = \partial_i(ZV), \quad (\bar{\partial}^\mu Z)V = \partial_i V \cdot \frac{\partial Y}{\partial V_{\mu i}} \quad (\text{B.71})$$

for a scalar  $V(\phi; x)$ . They entail

$$\partial_\mu \llbracket V \rrbracket = \llbracket \partial_i V \partial_\mu \phi^i \rrbracket + \llbracket \bar{\partial}_\mu V \rrbracket. \quad (\text{B.72})$$

Moreover the invariance of the regularization under reparameterizations of the target manifold allows one to convert the reparameterization invariance of the basic Lagrangian (B.57) into a “diffeomorphism Ward identity” [201, 162]:

$$\frac{1}{\sqrt{\bar{\gamma}}} \partial^\mu \llbracket \sqrt{\bar{\gamma}} \lambda J_\mu(v) \rrbracket = \left[ \left[ \frac{1}{2} \mathcal{L}_v \mathfrak{h}_{ij} \partial^\mu \phi^i \partial_\mu \phi^j + \partial^\mu \phi^i \mathcal{L}_v V_{\mu i} + \frac{1}{2} R^{(2)}(\bar{\gamma}) \mathcal{L}_v \Phi + \mathcal{L}_v F \right] \right] - \lambda v^i \cdot \frac{\delta S_B}{\delta \phi^i}, \quad (\text{B.73})$$

with  $\lambda J_\mu(v) = \partial_\mu \phi^i v_i + v^i V_{\mu i}$ . The Lie derivative terms on the right-hand-side are the response of the couplings/sources under an infinitesimal diffeomorphism  $\phi^j \rightarrow \phi^j + v^j(\phi)$ . Thus  $J_\mu(v)$  may be viewed as a “diffeomorphism current”. The last term on the right-hand-side is the (by itself finite) “equations of motion operator”. In deriving Equation (B.73) the following useful consistency conditions arise

$$\begin{aligned} \mathfrak{h}_{ij}^B v^j &= \mu^{d-2} \left[ Z^V(\mathfrak{h})_i^j v_j + N_i^{jk}(\mathfrak{h}) \mathcal{L}_v \mathfrak{h}_{jk} \right], \\ v^i V_{\mu i}^B &= \mu^{d-2} \left[ \mathcal{L}_v \mathfrak{h}_{ij} \cdot \frac{\partial Y}{\partial (\bar{\partial}^\mu \mathfrak{h}_{ij})} + v_i \cdot \frac{\partial Y}{\partial V_i^\mu} + Z(v^i V_{\mu i}) \right]. \end{aligned} \quad (\text{B.74})$$

So far the renormalization was done at a fixed normalization scale  $\mu$ . The scale dependence of the renormalized couplings/sources  $\mathcal{G} = \{\mathfrak{h}_{ij}, V_{\mu i}, D, F\}$  is governed by a set of renormalization functions which follow from Equation (B.61). For a counter tensor of the form (B.62) it is convenient to introduce

$$\dot{\mathcal{O}} = - \sum_{l \geq 1} \left( \frac{\lambda}{2\pi} \right)^l l \mathcal{O}^{(1,l)}, \quad (\text{B.75})$$

which in view of Equation (B.62) can be regarded as a parametric derivative of  $\mathcal{O}^{(1)}$ . Then

$$\begin{aligned} \mu \frac{d}{d\mu} \mathfrak{h}_{ij} &= \beta_{ij} := (2-d) \mathfrak{h}_{ij} - \dot{T}_{ij}, \\ \mu \frac{d}{d\mu} V_{\mu i} &= \gamma^V := (2-d) V_{\mu i} - (\dot{Z}^V)_i^j V_{\mu j} - \dot{N}_i^{jk} \bar{\partial}_\mu \mathfrak{h}_{jk}, \\ \mu \frac{d}{d\mu} D &= \gamma^D := (2-d - \dot{Z}) D - \dot{U}, \\ \mu \frac{d}{d\mu} F &= \gamma^F := (2-d - \dot{Z}) F - \dot{Y}. \end{aligned} \quad (\text{B.76})$$

The associated renormalization group operator is

$$\mathcal{D} = \mu \frac{\partial}{\partial \mu} + \beta \cdot \frac{\partial}{\partial \mathfrak{h}} + \gamma_\mu^V \cdot \frac{\partial}{\partial V_\mu} + \gamma^D \cdot \frac{\partial}{\partial D} + \gamma^F \cdot \frac{\partial}{\partial F}. \quad (\text{B.77})$$

For example the dimension 0 composite operators in Equation (B.70) obey

$$\mathcal{D} \llbracket V(\phi) \rrbracket = \llbracket (d-2 + \dot{Z} + \mathcal{D}) V \rrbracket, \quad (\text{B.78})$$

and similar equations hold for the dimension 1, 2 composite operators.

An important application of this framework is the determination of the Weyl anomaly as an ultraviolet finite composite operator. We shall only need the version without vector and scalar functionals. The result then reads [201, 220, 162]

$$\bar{\gamma}^{\mu\nu} \llbracket T_{\mu\nu} \rrbracket = \frac{1}{2} \llbracket B_{ij}(\mathfrak{h}/\lambda) \bar{\gamma}^{\mu\nu} \partial_\mu \phi^i \partial_\nu \phi^j \rrbracket + \frac{1}{2} R^{(2)}(\bar{\gamma}) \llbracket B^D \rrbracket. \quad (\text{B.79})$$

Here the so-called Weyl anomaly coefficients enter:

$$\begin{aligned}\lambda B_{ij}(\mathfrak{h}/\lambda) &:= \lambda \beta_{ij}(\mathfrak{h}/\lambda)|_{d=2} + \mathcal{L}_S \mathfrak{h}_{ij}, \\ \lambda B^D(D, \mathfrak{h}/\lambda) &:= \lambda \gamma^D(\mathfrak{h}/\lambda)|_{d=2} + S^j \partial_j D,\end{aligned}\tag{B.80}$$

where  $\beta_{ij}$  and  $\gamma^D$  are the renormalization group functions of Equation (B.76) and

$$S_i := W_i + \partial_i D \quad \text{with } W_i := N^{(1)}(\mathfrak{h})_i{}^{jk} \mathfrak{h}_{jk} = \left(\frac{\lambda}{2\pi}\right)^3 \frac{1}{32} \partial_i (R_{klmn} R^{klmn}) + O(\lambda^4).\tag{B.81}$$

These expressions hold in dimensional regularization, minimal subtraction, and the background field expansion in terms of normal coordinates. Terms proportional to the equations of motion operator  $\frac{\delta S_B}{\delta \phi^j}$  have been omitted. The normal-products (B.70) are normalized such that the expectation value of an operator contains as its leading term the value of the corresponding functional on the background,  $\langle \mathcal{O}(\phi) \rangle = \mathcal{O}(\bar{\phi}) + \dots$ , where the subleading terms are in general nonlocal and depend on the scale  $\mu$ . For the expectation value of the trace anomaly this produces a rather cumbersome expression (see e.g. [220]). As stressed in [201] the result (B.70), in contrast, allows one to use  $B_{ij}(\mathfrak{h}) = 0$  as a simple criterion to select functionals which ‘minimize’ the conformal anomaly.

The Weyl anomaly coefficients (and the anomaly itself) can be shown to be invariant under field redefinitions of the form

$$\phi_B^j \longrightarrow \phi_B^j + \frac{1}{2-d} V^j(\phi, \lambda),\tag{B.82}$$

with  $V^j(\phi, \lambda) = \sum_{l \geq 1} (\frac{\lambda}{2\pi})^l V_l^j(\phi)$  functionally *independent* of the metric. Roughly speaking Equation (B.82) changes the beta function by a Lie derivative term that is compensated by a contribution of the diffeomorphism current to the anomaly which amounts to  $W_j \rightarrow W_j - V_j$  [201]. It is important to distinguish these diffeomorphisms from field renormalizations like Equation (3.65, 3.70) that depend on the metric. Although formally Equation (B.82) amounts to  $\Xi^j(\phi, \lambda) \rightarrow \Xi^j(\phi, \lambda) + V^j(\phi, \lambda)$  in Equation (3.65); clearly one cannot cancel one against the other. The distinction is also highlighted by considering the change in the metric counter terms

$$T_{ij}^{(1)}(\mathfrak{h}) \longrightarrow T_{ij}^{(1)}(\mathfrak{h}) - \mathcal{L}_V \mathfrak{h}_{ij},\tag{B.83}$$

under Equation (B.82). Without further specifications this would not be legitimate for a  $\mathfrak{h}$ -dependent vector. Although the Lie derivative term in Equation (B.83) drops out when recomputing  $\beta_{ij}$  directly as a parametric derivative, in combinations like

$$\beta_{ij}(\phi_B) \partial^\mu \phi_B^i \partial_\mu \phi_B^j = \beta_{ij}(\phi) \partial^\mu \phi^i \partial_\mu \phi^j + \frac{1}{2-d} \mathcal{L}_V \beta_{ij}(\phi) \partial^\mu \phi^i \partial_\mu \phi^j + \dots\tag{B.84}$$

the term  $(2-d)\mathfrak{h}_{ij}$  in the metric beta function of Equation (B.76) induces an effective shift

$$\beta_{ij}(\mathfrak{h}) \longrightarrow \beta_{ij}(\mathfrak{h}) + \mathcal{L}_V \mathfrak{h}_{ij}.\tag{B.85}$$

Similarly  $W_i$  is shifted to  $W_i - V_i$  and the Weyl anomaly coefficients are invariant.

In the context of Riemannian sigma-models  $D$  is usually interpreted as a ‘‘string dilaton’’ for the systems (B.57) defined on a curved base space. If one is interested in the renormalization of Equation (B.57) on a flat base space,  $D$  on the other hand plays the role of a potential for the improvement term of the energy momentum tensor. This role of  $D$  can be made manifest by rewriting Equation (B.79) by means of the diffeomorphism Ward identity. Returning to a flat base space one finds [201, 162]

$$\llbracket T^\mu{}_\mu \rrbracket = \frac{1}{2} \llbracket \beta_{ij}(\mathfrak{h}/\lambda) \partial^\mu \phi^i \partial_\mu \phi^j \rrbracket + \partial^\mu \partial_\mu \llbracket D \rrbracket + \partial^\mu \llbracket \partial_\mu \phi^i W_i \rrbracket,\tag{B.86}$$

where again terms proportional to the equations of motion operator have been omitted. Here  $\partial^2[[D]]$  is the ‘naive’ improvement term while the additional total divergence is induced by operator mixing.

The functions  $B^D$  and  $B_{ij}(g)$  are linked by an important consistency condition, the Curci–Paffuti relation [57]. We present it in two alternative versions,

$$\begin{aligned}\partial_i \dot{U} &= \dot{N}_i{}^{jk} \dot{T}_{jk} - \dot{T} \cdot \frac{\partial}{\partial g} W_i + (\dot{Z}^V)_i{}^j W_j, \\ \partial_i B^D &= \dot{N}_i{}^{jk} B_{jk} - B \cdot \frac{\partial}{\partial g} S_i + B_{ij} S^j.\end{aligned}\tag{B.87}$$

The first version displays the fact that the identity relates various  $\mathfrak{h}$ -dependent counter terms without  $D$  entering. In the second version  $D$  is introduced in a way that yields an identity among the Weyl anomaly coefficients. It has the well-known consequence that  $B^D$  is constant when  $B_{ij}$  vanishes:

$$B_{ij} = 0 \quad \implies \quad B^D = c_T/6,\tag{B.88}$$

where  $c_T$  is the central charge of energy momentum tensor derived from Equation (B.57).

## C The Effective Average Action

Some general properties of the effective average action  $\Gamma_{\Lambda,k}$  have been anticipated in Section 2.2 in relation to the UV renormalization problem. In this appendix, we discuss the effective average action and its properties in more detail. First we give the definition of  $\Gamma_{\Lambda,k}$  from the functional integral with suitable mode cutoff kernels  $\mathcal{R}_{\Lambda,k}$ . Then we derive the basic flow equation (2.11) and discuss its properties. Finally the decoupling properties of  $\Gamma_{\Lambda,k}$  are exemplified, which are useful in particular for “renormalization group improvement”.

Again we illustrate the concepts below for a scalar quantum field theory on flat space. Part of the rationale for using the effective average action, however, stems from the fact that via the background field method it can be generalized to gauge theories as well. See [13] for an alternative computationally tested approach.

For orientation we briefly describe the route that led to the effective average action and its generalization to gauge theories. Initially the average action proper [193, 186, 185] was introduced as the generating functional for the correlators of fields averaged over a Euclidean spacetime volume  $k^{-d}$ , in  $d$  dimensions. Here “averaging” is to be understood in the literal sense; the defining functional integral (over a scalar  $\chi$ , say) contains a smeared delta functional  $\delta[\phi - \langle\chi\rangle_k]$  which forces  $\langle\chi\rangle_k$ , the average of  $\chi$  over a ball of radius  $1/k$ , to be equal to an externally prescribed field  $\phi$ . This construction is a continuum counterpart of a Kadanoff block spin transformation. For  $k \rightarrow 0$  the average action proper approaches the constraint effective potential studied earlier [161].

While the average action proper for non-gauge theories has a clear physical interpretation it has proven difficult to generalize it to gauge theories. With certain modifications this is possible in the Abelian case [186, 185] but the construction fails for non-Abelian Yang–Mills theories. This was the motivation for introducing the effective average action [228, 186, 185, 187]. It reinterprets the averaging in the non-gauge case as a cutting off of Fourier modes, the eigenfunctions of the ordinary Laplacian, and replaces it in the Yang–Mills case by a corresponding cutting-off of the eigenmodes of the covariant Laplacian  $D^\mu D_\mu$ . Contrary to the old average action the new one is defined in terms of a Legendre transform and therefore encodes the information about the multi-point functions in the more condensed 1PI form (see also [152]). The price to pay is that the simple averaging is replaced with the less intuitive weighing of field modes according to their  $D^\mu D_\mu$  eigenvalue, with the corresponding change in the meaning of “long” and “short” wavelength modes. However the effective average action has better effective field theory properties and satisfies a closed functional evolution equation. Both aspects have been tested in Yang–Mills theories, we refer to [229, 29] for a review and further literature.

### C.1 Definition and basic properties

The construction of  $\Gamma_{\Lambda,k}[\phi]$  starts out from a modified form,  $W_{\Lambda,k}[J]$ , of the standard generating functional  $W_\Lambda[J]$ :

$$\exp\{W_{\Lambda,k}[J]\} = \int \mathcal{D}\chi \exp \left\{ -S_\Lambda[\chi] - C_{\Lambda,k}[\chi] + \int dx \chi(x)J(x) \right\}. \quad (\text{C.1})$$

The extra factor  $\exp\{-C_{\Lambda,k}[\chi]\}$  suppresses the “IR modes” with  $p^2 < k^2$ . The modified  $W_{\Lambda,k}[J]$  is easily seen to be still a convex functional of the source. The corresponding  $\Lambda$ ,  $k$ , and  $J$ -dependent expectations of some (smooth) observable  $\mathcal{O}(\chi)$  are defined as in Equation (B.1)

$$\langle \mathcal{O} \rangle_{\Lambda,k} = \int \mathcal{D}\chi \mathcal{O}(\chi) \exp \left\{ -S_\Lambda[\chi] - C_{\Lambda,k}[\chi] + \int dx \chi(x)J(x) \right\}. \quad (\text{C.2})$$

The cutoff functional  $C_{\Lambda,k}$  is a quadratic form in the fields and has already been displayed in Equation (2.10). In the literature it is often denoted by  $\Delta_k S$  to indicate that it should be thought of as modifying the bare action.

The kernel  $\mathcal{R}_{\Lambda,k}$  defining  $C_{\Lambda,k}$  is conveniently chosen such that both  $\mathcal{R}_{\Lambda,k}$  and  $k\partial_k\mathcal{R}_{\Lambda,k}$  define a trace-class operator [157] (see Equation (2.10)). Once the trace-class condition is satisfied one can adjust the other features of the kernel to account for the mode suppression. These features are arbitrary to some extent; what matters is the limiting behavior for  $p^2, q^2 \gg k^2$  and (with foresight)  $\Lambda \rightarrow \infty$ . In the simplest case we require that  $\mathcal{R}_{\Lambda,k}(p, q)$  is smooth in all variables and of the factorized form

$$\mathcal{R}_{\Lambda,k}(p, q) = \mathcal{R}_k(p^2)\delta_\Lambda(p + q), \quad \lim_{\Lambda \rightarrow \infty} \delta_\Lambda(p) = \delta(p), \quad (\text{C.3})$$

$$\mathcal{R}_k(p^2) \approx \begin{cases} k^2 & \text{for } p^2 \ll k^2, \\ 0 & \text{for } k^2 \ll p^2. \end{cases} \quad (\text{C.4})$$

In the first condition  $\delta_\Lambda(p)$  is a smooth approximation to the delta distribution, normalized such that  $\int \frac{dp}{(2\pi)^d} \delta_\Lambda(p) = 1$ . In Fourier space the finiteness of the trace then amounts to  $\text{Tr}[\mathcal{R}_{\Lambda,k}] = \delta_\Lambda(0) \int \frac{dp}{(2\pi)^d} \mathcal{R}_k(p^2) < \infty$ . The condition (C.4) guarantees that the large momentum modes are integrated out in the usual way, while the  $\mathcal{R}_k(p^2) \approx k^2$  behavior for small  $p^2$  leads to a suppression of the small momentum modes by a soft mass-like IR cut-off. Indeed, if the bare action has the structure  $S_\Lambda[\chi] = \int dx \{ \frac{1}{2}(\partial_\mu\chi)^2 + \frac{1}{2}m_\Lambda^2\chi^2 + \text{interactions} \}$ , the addition of a  $C_{\Lambda,k}[\chi]$  term as in Equation (2.10) leads to

$$S_\Lambda[\chi] + C_{\Lambda,k}[\chi] = \frac{1}{2} \int \frac{dp}{(2\pi)^d} [p^2 + m^2 + \mathcal{R}_k(p^2)] |\hat{\chi}(p)|^2 + \dots, \quad (\text{C.5})$$

where  $m$  is the renormalized mass and the dots indicate the interaction terms and terms which vanish for  $\Lambda \rightarrow \infty$ . Obviously the cutoff function  $\mathcal{R}_k(p^2)$  has the interpretation of a momentum dependent mass square which vanishes for  $p^2 \gg k^2$  and assumes the constant value  $k^2$  for  $p^2 \ll k^2$ . How  $\mathcal{R}_k(p^2)$  is assumed to interpolate between these two regimes is a matter of calculational convenience. In practical calculations one often uses the exponential cutoff  $\mathcal{R}_k(p^2) = p^2[\exp(p^2/k^2) - 1]^{-1}$ , but many other choices are possible [29, 146].

Next one introduces the Legendre transform of  $W_{\Lambda,k}$ ,

$$\tilde{\Gamma}_{\Lambda,k}[\phi] := \sup_{\{J\}} \left\{ \int dx J(x)\phi(x) - W_{\Lambda,k}[J] \right\}, \quad (\text{C.6})$$

which is a convex functional of  $\phi$ . Making the usual simplifying assumption that  $W_{\Lambda,k}[J]$  admits a series expansion in powers of  $J$ , a formal inversion of the series  $\delta W[J]/\delta J = \phi$  defines a unique configuration  $J = J_*[\phi]$  with the property  $J_*[\phi = 0] = 0$  and  $\phi(x) = \langle \chi(x) \rangle = \delta W_{\Lambda,k}[J_*]/\delta J_*(x)$ . The actual effective average action is defined by

$$\Gamma_{\Lambda,k}[\phi] := \tilde{\Gamma}_{\Lambda,k}[\phi] - C_{\Lambda,k}[\phi]. \quad (\text{C.7})$$

The subtraction of the mode suppression term is essential for the properties listed below. The main properties of the effective average action are:

1. If the bare action  $S_\Lambda$  is quadratic (free field theory) the action  $\Gamma_{\Lambda,k}$  (but not  $\tilde{\Gamma}_{\Lambda,k}$ ) is independent of  $k$  and equals the bare one:  $\Gamma_{\Lambda,k} = S_\Lambda$ , for  $0 \leq k < \Lambda$ .
2. It satisfies the functional integro-differential equation for the standard effective action with  $S_\Lambda + C_{\Lambda,k}$  playing the role of the bare action, i.e.

$$\exp\{-\Gamma_{\Lambda,k}[\phi]\} = \int \mathcal{D}\chi \exp \left\{ -S_\Lambda[\chi] - C_{\Lambda,k}[\chi - \phi] + \int dx (\chi - \phi)(x) \frac{\delta \Gamma_{\Lambda,k}[\phi]}{\delta \phi(x)} \right\}. \quad (\text{C.8})$$

Equation (C.8) readily follows by converting Equation (C.1) via the definitions using the relation  $J_*[\phi] = \delta \tilde{\Gamma}_{\Lambda,k}[\phi]/\delta \phi = -\mathcal{R}_{\Lambda,k}\phi + \delta \Gamma_{\Lambda,k}[\phi]/\delta \phi$ , which is ‘dual’ to  $\delta W_{\Lambda,k}[J_*]/\delta J_* = \phi$ .



3. It interpolates between the  $S_\Lambda[\phi]$  and the UV regularized standard effective action  $\Gamma_\Lambda[\phi]$ , according to

$$S_\Lambda[\phi] \xleftarrow{k \rightarrow \Lambda} \Gamma_{\Lambda,k}[\phi] \xrightarrow{k \rightarrow 0} \Gamma_\Lambda[\phi]. \tag{C.9}$$

The first relation,  $\lim_{k \rightarrow 0} \Gamma_{\Lambda,k} = \Gamma_\Lambda$ , follows trivially from Equation (B.8) and the fact that  $\mathcal{R}_k(p^2)$  vanishes for all  $p^2 > 0$  when  $k \rightarrow 0$ . The  $k \rightarrow \Lambda$  limit of Equation (2.8) is more subtle. A formal argument for  $\lim_{k \rightarrow \Lambda} \Gamma_k \approx S_\Lambda$  is as follows. Since  $\mathcal{R}_k(p^2)$  approaches  $k^2$  for  $k \approx \Lambda$ , and  $\Lambda$  large, the second exponential on the right-hand-side of Equation (2.8) becomes

$$\exp \left\{ -k^2 \int \frac{dp}{(2\pi)^d} \frac{dq}{(2\pi)^d} \delta_\Lambda(p+q) \widehat{(\chi - \phi)}(p) \widehat{(\chi - \phi)}(q) \right\}. \tag{C.10}$$

For  $k \approx \Lambda$  this approaches a delta-functional  $\delta[\chi - \phi]$ , up to an irrelevant normalization. The  $\chi$  integration can be performed trivially then and one ends up with  $\lim_{k \rightarrow \Lambda} \Gamma_k[\phi] \approx S_\Lambda[\phi]$ , for  $\Lambda$  large. In a more careful treatment [29] one shows that the saddle point approximation of the functional integral in Equation (2.8) about the point  $\chi = \phi$  becomes exact in the limit  $k \approx \Lambda \rightarrow \infty$ . As a result  $\lim_{k \approx \Lambda \rightarrow \infty} \Gamma_{\Lambda,k}$  and  $S_\Lambda$  differ at most by the infinite mass limit of a one-loop determinant, which is ignored in Equation (2.7) since it plays no role in typical applications (see [188] for a more careful discussion).

## C.2 Flow equation

As anticipated, the scale dependence of  $\Gamma_{\Lambda,k}$  is governed by a functional differential equation [228, 229, 29],

$$k \frac{\partial}{\partial k} \Gamma_{\Lambda,k}[\phi] = \frac{1}{2} \text{Tr} \left[ k \frac{\partial}{\partial k} \mathcal{R}_{\Lambda,k} \left( \Gamma_{\Lambda,k}^{(2)}[\phi] + \mathcal{R}_{\Lambda,k} \right)^{-1} \right]. \tag{C.11}$$

This is the *functional renormalization group equation* (FRGE) for the effective average action. The ingredients have already been explained in Section 2.3; here we present its derivation. An instructive derivation of Equation (C.11) starting directly from Equation (C.8) can be found in [44] (see also [152]). A technically quicker one proceeds as follows [228]:

Taking the  $k$ -derivative of Equation (C.6) with Equations (C.1, 2.10) inserted one finds

$$k \frac{\partial}{\partial k} \tilde{\Gamma}_{\Lambda,k}[\phi] = \frac{1}{2} \int dx dy k \frac{\partial}{\partial k} \mathcal{R}_{\Lambda,k}(x, y) \langle \chi(x) \chi(y) \rangle_{\Lambda,k}, \tag{C.12}$$

with the  $J$  and  $k$ -dependent expectation values defined as in Equation (C.2). Next it is convenient to introduce the connected 2-point function  $G_{\Lambda,k}(x, y) := \delta^2 W_{\Lambda,k}[J_*] / \delta J_*(x) \delta J_*(y)$  and  $\tilde{\Gamma}_k^{(2)}(x, y) := \delta^2 \tilde{\Gamma}_{\Lambda,k}[\phi] / \delta \phi(x) \delta \phi(y)$ . Since  $W_{\Lambda,k}$  and  $\tilde{\Gamma}_{\Lambda,k}$  are related by a Legendre transformation,  $G_{\Lambda,k}(x, y)$  and  $\tilde{\Gamma}_{\Lambda,k}^{(2)}(x, y)$  are mutually inverse integral kernels, i.e.  $G_{\Lambda,k} \tilde{\Gamma}_{\Lambda,k}^{(2)} = 1$  for the associated integral operators. Taking two  $J$ -derivatives of Equation (C.1) one obtains  $\langle \chi(x) \chi(y) \rangle_{\Lambda,k} = G_{\Lambda,k}(x, y) + \phi(x) \phi(y)$ . Substituting this expression for the two-point function into Equation (C.12) we arrive at

$$k \frac{\partial}{\partial k} \tilde{\Gamma}_{\Lambda,k}[\phi] = \frac{1}{2} \text{Tr} \left[ k \frac{\partial}{\partial k} \mathcal{R}_k G_{\Lambda,k} \right] + \frac{1}{2} \int dx dy \phi(x) k \frac{\partial}{\partial k} \mathcal{R}_{\Lambda,k}(x, y) \phi(y). \tag{C.13}$$

In terms of  $\Gamma_{\Lambda,k}$  this translates into  $k \frac{\partial}{\partial k} \Gamma_{\Lambda,k}[\phi] = \frac{1}{2} \text{Tr} [k \frac{\partial}{\partial k} \mathcal{R}_{\Lambda,k} G_{\Lambda,k}]$ . The derivation is completed by noting that  $[\tilde{\Gamma}_{\Lambda,k}^{(2)}]^{-1} = (\Gamma_{\Lambda,k}^{(2)} + \mathcal{R}_{\Lambda,k})^{-1}$ , where the second equality follows by differentiating Equation (C.7).

We add some comments on the FRGE (C.11):

1. The right-hand-side of Equation (C.11) can be also rewritten in the form of a one-loop expression

$$k \frac{\partial}{\partial k} \Gamma_{\Lambda,k}[\phi] = \frac{1}{2} \frac{D}{D \ln k} \text{Tr} \ln \left( \Gamma_{\Lambda,k}^{(2)}[\phi] + \mathcal{R}_{\Lambda,k} \right). \quad (\text{C.14})$$

Here the scale derivative  $D/D \ln k$  acts only on the  $k$ -dependence of  $\mathcal{R}_{\Lambda,k}$ , not on  $\Gamma_{\Lambda,k}^{(2)}$ . The  $\text{Tr} \ln(\dots) = \ln \det(\dots)$  expression in Equation (C.11) differs from a standard one-loop determinant in two ways: It contains the Hessian of the actual effective action rather than that of the bare action  $S_\Lambda$ , and it has a built in IR regulator  $\mathcal{R}_{\Lambda,k}$ . These modifications make Equation (C.11) an *exact* equation.

2. Observe that the FRGE (C.6) is independent of the bare action  $S_\Lambda$ , which enters only via the initial condition  $\Gamma_{\Lambda,\Lambda} = S_\Lambda$  (for large  $\Lambda$ ). In the FRGE approach the calculation of the path integral for  $W_{\Lambda,k}[J]$  is replaced with the task of integrating this RG equation from  $k = \Lambda$ , where the initial condition  $\Gamma_{\Lambda,\Lambda} = S_\Lambda$  is imposed, down to  $k = 0$ , where the effective average action equals the ordinary effective action  $\Gamma_\Lambda$ . The role of the bare action in the removal of cutoff  $\Lambda$  and its relation to the UV renormalization problem has already been described in Section 2.3. Here we repeat again that the *explicit*  $\Lambda$  dependence entering via  $\mathcal{R}_{\Lambda,k}$  is harmless, and  $\mathcal{R}_{\Lambda,k}$  can essentially be replaced with  $\mathcal{R}_{\infty,k}$ .

To see this let us momentarily write  $G_{\Lambda,k}(p, q)$  for the kernel of  $(\Gamma_{\Lambda,k}^{(2)}[\phi] + \mathcal{R}_{\Lambda,k})^{-1}$  in Fourier space. By assumption  $G_{\Lambda,k}(p, q)$  is a family of functions which remains pointwise bounded as  $\Lambda \rightarrow \infty$  (but the falloff in  $p, q$  may not be strong enough so as to define a bounded operator in the limit). The right-hand-side of Equation (C.11) then is proportional to  $\int dp dq \delta_\Lambda(p + q) k \frac{\partial}{\partial k} \mathcal{R}_k(p^2) G_{\Lambda,k}(p, q)$ , which by Equation (C.3) behaves like  $\int dp k \frac{\partial}{\partial k} \mathcal{R}_k(p^2) G_{\infty,k}(p, -p)$  for  $\Lambda \rightarrow \infty$ . On the other hand by Equation (C.4) the derivative  $k \frac{\partial}{\partial k} \mathcal{R}_k(p^2)$  has support mostly on a thin shell around  $p^2 \approx k^2$ , so that the (potentially problematic) large  $p$  behavior of  $G_{\infty,k}(p, -p)$  is irrelevant.

3. By repeated differentiation of Equation (C.11) and evaluation at an extremizing configuration  $\phi_*$  one obtains a coupled infinite system of flow equations for the  $n$ -point functions (B.7). For example for a translation invariant theory [34],

$$k \partial_k \Gamma_{\Lambda,k}^{(2)}(p, -p) = \int \frac{dq}{(2\pi)^d} k \partial_k \mathcal{R}_{\Lambda,k}(q, -q) \left\{ G_{\Lambda,k}(q^2) \Gamma_{\Lambda,k}^{(3)}(p, q, -p - q) \right. \\ \times G_{\Lambda,k}((p + q)^2) \Gamma_{\Lambda,k}^{(3)}(-p, p + q, -q) G_{\Lambda,k}(q^2) \\ \left. - \frac{1}{2} G_{\Lambda,k}(q^2) \Gamma_{\Lambda,k}^{(4)}(p, -p, q, -q) G_{\Lambda,k}(q^2) \right\}. \quad (\text{C.15})$$

Here we defined  $\Gamma_{\Lambda,k}^{(n)}(x_1, \dots, x_n)$  in analogy to Equation (B.7) and wrote  $\Gamma_{\Lambda,k}^{(n)}(p_1, \dots, p_n)$   $(2\pi)^d \delta(p_1 + \dots + p_n)$  for its Fourier transform. Similarly for  $G_{\Lambda,k}^{-1}(q^2) := \Gamma_{\Lambda,k}^{(2)}(q, -q) + \mathcal{R}_{\Lambda,k}(q, -q)$ . One sees that lower multipoint functions couple to higher order ones in a way so that only the infinite system closes. However if all external momenta are small,  $p_i^2 < k^2$ , the  $k \partial_k \mathcal{R}_{\Lambda,k}$  insertion will ensure that also the internal momentum is small. This is the rationale for the derivative (small momentum) expansion. For an approximation suited for uniformly large momenta  $p_i^2 \geq k^2$  see [34].

4. The above FRGE is in the spirit of Wilson–Kadanoff renormalization ideas, but with the iterated coarse graining procedure replaced by a direct mode cutoff. Since the kernel in Equation (A.2) cuts off momenta with  $p^2 < k^2$ , the right-hand-side of Equation (A.2) corresponds to Equation (C.2) evaluated at  $J = 0$ . One could also have derived a flow equation for (a suitable variant of)  $W_{\Lambda,k}[0]$ .

However, there are conceptual differences between the effective average action and a genuine Wilsonian action  $S_l = S_l^W$ , as discussed in Appendix A.

First, in the literature the running scale  $l$  on which a Wilsonian action depends is frequently referred to as an *ultraviolet* cutoff and is denoted by  $\Lambda_W$ . This is due to a difference in perspective: If all modes of the original system with momenta between infinity (or  $\Lambda$ ) are integrated out,  $l = \Lambda_W$  is an infrared cutoff for the original model, but it plays the role of an UV cutoff for the “residual theory” of the modes below this scale, which are to be integrated out still. For them  $S_{\Lambda_W}^W$  has the status of a bare action.

Second, the Wilsonian action  $S_{\Lambda_W}^W$  describes a set of different actions, parameterized by  $\Lambda_W$  and subject to a flow equation like Equation (A.4), for one and the same system; the Greens functions are independent of  $\Lambda_W$  and have to be computed from  $S_{\Lambda_W}^W$  by further functional integration. In contrast the effective average action  $\Gamma_{\Lambda,k}$  can be thought of as the standard effective action for a family of different systems; for any value of  $k$  it equals the standard effective action (generating functional for the vertex- or 1-PI Green’s functions) for a model with bare action  $S_\Lambda + C_{\Lambda,k}$ . The latter is of course not subject to a Wilsonian type flow equation like Equation (A.4). In particular the multi-point functions do depend on  $k$ . This is a desired property, however, as these  $k$ -dependent Green’s functions are supposed to provide an effective field theory description of the physics at scale  $k$ , without further functional integration. See [29] for a detailed discussion.

There exists a variety of different functional renormalization group equations. We refer to [21, 166, 29, 146, 229] for reviews. To a certain extent they contain the same information but encoded in different ways (see e.g. [147]); the differences become important in approximations (the ‘truncations’ described below) where simple truncations adapted to a certain application in one FRGE might correspond to more complicated and less adapted truncations in another. We use the effective average action [228, 229, 29] here because of its effective field theory properties and because via the background field method it has been extended to gravity (see [179]). FRGEs invariant under field reparameterizations have been developed in [165, 44] but have not yet been applied in computations.

5. To solve the functional flow equation (C.11) approximations are indispensable. One which does not rely on a perturbative expansion is by truncation of the space of candidate continuum functionals  $\Gamma_k^{\text{trunc}}[\phi]$  to one where the initial value problem for the flow equation (2.12) can be solved in reasonably closed form. In this case one can then by ‘direct inspection’ determine the initial data for which a global solution exists. The existence of a nontrivial unstable manifold for  $\Gamma_k^{\text{trunc}}$  can then be taken as witnessing the renormalizability of an implicitly defined ‘hierarchical’ dynamics (see Section 2.3).

Concretely the truncation is usually done by assuming an ansatz of the form

$$\Gamma_k^{\text{trunc}}[\phi] = \sum_{\alpha=1}^N u_\alpha(k) P_\alpha[\phi], \quad (\text{C.16})$$

where the  $u_\alpha$ ’s are scale dependent (‘running’) parameters as in the previous general discussion, and the  $k$ -independent functionals  $P_\alpha[\phi]$  span the subspace selected. The ‘art’ of course consists in choosing a set of  $P_\alpha[\phi]$  small enough to be computationally manageable and yet such that the projected flow encapsulates the essential physics features of the exact flow. The projected RG flow then is described by a set of ordinary differential equations for the parameters  $u_\alpha(k)$ . Schematically those equations arise as follows. Let us assume the finite set  $P_\alpha[\phi]$ ,  $\alpha = 1, \dots, N$ , can be extended to a ‘basis’ (in the sense discussed in Section 2.1)  $P_\alpha$ ,  $\alpha \in \mathbb{N}$ , of the full space of functionals. Expanding the  $\phi$  dependence of  $\text{Tr}[\dots]$  on the

right-hand-side of the FRGE (C.11) in this basis an expression of the form

$$\frac{1}{2} \text{Tr}[\dots] = \sum_{\alpha=1}^{\infty} b_{\alpha}(u; k) P_{\alpha}[\phi] \quad (\text{C.17})$$

arises. Here the  $b_{\alpha}(u; k)$  play the role of generalized beta functions for the parameters  $u_{\alpha}$ ,  $\alpha \in \mathbb{N}$ . Neglecting, (i) the terms with  $\alpha > N$  in Equation (C.17) and (ii) the dependence of  $b_1, \dots, b_N$  on the parameters  $u_{\alpha}$ ,  $\alpha > N$ , a closed system of ordinary differential equations arises:  $k \frac{\partial}{\partial k} u_{\alpha} = b_{\alpha}(u_1, \dots, u_N; k)$ ,  $\alpha = 1, \dots, N$ . As in Equation (A.3) one can reparameterize the  $\Gamma_k[\phi]$  in Equation (C.14) in terms of inessential parameters and couplings  $g_i$ ,  $i = 1, \dots, n$ , with the latter ones made dimensionless. The corresponding  $b_i$  functions then become the beta functions proper and the system of differential equations reads

$$k \partial_k g_i(k) = \beta_i(g_1, \dots, g_n), \quad i = 1, \dots, n. \quad (\text{C.18})$$

The  $\beta_i$ 's have no explicit  $k$  dependence and define a ‘time independent’ vector field on the space of couplings  $\{g_1, \dots, g_n\}$ .

Another approximation procedure for the solution of Equation (2.12) is the *local potential approximation* [21, 111, 147]. Here the functionals  $\Gamma_k$ 's are constrained to contain only a standard kinetic term plus arbitrary non-derivative terms

$$\Gamma_k^{\text{trunc}}[\phi] := \int dx \left\{ \frac{1}{2} (\partial\phi(x))^2 + U_k(\phi(x)) \right\}. \quad (\text{C.19})$$

Since the potential function  $\phi \mapsto U(\phi)$  could be (Taylor-) expanded one can view (C.19) as a simple infinite parametric version of Equation (C.16). The truncated flow equations for Equation (C.19) now amount to a partial differential equation (in two variables) for the running potential  $U_k(\phi)$ . It is obtained by inserting Equation (C.19) into the FRGE and projecting the trace onto functionals of the form (C.19). This is most easily done by inserting a constant field  $\phi = \varphi = \text{const}$  into both sides of the equation since it gives a nonvanishing value precisely to the non-derivative  $P_{\alpha}$ 's. As  $[\Gamma_k^{\text{trunc}}]^{(2)}(\varphi) = -\partial^2 + U_k''(\varphi)$ ,  $U''(\varphi) := d^2 U_k / d\varphi^2$  has no explicit  $x$  dependence the trace is easily evaluated in momentum space. This leads to the following partial differential equation:

$$k \partial_k U_k(\varphi) = \frac{1}{2} \int \frac{dp}{(2\pi)^d} \frac{k \partial_k \mathcal{R}_k(p^2)}{p^2 + \mathcal{R}_k(p^2) + U_k''(\varphi)}. \quad (\text{C.20})$$

This equation describes how the classical (or microscopic) potential  $U_{\infty} = V_{\text{class}}$  evolves into the standard effective potential  $U_0 = V_{\text{eff}}$ . Remarkably, for an appropriate choice of  $\mathcal{R}_k$  the limit  $\lim_{k \rightarrow 0} U_k$  is automatically a convex function of  $\phi$ , a feature the full effective action must have but which is usually destroyed in perturbation theory. For a detailed discussion of this point we refer to [29]. Generally convexity can be used as guideline to identify good truncations.

A slight extension of the local potential approximation is to allow for a ( $\phi$ -independent) wave function renormalization, i.e. a running prefactor of the kinetic term:  $\Gamma_k^{\text{trunc}}[\phi] = \int dx \left\{ \frac{1}{2} Z_k (\partial\phi)^2 + U_k(\phi) \right\}$ . Using truncations of this type one should employ a slightly different normalization of  $\mathcal{R}_k(p^2)$ , namely  $\mathcal{R}_k(p^2) \approx Z_k k^2$  for  $p^2 \ll k^2$ . Then  $\mathcal{R}_k$  combines with  $\Gamma_k^{(2)}$  to the inverse propagator  $[\Gamma_k^{\text{trunc}}]^{(2)}[\phi] + \mathcal{R}_k = Z_k (p^2 + k^2) + \dots$ , as required if the IR cutoff is to give rise to a (mass)<sup>2</sup> of size  $k^2$  rather than  $k^2/Z_k$ . In particular in theories with more than one field it is important that all fields are cut off at the same  $k^2$ . This is achieved by a cutoff function of the form

$$\mathcal{R}_k(p^2) = k^2 Z_k \mathcal{R}^{(0)}(p^2/k^2), \quad (\text{C.21})$$

with  $\mathcal{R}^{(0)} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  as in Equation (4.15). Here  $\mathcal{Z}_k$  is in general a matrix in field space. In the sector of modes with inverse propagator  $Z_k^{(i)} p^2 + \dots$  the matrix  $\mathcal{Z}_k$  is chosen diagonal with entries  $\mathcal{Z}_k = Z_k^{(i)}$ . In a scalar field these  $\mathcal{Z}_k$  factors are automatically positive and the flow equations in the various truncations are well-defined.

### C.3 Decoupling properties

As compared to other generating functionals the effective average action has particularly benign *decoupling properties*. These are crucial for practical applications and provide a powerful tool for extracting physics information from  $\Gamma_{\Lambda,k}$ , in particular in the context of “renormalization group improvement”. In the following we suppress the UV cutoff  $\Lambda$  as it plays no role in the discussion.

For an illustration [189, 191, 190] consider a truncated solution of the FRGE (C.11),

$$\Gamma_k[\phi] = \int d^4x \left\{ \frac{1}{2} Z(k) \partial_\mu \phi \partial^\mu \phi + \frac{1}{2} m^2(k) \phi^2 + \frac{1}{12} \lambda(k) \phi^4 + \dots \right\}. \quad (\text{C.22})$$

To begin with we neglect the running of the kinetic term (“local potential ansatz”) and set  $Z(k) \equiv 1$ . For functionals of this type, and in a momentum basis where  $-\partial^2$  corresponds to  $p^2$ , the denominator appearing under the trace of Equation (C.11) reads

$$\Gamma_k^{(2)} + \mathcal{R}_k = p^2 + m^2(k) + k^2 + \lambda(k) \phi^2 + \dots \quad (\text{C.23})$$

Here we used a simple mass-type cutoff  $\mathcal{R}_k = k^2$  which is sufficient to make the point. In a loop calculation of  $\Gamma_k$  it is the inverse of Equation (C.23), evaluated at the vacuum expectation value  $\phi$  which appears as the effective propagator in all loops. It contains an IR cutoff at the scale  $k$ , a mass term  $k^2$  which adds to  $m^2(k)$  in the special case considered here. (In general  $\mathcal{R}_k \equiv \mathcal{R}_k(p^2)$  introduces a  $p^2$ -dependent mass.)

The  $p_\mu$ -modes (plane waves) are integrated out efficiently only in the domain  $p^2 \gtrsim m^2 + k^2 + \lambda \phi^2 + \dots$ . In the opposite case all loop contributions are suppressed by the effective mass square  $m^2 + k^2 + \lambda \phi^2 + \dots$ . It is the sum of the “artificial” cutoff  $k^2$ , introduced in order to effect the coarse graining, and the “physical” cutoff terms  $m^2(k) + \lambda(k) \phi^2 + \dots$ . As a consequence,  $\Gamma_k$  displays a significant dependence on  $k$  only if  $k^2 \gtrsim m^2(k) + \lambda(k) \phi^2 + \dots$  because otherwise  $k^2$  is negligible relative to  $m^2 + \lambda \phi^2 + \dots$  in all propagators; it is then the physical cutoff scale  $m^2 + \lambda \phi^2 + \dots$  which delimits the range of  $p^2$ -values which are integrated out.

Typically, for  $k$  very large,  $k^2$  is larger than the physical cutoffs so that  $\Gamma_k$  “runs” very fast. Lowering  $k$  it might happen that, at some  $k = k_{\text{dec}}$ , the “artificial” cutoff  $k$  becomes smaller than the running mass  $m(k)$ . At this point the physical mass starts playing the role of the actual cutoff; its effect overrides that of  $k$  so that  $\Gamma_k$  becomes approximately independent of  $k$  for  $k < k_{\text{dec}}$ . As a result,  $\Gamma_k \approx \Gamma_{k_{\text{dec}}}$  for all  $k$  below the threshold  $k_{\text{dec}}$ , and in particular the ordinary effective action  $\Gamma = \Gamma_0$  does not differ from  $\Gamma_{k_{\text{dec}}}$  significantly. This is the prototype of a “decoupling” or “freezing” phenomenon [208].

The situation is more interesting when  $m^2$  is negligible and  $k^2$  competes with  $\lambda \phi^2$  for the role of the actual cutoff. (Here we assume that  $\phi$  is  $x$ -independent.) The running of  $\Gamma_k$ , evaluated at a fixed  $\phi$ , stops once  $k \lesssim k_{\text{dec}}(\phi)$  where the by now field dependent decoupling scale obtains from the implicit equation  $k_{\text{dec}}^2 = \lambda(k_{\text{dec}}) \phi^2$ . Decoupling occurs for sufficiently large values of  $\phi$ , the RG evolution below  $k_{\text{dec}}$  is negligible then; hence, at  $k = 0$ ,

$$\Gamma[\phi] = \Gamma_k[\phi] \Big|_{k=k_{\text{dec}}(\phi)}. \quad (\text{C.24})$$

Equation (C.24) is an extremely useful tool for effectively going beyond the truncation (C.22) without having to derive and solve a more complicated flow equation. In fact, thanks to the

additional  $\phi$ -dependence which comes into play via  $k_{\text{dec}}(\phi)$ , Equation (C.24) can predict certain terms which are contained in  $\Gamma$  even though they are not present in the truncation ansatz.

A simple example illustrates this point. For  $k$  large, the truncation (C.22) yields a logarithmic running of the  $\Phi^4$ -coupling:  $\lambda(k) \propto \ln(k)$ . As a result, Equation (C.24) suggests that  $\Gamma$  should contain a term  $\propto \ln(k_{\text{dec}}(\phi)) \phi^4$ . Since, in leading order,  $k_{\text{dec}} \propto \phi$ , this leads us to the prediction of a  $\phi^4 \ln(\phi)$ -term in the conventional effective action. This prediction, including the prefactor of the term, is known to be correct: The Coleman–Weinberg potential of massless  $\phi^4$ -theory does indeed contain this  $\phi^4 \ln(\phi)$ -term. Note that this term is not analytic in  $\phi$ , so it lies outside the space of functionals spanned by the a power series ansatz like Equation (C.22).

This example illustrates the power of decoupling arguments. They can be applied even when  $\phi$  is taken  $x$ -dependent as it is necessary for computing  $n$ -point functions by differentiating  $\Gamma_k[\phi]$ . The running inverse propagator is given by  $\Gamma_k^{(2)}(x-y) = \delta^2 \Gamma_k / \delta\phi(x) \delta\phi(y)$ , for example. Here a new potential cutoff scale enters the game: the momentum  $q$  dual to the distance  $x-y$ . When it serves as the operative IR cutoff in the denominator of the multiply differentiated FRGE, the running of  $\Gamma_k^{(2)}(q)$ , the Fourier transform of  $\Gamma_k^{(2)}(x-y)$ , stops once  $k^2$  is smaller than  $k_{\text{dec}}^2 = q^2$ . Hence  $\Gamma_k^{(2)}(q) \approx \Gamma_k^{(2)}(q)|_{k=\sqrt{q^2}}$  for  $k^2 \lesssim q^2$ , provided no other physical scales intervene. As a result, if one allows for a running  $Z$ -factor in the truncation (C.22) one predicts a propagator of the type  $[Z(\sqrt{q^2})q^2]^{-1}$  in the standard effective action. Note that generically it corresponds to a nonlocal term  $\propto \int \phi Z(\sqrt{-\partial^2}) \partial^2 \phi$  in  $\Gamma$ , even though the truncation ansatz was local.

In the context of the effective average action formalism for gravity this kind of reasoning [135, 134] also underlies the evaluation of the UV behavior of the propagators in the “anomalous dimension argument” of Section 2.4. If  $\eta_N$  is approximately constant, the graviton  $Z$ -factor varies as  $Z_N(k) \sim k^{-\eta_N}$ , and the corresponding propagator  $[Z(\sqrt{q^2})q^2]^{-1}$  is proportional to  $(q^2)^{-1+\eta_N/2}$  in momentum space and to  $(\sqrt{x^2})^{2-d-\eta_N}$  in position space.

In the literature similar arguments have been used for the “renormalization group improvement” of cosmological [37, 38, 183, 28] and black hole spacetimes [36, 35, 40] on the basis of the effective average action (see also [189, 191, 190] for a discussion of different improvement schemes).



## References

- [1] Abbott, L.F., Grisaru, M.T., and Schaefer, R.K., “The background field method and the  $S$ -matrix”, *Nucl. Phys. B*, **229**, 372–380, (1983). [B.2.3](#), [B.2.3](#)
- [2] Adams, A., Arkani-Hamed, N., Dubovsky, S., Nicolis, A., and Rattazzi, R., “Causality, Analyticity and an IR Obstruction to UV Completion”, (2006). URL (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0602178>. [1.1](#)
- [3] Aida, T., Kitazawa, Y., Kawai, H., and Ninomiya, M., “Conformal invariance and renormalization group in quantum gravity near two-dimensions”, *Nucl. Phys. B*, **427**, 158, (1994). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9404171>. [a](#), [a](#)
- [4] Alekseev, G.A., “Monodromy data parametrization of the spaces of local solutions of integrable reductions of Einstein’s field equations”, *Theor. Math. Phys.*, **143**, 720, (2005). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/0503043>. *Teor. Mat. Fiz.* 143 (2005) 278. [2](#), [3.1.5](#)
- [5] Ambjørn, J., Jurkiewicz, J., and Loll, R., “Dynamically triangulating Lorentzian quantum gravity”, *Nucl. Phys. B*, **610**, 347, (2001). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0105267>. [1.1](#), [1.4.1](#)
- [6] Ambjørn, J., Jurkiewicz, J., and Loll, R., “Emergence of a 4D world from causal quantum gravity”, *Phys. Rev. Lett.*, **93**, 131301, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0404156>. [1.4.1](#)
- [7] Ambjørn, J., Jurkiewicz, J., and Loll, R., “Spectral dimension of the universe”, *Phys. Rev. Lett.*, **95**, 171301, (2005). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0505113>. [1.4.1](#)
- [8] Ambjørn, J., Jurkiewicz, J., and Watabiki, Y., “Dynamical triangulations, a gateway to quantum gravity?”, *J. Math. Phys.*, **36**, 6299–6339, (1995). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9503108>. [1.4.1](#), [2.3.5](#), [2.4.4](#)
- [9] Anderson, I.M., and Torre, C.G., “Classification of generalized symmetries for the vacuum Einstein equations”, *Commun. Math. Phys.*, **176**, 479, (1996). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/9404030>. [1.1](#)
- [10] Anselmi, D., “Absence of higher derivatives in the renormalization of propagators in quantum field theories with infinitely many couplings”, *Class. Quantum Grav.*, **20**, 2355–2378, (2003). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0212013>. [1.1](#), [1.5](#), [2.3.2](#), [2.3.2](#), [2.3.2](#)
- [11] Anselmi, D., “Infinite reduction of couplings in non-renormalizable quantum field theory”, *J. High Energy Phys.*, **2005**(08), 029, (2005). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0503131>. [1.1](#), [3.3](#)
- [12] Antoniadis, I., and Mottola, E., “Four dimensional quantum gravity in the conformal sector”, *Phys. Rev. D*, **45**, 2013, (1992). [2.3.3](#)



- [13] Arnone, S., Morris, R.T., and Rosten, O.J., “A generalised manifestly gauge invariant exact renormalisation group for SU(N) Yang–Mills”, (2005). URL (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0507154>. **C**
- [14] Ashtekar, A., and Lewandowski, J., “Representation theory of analytic holonomy  $C^*$  algebras”, in Baez, J.C., ed., *Knots and Quantum Gravity*, Proceedings of a workshop held at UC Riverside on May 14–16, 1993, vol. 1 of Oxford Lecture Series in Mathematics and its Applications, 21–61, (Clarendon Press; Oxford University Press, Oxford, U.K.; New York, U.S.A., 1994). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/9311010>. **1.4.2**
- [15] Ashtekar, A., and Lewandowski, J., “Background independent quantum gravity: A status report”, *Class. Quantum Grav.*, **21**, R53–R152, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/0404018>. **1.4.2**
- [16] Atance, M., and Cortes, J.L., “Effective field theory of gravity, reduction of couplings and the renormalization group”, *Phys. Rev. D*, **54**, 4973, (1996). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-ph/9605455>. **1.1**
- [17] Avramidi, I.G., “Covariant techniques for computation of the heat kernel”, *Rev. Math. Phys.*, **11**, 947–980, (1999). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9704166>. **2.4.4**
- [18] Avramidi, I.G., *Covariant Methods for the Calculation of the Effective Action in Quantum Field Theory and Investigation of Higher-Derivative Quantum Gravity*, Ph.D. Thesis, (Moscow State University, Moscow, USSR, 1986). URL (cited on 05 October 2006): <http://arXiv.org/abs/hep-th/9510140>. English translation by the author 1995. **2.3.2**
- [19] Avramidi, I.G., and Barvinsky, A.O., “Asymptotic Freedom In Higher Derivative Quantum Gravity”, *Phys. Lett. B*, **159**, 269, (1985). **1.2, b, 2.3.2**
- [20] Babelon, O., and Viallet, C., “The Riemannian geometry of the configuration space of gauge theories”, *Commun. Math. Phys.*, **81**, 515, (1981). **2.3.3**
- [21] Bagnuls, C., and Bervillier, C., “Exact renormalization group equations: An introductory review”, *Phys. Rep.*, **348**, 91, (2001). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0002034>. **2.2, A, 2, 3, 4, 5**
- [22] Balog, J., and Niedermaier, M., “Off-shell dynamics of the O(3) NLS model beyond Monte Carlo and perturbation theory”, *Nucl. Phys. B*, **500**, 421, (1997). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9612039>. **2.1**
- [23] Barbero G, J.F., Mena Marugán, G.A., and Villasenor, E.J.S., “Particles and vacuum for perturbative and non-perturbative Einstein–Rosen gravity”, *Phys. Rev. D*, **70**, 044028, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/0406087>. **3.2.1**
- [24] Baumann, K., “On canonical irreducible quantum field theories describing bosons and fermions”, *J. Math. Phys.*, **29**, 1225, (1988). **1.4.2**
- [25] Becchi, C., and Collina, R., “Further comments on the background field method and gauge-invariant effective actions”, *Nucl. Phys. B*, **562**, 412–430, (1999). **B.2.3, B.2.3**

- [26] Belinski, V., and Verdaguer, E., *Gravitational solitons*, (Cambridge University Press, Cambridge, U.K.; New York, U.S.A., 2001). [c, 2](#)
- [27] Ben-Avraham, D., and Havlin, S., *Diffusion and reactions in fractals and disordered systems*, (Cambridge University Press, Cambridge, U.K.; New York, U.S.A., 2000). [2.4.4](#)
- [28] Bentivegna, E., Bonanno, A., and Reuter, M., “Confronting the IR fixed point cosmology with high redshift supernova data”, *J. Cosmol. Astropart. Phys.*, **2004**(01), 001, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/astro-ph/0303150>. [C.3](#)
- [29] Berges, J., Tetradis, N., and Wetterich, C., “Non-perturbative renormalization flow in quantum field theory and statistical physics”, *Phys. Rep.*, **363**, 223, (2002). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-ph/0005122>. [2.2, 2.2, C, C.1, 3, C.2, 4, 5](#)
- [30] Bern, Z., “Perturbative Quantum Gravity and its Relation to Gauge Theory”, *Living Rev. Relativity*, **5**, lrr-2002-5, (2002). URL (cited on 15 May 2006): <http://www.livingreviews.org/lrr-2002-5>. [1.1](#)
- [31] Bern, Z., Mottola, E., and Blau, S.K., “General covariance of the path integral for quantum gravity”, *Phys. Rev. D*, **43**, 1212, (1991). [2.3.3, 2.3.3, 2.3.3](#)
- [32] Bjerrum-Bohr, N.E.J., Donoghue, F.J., and Holstein, B.R., “Quantum gravitational corrections to the nonrelativistic scattering potential of two masses”, *Phys. Rev. D*, **67**, 1–12, 084033, (2003). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0211072>. Erratum: *Phys. Rev. D*, **71**, 069903, (2005). [1.1, 1.1, 1.5](#)
- [33] Bjerrum-Bohr, N.E.J., Donoghue, J.F., and Holstein, B.R., “On the parameterization dependence of the energy momentum tensor and the metric”, (2006). URL (cited on 14 November 2006): <http://arXiv.org/abs/gr-qc/0610096>. [1.5](#)
- [34] Blaizot, J.P., Mendez-Galain, R., and Wschebor, N., “Non perturbative renormalisation group and momentum dependence of n-point functions (I)”, (2005). URL (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0512317>. [3, 3](#)
- [35] Bonanno, A., and Reuter, M., “Quantum gravity effects near the null black hole singularity”, *Phys. Rev. D*, **60**, 084011, (1999). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/9811026>. [C.3](#)
- [36] Bonanno, A., and Reuter, M., “Renormalization group improved black hole spacetimes”, *Phys. Rev. D*, **62**, 043008, (2000). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0002196>. [C.3](#)
- [37] Bonanno, A., and Reuter, M., “Cosmology of the Planck era from a renormalization group for quantum gravity”, *Phys. Rev. D*, **65**, 043508, (2002). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0106133>. [C.3](#)
- [38] Bonanno, A., and Reuter, M., “Cosmological perturbations in renormalization group derived cosmologies”, *Int. J. Mod. Phys. D*, **13**, 107, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/astro-ph/0210472>. [C.3](#)

- [39] Bonanno, A., and Reuter, M., “Proper time flow equation for gravity”, *J. High Energy Phys.*, **2005**(02), 035, (2005). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0410191>. [d](#), [4.4.4](#)
- [40] Bonanno, A., and Reuter, M., “Spacetime structure of an evaporating black hole in quantum gravity”, *Phys. Rev. D*, **73**, 083005, (2006). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0602159>. [C.3](#)
- [41] Bonneau, G., and Delduc, F., “Nonlinear renormalization and the equivalence theorem”, *Nucl. Phys. B*, **266**, 536, (1986). [2.1](#), [3.3](#)
- [42] Boulware, D., “Gauge dependence of the effective action”, *Phys. Rev. D*, **23**, 389–396, (1981). [6](#)
- [43] Bovie, A., and Felder, G., “Skeleton inequalities and the asymptotic nature of perturbation theory for  $\phi^4$  theories in two-dimensions and three-dimensions”, *Commun. Math. Phys.*, **93**, 259, (1984). [2.1](#)
- [44] Branchina, V., Meissner, A.K., and Veneziano, G., “The price of an exact, gauge-invariant RG-flow equation”, *Phys. Lett. B*, **574**, 319, (2003). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0309234>. [C.2](#), [4](#)
- [45] Breitenlohner, P., Gibbons, G., and Maison, D., “4-Dimensional Black Holes from Kaluza–Klein Reduction”, *Commun. Math. Phys.*, **120**, 295–333, (1988). [c](#), [3.1.1](#), [3.1.1](#), [3.1.2](#), [3.1.2](#), [3.1.2](#)
- [46] Breitenlohner, P., and Maison, D., “On nonlinear sigma-models arising in (super-)gravity”, *Commun. Math. Phys.*, **209**, 785, (2000). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/9806002>. [c](#), [3.1.1](#), [3.1.1](#)
- [47] Brezin, E., Hikami, S., and Zinn-Justin, J., “Generalized non-linear sigma-models with gauge invariance”, *Nucl. Phys. B*, **165**, 528, (1980). [3.4](#)
- [48] Brydges, C.D., Mitter, K.P., and Scoppola, B., “Critical  $\phi_{3,\epsilon}^4$ ”, *Commun. Math. Phys.*, **240**, 281, (2003). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0206040>. [2.2](#)
- [49] Buchbinder, I.L., Odintsov, S.D., and Shapiro, I.L., *Effective Action in Quantum Gravity*, (Institute of Physics Publishing, Bristol, U.K., 1992). [2.3.2](#), [B.2.1](#)
- [50] Burgess, C.P., “Quantum Gravity in Everyday Life: General Relativity as an Effective Field Theory”, *Living Rev. Relativity*, **7**, lrr-2004-5, (2004). URL (cited on 15 May 2006): <http://www.livingreviews.org/lrr-2004-5>. [1.1](#)
- [51] Burgess, C.P., and Kunstatter, G., “On the physical interpretation of the Vilkovisky–DeWitt effective action”, *Mod. Phys. Lett. A*, **2**, 875–886, (1987). Erratum: *Mod. Phys. Lett. A*, **2**, 1003, (1987). [2.3.4](#), [B.2.2](#), [B.2.2](#)
- [52] Chow, B., and Knopf, D., *The Ricci Flow: An Introduction*, vol. 110 of Mathematical Surveys and Monographs, (American Mathematical Society, Providence, U.S.A., 2004). [6](#)
- [53] Christensen, S.M., and Duff, M.J., “Quantum gravity in  $2 + \epsilon$  dimensions”, *Phys. Lett. B*, **79**, 213, (1978). [a](#)

- [54] Codello, A., and Percacci, R., “Fixed points of higher derivative gravity”, (2006). URL (cited on 05 October 2006):  
<http://arXiv.org/abs/hep-th/0607128>. 1.2, 2.3.2
- [55] Cooper, A., Feldman, J., and Rosen, L., “Legendre transforms and  $r$ -particle irreducibility in quantum field theory: the formal power series framework”, *Ann. Phys. (N.Y.)*, **137**, 213, (1981). B.1
- [56] Cremmer, E., Julia, B., Lü, H., and Pope, C., “Higher dimensional origin of  $D = 3$  coset symmetries”, (1999). URL (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/9909099>. c, 3.1.1, 3.1.2, 3.1.2
- [57] Curci, G., and Paffuti, G., “Consistency between the string background field equations of motion and the vanishing of the conformal anomaly”, *Nucl. Phys. B*, **286**, 399, (1987). c, B.3, B.3, B.3, B.3, B.3
- [58] de Bakker, B.V., and Smit, J., “Gravitational binding in 4D dynamical triangulation”, *Nucl. Phys. B*, **484**, 476–492, (1997). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-lat/9604023>. 2.3.5
- [59] de Berredo-Peixoto, G., and Shapiro, I.L., “Higher derivative quantum gravity with Gauss–Bonnet term”, *Phys. Rev. D*, **71**, 064005, (2005). Related online version (cited on 05 October 2006):  
<http://arXiv.org/abs/hep-th/0412249>. 1.2, b, 2.3.2, 2.3.2, 2.3.2
- [60] de Calan, C., Faria da Veiga, P., Magnen, J., and Seneor, R., “Constructing the three-dimensional Gross–Neveu model with a large number of flavor components”, *Phys. Rev. Lett.*, **66**, 3233, (1991). 1.3
- [61] de Wit, B., Grisaru, M.T., Nicolai, H., and Rabinovici, E., “Two loop finiteness of  $d = 2$  supergravity”, *Phys. Lett.*, **286**, 78, (1992). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/9205012>. 3.3
- [62] Deser, S., and Jackiw, R., “Energy-momentum improvements in two dimensions”, *Int. J. Mod. Phys. B*, **10**, 1499, (1996). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/9510145>. 3.1.4, 3.1.4
- [63] Deser, S., and van Nieuwenhuizen, P., “One loop divergencies of quantized Einstein–Maxwell fields”, *Phys. Rev. D*, **10**, 401, (1974). 1.1
- [64] DeWitt, B.S., “Quantum Theory of Gravity. I. The Canonical Theory”, *Phys. Rev.*, **160**, 1113–1148, (1967). 1.1
- [65] DeWitt, B.S., “Quantum Theory of Gravity. II. The Manifestly Covariant Theory”, *Phys. Rev.*, **162**, 1195–1239, (1967). 1.1
- [66] DeWitt, B.S., “Quantum Theory of Gravity. III. Applications of the Covariant Theory”, *Phys. Rev.*, **162**, 1239–1256, (1967). 1.1
- [67] DeWitt, B.S., “Approximative Effective Action for Quantum Gravity”, *Phys. Rev. Lett.*, **47**, 1647–1650, (1981). 6
- [68] DeWitt, B.S., *The Global Approach to Quantum Field Theory, 2 vols.*, vol. 114 of International Series of Monographs on Physics, (Oxford University Press, Oxford, U.K.; New York, U.S.A., 2003). 6, B.1, B.2.1, B.2.4

- [69] Dimock, J., “Asymptotic perturbation expansion in the  $P_2(\phi)$  theory”, *Commun. Math. Phys.*, **35**, 347, (1974). [2.1](#)
- [70] Dittrich, B., “Partial and Complete Observables for Canonical General Relativity”, (2005). URL (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/0507106>. [1.1](#), [3.1.5](#)
- [71] Dittrich, W., “Bloch–Nordsieck approximation in linearized quantum gravity”, in Semikhatov, A., Vasiliev, M., and Zaikin, V., eds., *Quantization, Gauge theory, and Strings, Vol. II*, Proceedings of the International Conference dedicated to the memory of Professor Efim Fradkin, Moscow, Russia, June 5–10, 2000, (Scientific World, Moscow, Russia, 2001). [c](#), [3.2.1](#)
- [72] Dou, D., and Percacci, R., “The running gravitational couplings”, *Class. Quantum Grav.*, **15**, 3449, (1998). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9707239>. [d](#), [4.3](#)
- [73] Fabbrichesi, M., Pettorino, R., Veneziano, G., and Vilkovisky, G.A., “Planckian energy scattering and surface terms in the gravitational action”, *Nucl. Phys. B*, **419**, 147, (1994). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9309037>. [c](#), [3.2.1](#)
- [74] Fatibene, L., Ferraris, M., Francaviglia, M., and Raiteri, M., “Noether charges, Brown–York quasilocal energy and related topics”, *J. Math. Phys.*, **42**, 1173–1195, (2001). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/0003019>. [B.2](#)
- [75] Felder, G., “Renormalization group in the local potential approximation”, *Commun. Math. Phys.*, **111**, 101–121, (1987). [2.2](#)
- [76] Fels, M.E., and Torre, C.G., “The principle of symmetric criticality”, *Class. Quantum Grav.*, **19**, 641–676, (2002). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/0108033>. [1](#)
- [77] Fernández, R., Fröhlich, J., and Sokal, A.D., *Random walks, critical phenomena, and triviality in quantum field theory*, (Springer, Berlin, Germany; New York, U.S.A., 1992). [1.4.2](#), [A](#), [4](#)
- [78] Feynman, R.P., “Quantum theory of gravitation”, *Acta Phys. Pol.*, **24**, 697–722, (1963). [1.1](#)
- [79] Fischer, P., and Litim, D.F., “Fixed points of quantum gravity in extra dimensions”, *Phys. Lett. B*, **638**, 497, (2006). Related online version (cited on 05 October 2006): <http://arXiv.org/abs/hep-lat/0602203>. [4.4.4](#), [4.4.4](#)
- [80] Fokas, A., and Mohammadi, N., “Three-loop calculation of the beta function for the purely metric nonlinear sigma-model”, *Phys. Lett. B*, **198**, 359, (1987). [B.3](#), [B.3](#)
- [81] Ford, L.H., “Gravitational radiation by quantum systems”, *Ann. Phys. (N.Y.)*, **144**, 238–248, (1982). [1.1](#)
- [82] Forgács, P., and Niedermaier, M., “A fixed point for truncated quantum Einstein gravity”, (2002). URL (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0207028>. [1.1](#)

- [83] Fradkin, E.S., and Tseytlin, A.A., “Higher Derivative Quantum Gravity: One Loop Counterterms And Asymptotic Freedom”, *Nucl. Phys. B*, **201**, 469, (1982). Related online version (cited on 05 October 2006):  
<http://arXiv.org/abs/gr-qc/0601137>. 1.2, b, 2.3.2
- [84] Friedan, D.H., “Nonlinear models in  $2+\epsilon$  dimensions”, *Ann. Phys. (N.Y.)*, **163**, 318, (1985).  
c, 3.3, B.3, B.3
- [85] Garay, L.J., “Quantum gravity and minimum length”, *Int. J. Mod. Phys. A*, **10**, 145–166, (1995). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/gr-qc/9403008>. 1.5
- [86] Gastmans, R., Kallosh, R., and Truffin, C., “Quantum gravity near two-dimensions”, *Nucl. Phys. B*, **133**, 417, (1978). a
- [87] Gawedski, K., and Kupainen, A., “Exact renormalization in Gross–Neveu model of quantum fields”, *Phys. Rev. Lett.*, **54**, 2191, (1985). 2.1
- [88] Gawedski, K., and Kupainen, A., “Renormalizing the nonrenormalizable”, *Phys. Rev. Lett.*, **55**, 363, (1985). 2
- [89] Gawedski, K., and Kupiainen, A., “Gross–Neveu model through convergent perturbation expansions”, *Commun. Math. Phys.*, **102**, 1, (1985). 2.1
- [90] Geroch, R., “A Method for Generating Solutions of Einstein’s Equations”, *J. Math. Phys.*, **12**, 918–924, (1971). 3.1.2, 3.1.5
- [91] Geroch, R., “A Method for Generating New Solutions of Einstein’s Equation. II”, *J. Math. Phys.*, **13**, 394–404, (1972). 3.1.2, 3.1.5
- [92] Gibbons, G.W., and Hawking, S.W., eds., *Euclidean Quantum Gravity*, (World Scientific, Singapore; River Edge, U.S.A., 1993). 1.1
- [93] Gies, H., “Renormalizability of gauge theories in extra dimensions”, *Phys. Rev. D*, **68**, 085015, (2003). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/0305208>. 1.3
- [94] Gomis, J., and Weinberg, S., “Are nonrenormalizable gauge theories renormalizable?”, *Nucl. Phys. B*, **469**, 473–487, (1996). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/9510087>. 1.1, 1.1, 1.3, 1.5, 2.3.2, 2.3.2, 2.3.2, 2.4.1, 3.1
- [95] Goroff, M.H., and Sagnotti, A., “The ultraviolet behavior of Einstein gravity”, *Nucl. Phys. B*, **266**, 709–736, (1986). 2.3.2
- [96] Graham, S., “Three-loop beta function for the bosonic nonlinear sigma-model”, *Phys. Lett. B*, **197**, 543, (1987). B.3
- [97] Green, M.B., Schwarz, J.H., and Witten, E., *Superstring Theory, 2 vols.*, Cambridge Monographs on Mathematical Physics, (Cambridge University Press, Cambridge, U.K., New York, U.S.A., 1987). 1, 1.4.3
- [98] Griffiths, J.B., *Colliding plane waves in general relativity*, (Clarendon Press, Oxford, U.K.; New York, U.S.A., 1991). c, 2
- [99] Haag, R., *Local Quantum Physics: Fields, Particles, Algebras*, (Springer, Berlin, Germany; New York, U.S.A., 1996), 2nd edition. 1.4.2



- [100] Hamber, H.W., and Williams, R.M., “Newtonian potential in quantum Regge gravity”, *Nucl. Phys. B*, **435**, 361–398, (1995). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9406163>. 1.5
- [101] Hands, S., Kocić, A., and Kogut, J.B., “Four-Fermi Theories in Fewer Than Four Dimensions”, *Ann. Phys. (N.Y.)*, **224**, 29–89, (1993). 1.3
- [102] Hasenfratz, P., “The theoretical background and properties of perfect actions”, (1998). URL (cited on 15 May 2006): <http://arXiv.org/abs/hep-lat/9803027>. A, 5
- [103] Hauser, I., and Ernst, F.J., “Proof of a generalized Geroch conjecture for the hyperbolic Ernst equation”, *Gen. Relativ. Gravit.*, **33**, 195–293, (2001). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/0002049>. 2, 3.1.5
- [104] Hawking, S.W., “Quantum gravity and path integrals”, *Phys. Rev. D*, **18**, 1747, (1978). 1.1
- [105] Hofling, F., Nowak, C., and Wetterich, C., “Phase transition and critical behaviour of the  $d = 3$  Gross–Neveu model”, *Phys. Rev. B*, **66**, 205111, (2002). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/cond-mat/0203588>. 1.3
- [106] Holstein, B.R., and Donoghue, J.F., “Classical physics and quantum loops”, *Phys. Rev. Lett.*, **93**, 201602, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0405239>. 2.1, 2.1
- [107] Honerkamp, J., “Chiral multi-loops”, *Nucl. Phys. B*, **36**, 130, (1972). B.2.2
- [108] Honerkamp, J., “The question of invariant renormalizability of massless YM theory in a manifestly covariant approach”, *Nucl. Phys. B*, **48**, 269, (1972). B.2.3
- [109] Honerkamp, J., Krause, F., and Scheunert, M., “On the equivalence of standard and covariant perturbation series in non-polynomial pion Lagrangian field theory”, *Nucl. Phys. B*, **69**, 618, (1974). B.2.2
- [110] Howe, P., Papadopoulos, G., and Stelle, K.S., “Background field method and the nonlinear  $\sigma$ -model”, *Nucl. Phys.*, **296**, 26, (1988). c, 3.3, B.2.1, B.2.1, B.2.2, B.2.2, B.3, B.3, B.3, B.3, B.3, B.3
- [111] Huang, K., *Quantum Field Theory: From Operators to Path Integrals*, (Wiley-Interscience, New York, U.S.A., 1998). A, 5
- [112] Johnson, C.V., *D-Branes*, Cambridge Monographs on Mathematical Physics, (Cambridge University Press, Cambridge, U.K., New York, U.S.A., 2002). 1, 1.4.3
- [113] Kabat, D., and Ortiz, M., “Eikonal quantum gravity and Planckian scattering”, *Nucl. Phys. B*, **388**, 570, (1992). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9203082>. c, 3.2.1
- [114] Kadanoff, L., *Statistical Physics: Statics, Dynamics and Renormalization*, (World Scientific, Singapore; River Edge, U.S.A., 2000). A
- [115] Kawai, H., Kawamoto, N., Mogami, T., and Watabiki, Y., “Transfer matrix formalism for two-dimensional quantum gravity and fractal structures of space-time”, *Phys. Lett. B*, **306**, 19, (1993). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9302133>. 2.3.5, 2.4.4



- [116] Kawai, H., Kitazawa, Y., and Ninomiya, M., “Scaling exponents in quantum gravity near two-dimensions”, *Nucl. Phys. B*, **393**, 280, (1993). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/9206081>. 1.1, 1.1, a, a, 2.3.1, 2.4.2, 4.4.4
- [117] Kawai, H., Kitazawa, Y., and Ninomiya, M., “Ultraviolet stable fixed point and scaling relations in  $(2 + \epsilon)$ -dimensional quantum gravity”, *Nucl. Phys. B*, **404**, 684, (1993). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/9303123>. 1.1, a, a, 2.4.2
- [118] Kawai, H., and Ninomiya, M., “Renormalization group and quantum gravity”, *Nucl. Phys. B*, **336**, 115–145, (1990). a, a, 2.3.5, 2.4.4
- [119] Kirilin, G., “Quantum corrections to the Schwarzschild metric and reparameterization transformations”, (2006). URL (cited on 15 May 2006):  
<http://arXiv.org/abs/gr-qc/0601020>. 1.5
- [120] Klebanov, I.R., Kogan, I.I., and Polyakov, A.M., “Gravitational dressing of renormalization group”, *Phys. Rev. Lett.*, **71**, 3243–3246, (1993). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/9309106>. 3.4
- [121] Klein, C., and Richter, O., *Ernst Equation and Riemann Surfaces: Analytical and Numerical Methods*, vol. 685 of Lecture Notes in Physics, (Springer, Berlin, Germany; New York, U.S.A., 2005). c, 2
- [122] Knizhnik, G.V., Polyakov, M.A., and Zamolodchikov, A.B., “Fractal structure of 2d-Quantum Gravity”, *Mod. Phys. Lett. A*, **3**, 819, (1988). a
- [123] Kopper, C., “Renormalization theory based on flow equations”, (2005). URL (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/0508143>. 2.2, B.1
- [124] Korotkin, D., and Samtleben, H., “Poisson realization and quantization of the Geroch group”, *Class. Quantum Grav.*, **14**, L151, (1997). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/gr-qc/9611061>. 3.1.5
- [125] Kröger, H., “Fractal geometry in quantum mechanics, field theory, and spin systems”, *Phys. Rep.*, **323**, 81–181, (2000). 1.1, 2.4.2
- [126] Kubo, J., and Nunami, M., “Unrenormalizable theories are predictive”, *Eur. Phys. J. C*, **26**, 461, (2003). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/0112032>. 1.1
- [127] Kuchař, K., “Dirac constraint quantization of a parametrized field theory by anomaly-free operator representations of space-time diffeomorphisms”, *Phys. Rev. D*, **39**, 2263–2280, (1989). 3.4
- [128] Kuchař, K.V., Romano, D.J., and Varadarajan, M., “Dirac constraint quantization of a dilatonic model of gravitational collapse”, *Phys. Rev. D*, **55**, 795–808, (1997). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/gr-qc/9608011>. 3.4
- [129] Kunstatter, G., “The Path integral for gauge theories: A Geometrical approach”, *Class. Quantum Grav.*, **9**, S157, (1992). B.2.1, B.2.4

- [130] Latorre, J.I., and Morris, T.R., “Exact scheme independence”, *J. High Energy Phys.*, **2000**(11), 004, (2000). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0008123>. [2](#)
- [131] Lauscher, O., and Reuter, M., “Flow equation of quantum Einstein gravity in a higher-derivative truncation”, *Phys. Rev. D*, **66**, 025026, (2002). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0205062>. [d](#), [d](#), [1.3](#), [1.5](#), [2.4.2](#), [4.1.1](#), [4.3](#), [4.3](#), [4.3](#), [4.3](#), [4.4.2](#), [4.4.4](#), [4.4.4](#), [4.4.4](#), [4.4.5](#), [B.2.4](#)
- [132] Lauscher, O., and Reuter, M., “Is Quantum Einstein Gravity nonperturbatively renormalizable?”, *Class. Quantum Grav.*, **19**, 483, (2002). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0110021>. [4.1.1](#), [4.3](#), [3](#), [4](#), [4.4.4](#)
- [133] Lauscher, O., and Reuter, M., “Ultraviolet fixed point and generalized flow equations of Quantum Gravity”, *Phys. Rev. D*, **65**, 025013, (2002). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0108040>. [1.1](#), [1.1](#), [d](#), [d](#), [1.3](#), [2.4.2](#), [4.1.1](#), [4.3](#), [4.3](#), [4.3](#), [4.3](#), [4.4.2](#), [4.4.4](#), [4.4.4](#), [4.4.4](#), [B.2.4](#)
- [134] Lauscher, O., and Reuter, M., “Asymptotic safety in quantum Einstein gravity: Nonperturbative renormalizability and fractal spacetime structure”, (2005). URL (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0511260>. [1.1](#), [2.4.4](#), [2.4.4](#), [4.2](#), [C.3](#)
- [135] Lauscher, O., and Reuter, M., “Fractal spacetime structure in asymptotically safe gravity”, *J. High Energy Phys.*, **2005**(10), 050, (2005). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0508202>. [1.1](#), [2.4.4](#), [2.4.4](#), [4.2](#), [C.3](#)
- [136] Litim, D.F., “Fixed points of quantum gravity”, *Phys. Rev. Lett.*, **92**, 201301, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0312114>. [d](#), [2.3.2](#), [4.3](#), [4.4.4](#)
- [137] Litim, D.F., and Pawłowski, J.M., “Flow equations for Yang–Mills theories in general axial gauges”, *Phys. Lett. B*, **435**, 181–188, (1998). Related online version (cited on 05 October 2006): <http://arXiv.org/abs/hep-th/9802064>. [4.3](#)
- [138] Loll, R., “Discrete Approaches to Quantum Gravity in Four Dimensions”, *Living Rev. Relativity*, **1**, lrr-1998-13, (1998). URL (cited on 15 May 2006): <http://www.livingreviews.org/lrr-1998-13>. [1.4.1](#)
- [139] Lüscher, M., “Quantum nonlocal charges and absence of particle production in the two-dimensional nonlinear  $\sigma$ -model”, *Nucl. Phys. B*, **135**, 1, (1978). [3.1.5](#)
- [140] Marolf, D., and Mourão, J.M., “On the support of the Ashtekar–Lewandowski measure”, *Commun. Math. Phys.*, **170**, 583–606, (1995). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9403112>. [1.4.2](#)
- [141] Mattei, F., Rovelli, C., Speziale, S., and Testa, M., “From 3-geometry transition amplitudes to graviton states”, *Nucl. Phys. B*, **739**, 234–253, (2006). Related online version (cited on 05 October 2006): <http://arXiv.org/abs/gr-qc/0508007>. [3.2.1](#)

- [142] Mazur, P.O., and Mottola, E., “The gravitational measure, solution of the conformal factor problem and stability of the ground state of quantum gravity”, *Nucl. Phys. B*, **341**, 187, (1990). [2.3.3](#), [3.2.3](#)
- [143] Mead, C.A., “Possible Connection Between Gravitation and Fundamental Length”, *Phys. Rev.*, **135**, B849–B862, (1964). [1.5](#)
- [144] Misner, C., “Feynman quantization of general relativity”, *Rev. Mod. Phys.*, **29**, 497–509, (1957). [1.1](#)
- [145] Mitter, P.K., “The Exact Renormalization Group”, (2005). URL (cited on 15 May 2006): <http://arXiv.org/abs/math-ph/0505008>. [2.2](#)
- [146] Morris, T.R., “Elements of the continuous renormalization group”, *Prog. Theor. Phys. Suppl.*, **131**, 395, (1998). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9802039>. [2.2](#), [A](#), [C.1](#), [4](#)
- [147] Morris, T.R., “Equivalence of local potential approximations”, *J. High Energy Phys.*, **2005**(07), 027, (2005). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0503161>. [4](#), [5](#)
- [148] Morris, T.R., “Renormalizable extra-dimensional models”, *J. High Energy Phys.*, **2005**(01), 002, (2005). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-ph/0410142>. [1.3](#)
- [149] Mottola, E., “Functional integration over geometries”, *J. Math. Phys.*, **36**, 2470, (1995). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9502109>. [1.3](#), [2.3.3](#)
- [150] Nicolai, H., and Niedermaier, M., unknown status, (1990). [3.1.5](#)
- [151] Nicolai, H., Peeters, K., and Zamaklar, M., “Loop quantum gravity: an outside view”, *Class. Quantum Grav.*, **22**, R193–R247, (2005). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0501114>. [1.4.2](#), [1.5](#)
- [152] Nicoll, J.F., and Chang, T.S., “An exact one particle irreducible renormalization group generator for critical phenomena”, *Phys. Lett. A*, **62**, 287, (1977). [C](#), [C.2](#)
- [153] Niedermaier, M., “Quantized Einstein–Rosen waves, AdS(2), and spontaneous symmetry breaking”, *Phys. Lett. B*, **498**, 83, (2001). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0007227>. [3.4](#)
- [154] Niedermaier, M., “Renormalization and asymptotic safety in truncated quantum Einstein gravity”, *J. High Energy Phys.*, **2002**(12), 066, (2002). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0207143>. [c](#), [c](#), [1.3](#), [1.3](#), [3.1.2](#), [3.2.2](#), [3.3](#), [3.3](#), [3.3](#), [3.3](#), [3.3](#), [3.4](#), [3.4](#), [B.2.4](#)
- [155] Niedermaier, M., “Dimensionally reduced gravity theories are asymptotically safe”, *Nucl. Phys. B*, **673**, 131, (2003). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0304117>. [c](#), [3.2.2](#), [3.3](#), [3.3](#), [3.3](#), [3.4](#), [3.4](#), [3.4](#), [B.2.4](#)
- [156] Niedermaier, M., unknown format, (2005). Unpublished notes. [1.4.2](#), [2.3.4](#), [3.1.4](#), [3.2](#), [3.2.3](#), [3.4](#), [4.2](#), [B.2.3](#)

- [157] Niedermaier, M., “The asymptotic safety scenario in quantum gravity – an introduction”, (2006). URL (cited on 05 October 2006):  
<http://arXiv.org/abs/gr-qc/0610018>. 1, 1.1, 1.1, d, 2.2, 2.2, 2.3.2, 2.3.4, 2.3.4, 2.3.5, 2.3.5, 2.4.4, 2.4.4, C.1
- [158] Niedermaier, M., and Samtleben, H., “An algebraic bootstrap for dimensionally reduced quantum gravity”, *Nucl. Phys. B*, **579**, 437, (2000). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/9912111>. c, 3.1.5
- [159] Niedermayer, F., Niedermaier, M., and Weisz, P., “Questionable and unquestionable in the perturbation theory of non-Abelian models”, *Phys. Rev. D*, **56**, 2555–2565, (1997). Related online version (cited on 05 October 2006):  
<http://arXiv.org/abs/hep-lat/9612002>. 2.1
- [160] Oehme, R., and Zimmermann, W., “Relation between effective couplings for asymptotically free models”, *Commun. Math. Phys.*, **97**, 569, (1985). 1.1, 1.5, 3.3
- [161] O’Raifeartaigh, L., Wipf, A., and Yoneyama, H., “The Constraint Effective Potential”, *Nucl. Phys. B*, **271**, 653, (1986). C
- [162] Osborn, H., “Renormalization and composite operators in nonlinear  $\sigma$ -models”, *Nucl. Phys. B*, **294**, 595, (1987). c, B.2.2, B.3, B.3, B.3, B.3, B.3, B.3, B.3, B.3, B.3, B.3
- [163] Padmanabhan, T., “Limitations on the operational definition of space-time events and quantum gravity”, *Class. Quantum Grav.*, **4**, L107–L113, (1987). 1.5
- [164] Padmanabhan, T., “From Gravitons to Gravity: Myths and Reality”, (2004). URL (cited on 05 October 2006):  
<http://arXiv.org/abs/gr-qc/0409089>. 3.2.1
- [165] Pawłowski, J.M., “Geometrical effective action and Wilsonian flows”, (2003). URL (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/0310018>. 4
- [166] Pawłowski, J.M., “Aspects of the functional renormalisation group”, (2005). URL (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/0512261>. 2.2, 4
- [167] Peldan, P., “Actions for gravity, with generalizations: a review”, *Class. Quantum Grav.*, **11**, 1087, (1994). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/gr-qc/9305011>. 1.3
- [168] Penati, S., Santambrogio, A., and Zanon, D., “Renormalization group flows in sigma-models coupled to 2D dynamical gravity”, *Nucl. Phys. B*, **483**, 495, (1997). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/9605124>. 3.4
- [169] Percacci, R., “Further evidence for a gravitational fixed point”, *Phys. Rev. D*, **73**, 041501, (2006). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/0511177>. b, 4.3
- [170] Percacci, R., and Perini, D., “Asymptotic safety of gravity coupled to matter”, *Phys. Rev. D*, **68**, 044018, (2003). Related online version (cited on 15 May 2006):  
<http://arXiv.org/abs/hep-th/0304222>. d, 4.3

- [171] Percacci, R., and Perini, D., “Constraints on matter from asymptotic safety”, *Phys. Rev. D*, **67**, 081503, (2003). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0207033>. [d](#), [4.3](#)
- [172] Percacci, R., and Perini, D., “Should we expect a fixed point for Newton’s constant?”, *Class. Quantum Grav.*, **21**, 5035, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0401071>. [1.1](#), [1.5](#)
- [173] Perez, A., “On the regularization ambiguities in loop quantum gravity”, *Phys. Rev. D*, **73**, 044007, (2006). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/0509118>. [1.4.2](#), [1.5](#)
- [174] Perry, R.J., and Wilson, K.G., “Perturbative renormalizability with an infinite number of relevant and marginal operators”, *Nucl. Phys. B*, **403**, 587, (1993). [1.1](#), [1.1](#), [1.5](#), [3.3](#)
- [175] Peskin, M.E., and Schroeder, D.V., *An Introduction to Quantum Field Theory*, (Westview Press, Boulder, U.S.A., 1995). [1.1](#)
- [176] Polchinski, J., “Scale and conformal invariance in quantum field theory”, *Nucl. Phys. B*, **303**, 226–236, (1988). [3.4](#)
- [177] Polchinski, J.G., *String Theory, 2 vols.*, Cambridge Monographs on Mathematical Physics, (Cambridge University Press, Cambridge, U.K., New York, U.S.A., 1998). [1](#), [1.4.3](#)
- [178] Rebhan, A., “Feynman rules and  $S$ -matrix equivalence of the Vilkovisky–DeWitt effective action”, *Nucl. Phys. B*, **298**, 726–740, (1988). [B.2.1](#), [B.2.4](#)
- [179] Reuter, M., “Nonperturbative Evolution Equation for Quantum Gravity”, *Phys. Rev. D*, **57**, 971, (1998). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9605030>. [d](#), [d](#), [4.1](#), [3](#), [5](#), [4.3](#), [4.3](#), [4.3](#), [4.4.4](#), [B.2.4](#), [4](#)
- [180] Reuter, M., and Saueressig, F., “A class of nonlocal truncations in quantum Einstein gravity and its renormalization group behavior”, *Phys. Rev. D*, **66**, 125001, (2002). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0206145>. [4.3](#), [4.4.5](#)
- [181] Reuter, M., and Saueressig, F., “Renormalization group flow of quantum gravity in the Einstein–Hilbert truncation”, *Phys. Rev. D*, **65**, 065016, (2002). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0110054>. [4.3](#), [4.3](#), [4.4.1](#), [2](#), [4.4.3](#), [4.4.4](#), [4.4.4](#)
- [182] Reuter, M., and Saueressig, F., “Nonlocal quantum gravity and the size of the universe”, *Fortschr. Phys.*, **52**, 650, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0311056>. [4.4.5](#)
- [183] Reuter, M., and Saueressig, F., “From big bang to asymptotic de Sitter: Complete cosmologies in a quantum gravity framework”, *J. Cosmol. Astropart. Phys.*, **2005**(09), 012, (2005). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0507167>. [C.3](#)
- [184] Reuter, M., and Schwindt, J.M., “A minimal length from the cutoff modes in asymptotically safe quantum gravity”, *J. High Energy Phys.*, **2006**(01), 070, (2006). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0511021>. [4.2](#), [4.2](#)

- [185] Reuter, M., and Wetterich, C., “Average action for the Higgs model with Abelian gauge symmetry”, *Nucl. Phys. B*, **391**, 147, (1993). [C](#)
- [186] Reuter, M., and Wetterich, C., “Running gauge coupling in three-dimensions and the electroweak phase transition”, *Nucl. Phys. B*, **408**, 91, (1993). [C](#)
- [187] Reuter, M., and Wetterich, C., “Effective average action for gauge theories and exact evolution equations”, *Nucl. Phys. B*, **417**, 181, (1994). [4.1](#), [C](#)
- [188] Reuter, M., and Wetterich, C., “Quantum Liouville field theory as solution of a flow equation”, *Nucl. Phys. B*, **506**, 483, (1997). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9605039>. [3](#)
- [189] Reuter, M., and Weyer, H., “Quantum gravity at astrophysical distances?”, *J. Cosmol. Astropart. Phys.*, **2004**(12), 001, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0410119>. [C.3](#), [C.3](#)
- [190] Reuter, M., and Weyer, H., “Renormalization group improved gravitational actions: A Brans–Dicke approach”, *Phys. Rev. D*, **69**, 104022, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0311196>. [C.3](#), [C.3](#)
- [191] Reuter, M., and Weyer, H., “Running Newton constant, improved gravitational actions, and galaxy rotation curves”, *Phys. Rev. D*, **70**, 124028, (2004). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0410117>. [C.3](#), [C.3](#)
- [192] Riedel, E.K., Golner, G.R., and Newman, K.E., “Scaling-Field Representation of Wilson’s Exact Renormalization-Group Equation”, *Ann. Phys. (N.Y.)*, **161**, 178–238, (1985). [2](#), [3](#)
- [193] Ringwald, A., and Wetterich, C., “Average action for the  $N$ -component  $\phi^4$  theory”, *Nucl. Phys. B*, **334**, 506–526, (1990). [C](#)
- [194] Riva, V., and Cardy, J., “Scale and conformal invariance in field theory: A physical counterexample”, *Phys. Lett. B*, **622**, 339, (2005). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0504197>. [3.4](#)
- [195] Rivasseau, V., *From Perturbative to Constructive Renormalization*, (Princeton University Press, Princeton, U.S.A., 1991). [2.2](#), [A](#)
- [196] Robinson, S., “Normalization conventions for Newton’s constant and the Planck scale in arbitrary dimensions”, (2006). URL (cited on 05 October 2006): <http://arXiv.org/abs/hep-th/0609060>. [2.3.1](#)
- [197] Romano, J.D., and Torre, C.G., “Internal time formalism for spacetimes with two Killing vectors”, *Phys. Rev. D*, **53**, 5634–5650, (1996). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/9509055>. [c](#), [3.1.3](#), [3.1.4](#), [3.1.4](#)
- [198] Rosa, L., Vitale, P., and Wetterich, C., “Critical exponents of the Gross–Neveu model from the effective average action”, *Phys. Rev. Lett.*, **86**, 958, (2001). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0007093>. [1.3](#)
- [199] Rovelli, C., *Quantum Gravity*, Cambridge Monographs on Mathematical Physics, (Cambridge University Press, Cambridge, U.K.; New York, U.S.A., 2004). [1](#), [1.4.2](#), [1.5](#)



- [200] Salmhofer, M., *Renormalization: An Introduction*, (Springer, Berlin, Germany; New York, U.S.A., 1999). [2.2](#)
- [201] Shore, G.M., “A local renormalization group equation, diffeomorphisms and conformal invariance in sigma models”, *Nucl. Phys. B*, **286**, 349–377, (1987). [c, 5](#), [B.2.2](#), [B.3](#), [B.3](#), [B.3](#), [B.3](#), [B.3](#)
- [202] Shore, G.M., “Quantum gravitational optics”, *Contemp. Phys.*, **44**, 503, (2003). Related online version (cited on 05 October 2006): <http://arXiv.org/abs/gr-qc/0304059>. [2.3.5](#)
- [203] Smolin, L., “A fixed point for quantum gravity”, *Nucl. Phys. B*, **208**, 439–466, (1982). [1.1](#), [b, 2.4.3](#), [4.3](#)
- [204] Souma, W., “Non-trivial UV fixed point in quantum gravity”, *Prog. Theor. Phys.*, **102**, 181, (1999). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9907027>. [d, 4.4.4](#)
- [205] Souma, W., “Gauge and cutoff function dependence of the ultraviolet fixed point in quantum gravity”, (2000). URL (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/0006008>. [d, 4.3](#), [4.4.4](#)
- [206] Stelle, K.S., “Renormalization of higher derivative Quantum Gravity”, *Phys. Rev. D*, **16**, 953, (1977). [1.2](#), [b, 2.3.2](#), [2.3.2](#)
- [207] Stelle, K.S., “Classical Gravity With Higher Derivatives”, *Gen. Relativ. Gravit.*, **9**, 353, (1978). [1.2](#), [2.3.2](#)
- [208] Sumi, I.J., Souma, W., Aoki, I.K., Terao, H., and Morikawa, K., “Scheme dependence of the Wilsonian effective action and sharp cutoff limit of the flow equation”, (2000). URL (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/0002231>. [C.3](#)
- [209] ’t Hooft, G., “Graviton dominance in ultrahigh-energy scattering”, *Phys. Lett. B*, **198**, 61, (1987). [c, 3.2.1](#)
- [210] ’t Hooft, G., and Veltman, M.J.G., “One loop divergencies in the theory of gravitation”, *Ann. Inst. Henri Poincaré A*, **20**, 69–94, (1974). [1.1](#), [2.3.2](#)
- [211] Taylor, T.R., and Veneziano, G., “Quantum gravity at large distances and the cosmological constant”, *Nucl. Phys. B*, **345**, 210, (1990). [4.4.5](#)
- [212] Teitelboim, C., “The proper time gauge in quantum theory of gravitation”, *Phys. Rev. D*, **28**, 297, (1983). [1.1](#), [1.1](#)
- [213] Teitelboim, C., “Quantum mechanics of the gravitational field in asymptotically flat space”, *Phys. Rev. D*, **28**, 310, (1983). [1.1](#), [1.1](#)
- [214] Teitelboim, C., “The Hamiltonian structure of two-dimensional space-time and its relation with the conformal anomaly”, in Christensen, S.M., ed., *Quantum Theory of Gravity: Essays in Honor of the 60th Birthday of Bryce S. DeWitt*, (Adam Hilger, Bristol, U.K., 1984). [3.1.3](#), [3.1.4](#)
- [215] Thompson, R.T., and Ford, L.H., “Spectral line broadening and angular blurring due to spacetime geometry fluctuations”, *Phys. Rev. D*, **74**, 024012, (2006). Related online version (cited on 05 October 2006): <http://arXiv.org/abs/gr-qc/0601137>. [2.3.5](#)



- [216] Tomboulis, E., “ $1/N$  expansion and renormalization in quantum gravity”, *Phys. Lett. B*, **70**, 361–364, (1977). [1.1](#), [b](#), [2.4.3](#), [4.3](#)
- [217] Tomboulis, E., “Renormalizability and asymptotic freedom in quantum gravity”, *Phys. Lett. B*, **97**, 77–80, (1980). [1.1](#), [b](#), [2.4.3](#), [4.3](#)
- [218] Tomboulis, E.T., “Superrenormalizable gauge and gravitational theories”, (1997). URL (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9702146>. [1.5](#)
- [219] Torre, C.G., “Gravitational observables and local symmetries”, *Phys. Rev. D*, **48**, 2373–2376, (1993). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/gr-qc/9306030>. [1.1](#)
- [220] Tseytlin, A.A., “Sigma-model Weyl invariance conditions and string equations of motion”, *Nucl. Phys. B*, **294**, 383, (1987). [c](#), [5](#), [B.3](#), [B.3](#), [B.3](#), [B.3](#), [B.3](#), [B.3](#)
- [221] Tseytlin, A.A., “Finite  $\sigma$  models and exact string solutions with Minkowski signature metric”, *Phys. Rev. D*, **47**, 3421–3429, (1993). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9211061>. [3](#), [3.3](#)
- [222] van de Ven, A.E.M., “Two loop quantum gravity”, *Nucl. Phys. B*, **378**, 309–366, (1992). [2.3.2](#)
- [223] Vilkovisky, G.A., “The Gospel according to deWitt”, in Christensen, S.M., ed., *Quantum Theory of Gravity: Essays in Honor of the 60th Birthday of Bryce S. DeWitt*, (Adam Hilger, Bristol, U.K., 1984). [B.1](#), [B.2.1](#), [B.2.4](#)
- [224] Vilkovisky, G.A., “The unique effective action in quantum field theory”, *Nucl. Phys. B*, **234**, 125, (1984). [B.2.1](#)
- [225] Vilkovisky, G.A., “Heat kernel: Rencontre entre physiciens et mathématiciens”, (1991). URL (cited on 15 May 2006): <http://cdsweb.cern.ch/search.py?recid=234641>. [2.4.4](#)
- [226] Ward, B.F.L., “Quantum corrections to Newton’s law”, *Mod. Phys. Lett. A*, **17**, 2371, (2002). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-ph/0204102>. [1.5](#)
- [227] Weinberg, S., “Ultraviolet divergencies in quantum theories of gravitation”, in Hawking, S.W., and Israel, W., eds., *General Relativity: An Einstein Centenary Survey*, (Cambridge University Press, Cambridge, U.K.; New York, U.S.A., 1979). [1.1](#), [1.1](#), [1.1](#), [a](#), [4.4.1](#), [3](#)
- [228] Wetterich, C., “Exact evolution equation for the effective potential”, *Phys. Lett. B*, **301**, 90, (1993). [2.2](#), [C](#), [C.2](#), [C.2](#), [4](#)
- [229] Wetterich, C., “Effective average action in statistical physics and quantum field theory”, *Int. J. Mod. Phys. A*, **16**, 1951, (2001). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-ph/0101178>. [2.2](#), [2.2](#), [C](#), [C.2](#), [4](#)
- [230] Wilczek, F., “Quantum field theory”, *Rev. Mod. Phys.*, **71**, S85, (1999). Related online version (cited on 15 May 2006): <http://arXiv.org/abs/hep-th/9803075>. [1.1](#)
- [231] Wilson, K.G., “Renormalization Group and Critical Phenomena. I. Renormalization Group and the Kadanoff Scaling Picture”, *Phys. Rev. B*, **4**, 3174–3183, (1971). [A](#)

- [232] Wilson, K.G., “Renormalization Group and Critical Phenomena. II. Phase-Space Cell Analysis of Critical Behavior”, *Phys. Rev. B*, **4**, 3184–3205, (1971). [A](#)
- [233] Wilson, K.G., “Quantum Field-Theory Models in Less Than 4 Dimensions”, *Phys. Rev. D*, **7**, 2911–2926, (1973). [A](#)
- [234] Wilson, K.G., “The Renormalization Group and Critical Phenomena”, in Frängsmyr, T., and Ekspong, G., eds., *Nobel Lectures, Physics 1981–1990*, (World Scientific, Singapore; River Edge, U.S.A., 1993). Related online version (cited on 15 May 2006): <http://nobelprize.org/physics/laureates/1982/wilson-lecture.html>. Nobel lecture of December 8, 1982. [A](#)
- [235] York Jr, J.W., “Conformally invariant orthogonal decomposition of symmetric tensors on Riemannian manifolds and the initial-value problem of general relativity”, *J. Math. Phys.*, **14**, 456–464, (1973). [2.3.3](#)
- [236] Zimmermann, W., “Reduction in the number of coupling parameters”, *Commun. Math. Phys.*, **97**, 211, (1985). [1.1](#), [1.5](#), [3.3](#)
- [237] Zinn-Justin, J., *Quantum Field Theory and Critical Phenomena*, vol. 77 of International Series of Monographs on Physics, (Clarendon Press; Oxford University Press, Oxford, U.K.; New York, U.S.A., 1989), 3rd edition. [3.1.1](#), [3](#), [B.1](#), [B.1](#)