

Discrete and Continuum Quantum Gravity

Herbert W. Hamber*

*Max Planck Institute for Gravitational Physics (Albert Einstein Institute),
D-14476 Potsdam, Germany*

I review discrete and continuum approaches to quantized gravity, based on the covariant Feynman path integral approach.

*Electronic address: Herbert.Hamber@aei.mpg.de

Contents

I. INTRODUCTION	4
II. CONTINUUM FORMULATION	8
A. General Aspects	8
1. Massless Spin Two Field	8
2. Wave Equation	9
3. Feynman Rules	15
4. One-Loop Divergences	21
5. Higher Derivative Terms	26
6. Supergravity	35
B. Feynman Path Integral Formulation	39
1. Sum over Paths	40
2. Euclidean Rotation	42
3. Gravitational Functional Measure	44
4. Conformal Instability	48
C. Gravity in $2 + \epsilon$ Dimensions	51
1. Perturbatively Non-renormalizable Theories: The Sigma Model	52
2. Non-linear Sigma Model in the Large-N Limit	62
3. Self-coupled Fermion Model	66
4. The Gravitational Case	68
5. Phases of Gravity in $2 + \epsilon$ Dimensions	75
6. Running of $\alpha(\mu)$ in <i>QED</i> and <i>QCD</i>	81
III. LATTICE REGULARIZED QUANTUM GRAVITY	84
A. General Formulation	86
1. Volumes and Angles	86
2. Rotations, Parallel Transports and Voronoi Loops	89
3. Invariant Lattice Action	95
4. Lattice Diffeomorphism Invariance	101
5. Lattice Bianchi Identities	103
6. Gravitational Wilson Loop	104
7. Lattice Regularized Path Integral	108
8. An Elementary Example	113
9. Lattice Higher Derivative Terms	116
10. Scalar Matter Fields	121
11. Invariance Properties of the Scalar Action	127
12. Lattice Fermions, Tetrads and Spin Rotations	129
13. Alternate Discrete Formulations	131
B. Analytical Expansion Methods	135

1. Lattice Weak Field Expansion and Transverse-Traceless Modes	136
2. Strong Coupling Expansion	145
3. Discrete Gravity in the Large-d Limit	151
IV. NUMERICAL STUDIES IN FOUR DIMENSIONS	160
A. Observables, Phase Structure and Critical Exponents	162
1. Invariant Local Gravitational Averages	163
2. Invariant Correlations at Fixed Geodesic Distance	166
3. Wilson Lines and Static Potential	168
4. Scaling in the Vicinity of the Critical Point	172
5. Physical and Unphysical Phases	174
6. Numerical Determination of the Scaling Exponents	179
7. Renormalization Group and Lattice Continuum Limit	186
8. Curvature Scales	190
9. Gravitational Condensate	192
V. SCALE DEPENDENT GRAVITATIONAL COUPLING	194
A. Effective Field Equations	195
1. Scale Dependence of G	195
2. Poisson's Equation and Vacuum Polarization Cloud	198
3. Static Isotropic Solution	200
4. Cosmological Solutions	204
VI. CONCLUSIONS and OUTLOOK	210
Acknowledgments	212
References	213

I. INTRODUCTION

In this review article I will attempt to cover key aspects and open issues related to a consistent lattice regularized formulation of quantum gravity. Such a formulation can be viewed as an important and perhaps essential step towards a quantitative, controlled investigation of the physical content of the theory. The main emphasis of the present review will therefore rest on discrete and continuum space-time formulations of quantum gravity based on the covariant Feynman path integral approach, and their mutual interrelation.

The first part of the review will introduce the basic elements of a covariant formulation of continuum quantum gravity, with special emphasis on those issues which bear some immediate relevance for the remainder of the work. These include a discussion of the nature of the spin-two field, its wave equation and possible gauge choices, the Feynman propagator, the coupling of a spin two field to matter, and the implementation of a consistent local gauge invariance to all orders, ultimately leading to the Einstein action. Additional terms in the gravitational action, such as the cosmological constant and higher derivative contributions, are naturally introduced at this stage, and play some role later in the context of a full quantum theory.

A section on the perturbative (weak field) expansion introduces the main aspects of the background field method as applied to gravity, including the choice of field parametrization and gauge fixing terms. Later the results on the structure of one- and two-loop divergences in pure gravity are discussed, leading up to the conclusion of perturbative non-renormalizability for the Einstein theory in four dimensions. The relevant one-loop and two-loop counterterms will be recalled. One important aspect that needs to be emphasized is that these perturbative methods generally rely on a weak field expansion for the metric fluctuations, and are therefore not well suited for investigating the (potentially physically relevant) regime where metric fluctuations can be large.

Later the Feynman path integral for gravitation is introduced, in analogy with the closely related case of Yang-Mills theories. This, of course, brings up the thorny issue of the gravitational functional measure, expressing Feynman's sum over geometries, as well as important aspects related to the convergence of the path integral and derived quantum averages, and the origin of the conformal instability affecting the Euclidean case. An important point that needs to be emphasized is the strongly constrained nature of the theory, which depends, in the absence of matter, and as in pure Yang-Mills theories, on a single dimensionless parameter $G\lambda$, besides a required short distance cutoff.

Since quantum gravity is not perturbatively renormalizable, the following question arises nat-

urally: what other theories are not perturbatively renormalizable, and what can be done with them? The following parts will therefore summarize the methods of the $2 + \epsilon$ expansion for gravity, an expansion in the deviation of the space-time dimensions from two, where the gravitational coupling is dimensionless and the theory appears therefore power-counting renormalizable. As initial motivation, but also for illustrative and pedagogical purposes, the non-linear sigma model will be introduced first. The latter represents a reasonably well understood perturbatively non-renormalizable theory above two dimensions, characterized by a rich two-phase structure, and whose scaling properties in the vicinity of the fixed point can nevertheless be accurately computed (via the $2 + \epsilon$ expansion, as well as by other methods, including most notably the lattice) in three dimensions, and whose universal predictions are known to compare favorably with experiments. In the gravity context, to be discussed next, the main results of the perturbative expansion are the existence of a nontrivial ultraviolet fixed point close to the origin above two dimensions (a phase transition in statistical field theory language), and the predictions of universal scaling exponents in the vicinity of this fixed point.

The next sections deal with the natural lattice discretization for quantum gravity based on Regge's simplicial formulation, with a primary focus on the physically relevant four-dimensional case. The starting point there is a description of a discrete manifold in terms of edge lengths and incidence matrices, then moving on to a description of curvature in terms of deficit angles, thereby offering a re-formulation of continuum gravity in terms of the discrete Regge action and ensuing lattice field equations. The direct and clear correspondence between lattice quantities (edges, dihedral angles, volumes, deficit angles, etc.) and continuum operators (metric, affine connection, volume element, curvature tensor etc.) will be emphasized all along. The latter will be useful in defining, as an example, discrete formulations of curvature squared terms which arise in higher derivative gravity theories, or more generally as radiatively induced corrections. An important element in this lattice-to-continuum correspondence will be the development of the lattice weak field expansion, allowing in this context again a clear and precise identification between lattice and continuum degrees of freedom, as well as their gauge invariance properties, as illustrated in the weak field limit by the arbitrariness in the assignments of edge lengths used to cover a given physical geometry. The lattice analogs of gravitons arise naturally, and their transverse-traceless nature (in a suitable gauge) can easily be made manifest.

When coupling matter fields to lattice gravity one needs to introduce new fields localized on vertices, as well as appropriate dual volumes which enter the definition of the kinetic terms for those fields. The discrete re-parametrization invariance properties of the discrete matter action

will be described next. In the fermion case, it is necessary (as in the continuum) to introduce vierbein fields within each simplex, and then use an appropriate spin rotation matrix to relate spinors between neighboring simplices. In general the formulation of fractional spin fields on a simplicial lattice could have some use in the lattice discretization of supergravity theories. At this point it will also be useful to compare, and contrast, Regge's simplicial formulation to other discrete approaches to quantum gravity, such as the hypercubic (vierbein-connection) lattice formulation, and fixed-edge-length approaches such as dynamical triangulations.

The next sections deal with the interesting problem of what gravitational observables should look like, i.e. which expectation values of operators (or ratios thereof) have meaning and physical interpretation in the context of a manifestly covariant formulation, specifically in a situation where metric fluctuations are not necessarily bounded. Such averages naturally include expectation values of the (integrated) scalar curvature and other related quantities (involving for example curvature squared terms), as well as correlations of operators at fixed geodesic distance, sometimes referred to as bi-local operators. Another set of physical averages refer to the geometric nature of space-time itself, such as the fractal dimension. One more set of physical observables correspond to the gravitational analog of the Wilson loop (providing information about the parallel transport of vectors, and therefore on the effective curvature, around large near-planar loops), and the correlation between particle world-lines (providing information about the static gravitational potential). It is reasonable to expect that these quantities will play an important role in the physical characterization of the two phases of gravity, as seen both in the $2 + \epsilon$ and in the lattice formulation in four dimensions.

There are reasons to believe that ultimately the investigation of a strongly coupled regime of quantum gravity, where metric fluctuations cannot be assumed to be small, requires the use of numerical methods applied to the lattice theory. A discrete formulation combined with numerical tools can therefore be viewed as an essential step towards a quantitative and controlled investigation of the physical content of the theory: in the same way that a discretization of a complicated ordinary differential equation can be viewed as a mean to determine the properties of its solution with arbitrary accuracy. These methods are outlined next, together with a summary of the main lattice results, suggesting the existence of two phases (depending on the value of the bare gravitational coupling) and in agreement with the qualitative predictions of the $2 + \epsilon$ expansion. Specifically one finds a weak coupling degenerate polymer-like phase, and a strong coupling smooth phase with bounded curvatures in four dimensions. The somewhat technical aspect of the determination of universal critical exponents and non-trivial scaling dimensions, based on finite size methods, is

outline, together with a detailed (but by now standard) discussion of how the lattice continuum limit has to be approached in the vicinity of a non-trivial ultraviolet fixed point.

The determination of non-trivial scaling dimensions in the vicinity of the fixed point opens the door to a discussion of the renormalization group properties of fundamental couplings, i.e. their scale dependence, as well as the emergence of physical renormalization group invariant quantities, such as the gravitational correlation length and the closely related gravitational condensate. Such topics will be discussed next, with an eye towards perhaps more physical applications. These include a discussion on the physical nature of the expected quantum corrections to the gravitational coupling, based, in part on an analogy to qed and qcd, on the effects of a virtual graviton cloud (as already suggested in the $2 + \epsilon$ expansion context), and of how the two phases of lattice gravity relate to the two opposite scenarios of gravitational screening (for weak coupling, and therefore unphysical due to the branched polymer nature of this phase) versus anti-screening (for strong coupling, and therefore physical).

A final section touches on the general problem of formulating running gravitational couplings in a context that does not assume weak gravitational fields and close to flat space at short distance. The discussion includes a brief presentation on the topic of covariant running of G based on the formalism of non-local field equations, with the scale dependence of G expressed through the use of a suitable covariant d'Alembertian. Simple applications to standard metrics (static isotropic and homogeneous isotropic) are briefly summarized and their potential physical consequences and interpretation elaborated.

The review will end with a general outlook on future prospects for lattice studies of quantum gravity, some open questions and work that can be done to help elucidate the relationship between discrete and continuum models, such as extending the range of problems addressed by the lattice, and providing new impetus for further developments in covariant continuum quantum gravity.

Notation: Throughout this work, unless stated otherwise, the same notation is used as in (Weinberg, 1973), with the sign of the Riemann tensor reversed. The signature is therefore $-, +, +, +$. In the Euclidean case $t = -i\tau$ the flat metric is of course the Kronecker $\delta_{\mu\nu}$, with the same conventions as before for Riemann.

II. CONTINUUM FORMULATION

A. General Aspects

The Lagrangian for the massless spin-two field can be constructed in close analogy to what one does in the case of electromagnetism. In gravity the electromagnetic interaction $e j \cdot A$ is replaced by a term

$$\frac{1}{2} \kappa h_{\mu\nu}(x) T^{\mu\nu}(x) \quad (1)$$

where κ is a constant to be determined later, $T^{\mu\nu}$ is the conserved energy-momentum tensor

$$\partial_\mu T^{\mu\nu}(x) = 0 \quad (2)$$

associated with the sources, and $h_{\mu\nu}(x)$ describes the gravitational field. It will be shown later that κ is related to Newton's constant G by $\kappa = \sqrt{16\pi G}$.

1. Massless Spin Two Field

As far as the pure gravity part of the action is concerned, one has in principle four independent quadratic terms one can construct out of the first derivatives of $h_{\mu\nu}$, namely

$$\begin{aligned} \partial_\sigma h_{\mu\nu} \partial^\sigma h^{\mu\nu} & \quad , \quad \partial^\nu h_{\mu\nu} \partial_\sigma h^{\mu\sigma} \quad , \\ \partial^\nu h_{\mu\nu} \partial^\mu h_\sigma{}^\sigma & \quad , \quad \partial^\mu h_\nu{}^\nu \partial_\mu h_\sigma{}^\sigma \quad . \end{aligned} \quad (3)$$

The term $\partial_\sigma h_{\mu\nu} \partial^\nu h^{\mu\sigma}$ need not be considered separately, as it can be shown to be equivalent to the second term in the above list, after integration by parts. After combining these four terms into an action

$$\begin{aligned} \int dx \quad [& a \partial_\sigma h_{\mu\nu} \partial^\sigma h^{\mu\nu} + b \partial^\nu h_{\mu\nu} \partial_\sigma h^{\mu\sigma} \\ & + c \partial^\nu h_{\mu\nu} \partial^\mu h_\sigma{}^\sigma + d \partial^\mu h_\nu{}^\nu \partial_\mu h_\sigma{}^\sigma \\ & + \frac{1}{2} \kappa h_{\mu\nu} T^{\mu\nu}] \end{aligned} \quad (4)$$

and performing the required variation with respect to $h_{\alpha\beta}$, one obtains for the field equations

$$\begin{aligned} & 2a \partial_\sigma \partial^\sigma h_{\alpha\beta} \\ & + b (\partial_\beta \partial^\sigma h_{\alpha\sigma} + \partial_\alpha \partial^\sigma h_{\beta\sigma}) \\ & + c (\partial_\alpha \partial_\beta h_\sigma{}^\sigma + \eta_{\alpha\beta} \partial_\mu \partial_\nu h^{\mu\nu}) \\ & + 2d \eta_{\alpha\beta} \partial_\mu \partial^\mu h_\sigma{}^\sigma \\ & = \frac{1}{2} \kappa T_{\alpha\beta} \end{aligned} \quad (5)$$

with $\eta_{\alpha\beta} = \text{diag}(-1, 1, 1, 1)$. Consistency requires that the four-divergence of the above expression give zero on both sides, $\partial^\beta(\dots) = 0$. After collecting terms of the same type, one is led to the three conditions

$$\begin{aligned}(2a + b) \partial_\sigma \partial^\sigma \partial^\beta h_{\alpha\beta} &= 0 \\ (b + c) \partial_\alpha \partial^\beta \partial^\sigma h_{\beta\sigma} &= 0 \\ (c + 2d) \partial_\alpha \partial_\beta \partial^\beta h_\sigma{}^\sigma &= 0\end{aligned}\tag{6}$$

with unique solution (up to an overall constant, which can be reabsorbed into κ) $a = -\frac{1}{4}$, $b = \frac{1}{2}$, $c = -\frac{1}{2}$ and $d = \frac{1}{4}$. As a result, the quadratic part of the Lagrangian for the pure gravitational field is given by

$$\begin{aligned}\mathcal{L}_{sym} &= -\frac{1}{4} \partial_\sigma h_{\mu\nu} \partial^\sigma h^{\mu\nu} + \frac{1}{2} \partial^\nu h_{\mu\nu} \partial_\sigma h^{\mu\sigma} \\ &\quad - \frac{1}{2} \partial^\nu h_{\mu\nu} \partial^\mu h_\sigma{}^\sigma + \frac{1}{4} \partial^\mu h_\nu{}^\nu \partial_\mu h_\sigma{}^\sigma\end{aligned}\tag{7}$$

2. Wave Equation

One notices that the field equations of Eq. (5) take on a particularly simple form if one introduces trace reversed variables $\bar{h}_{\mu\nu}(x)$,

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h_\sigma{}^\sigma\tag{8}$$

and

$$\bar{T}_{\mu\nu} = T_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} T_\sigma{}^\sigma\tag{9}$$

In the following it will be convenient to write the trace as $h = h^\sigma{}_\sigma$ so that $\bar{h}^\sigma{}_\sigma = -h$, and define the d'Alembertian as $\square = \partial_\mu \partial^\mu = \nabla^2 - \partial_t^2$. Then the field equations become simply

$$\square h_{\mu\nu} - 2 \partial_\nu \partial^\sigma \bar{h}_{\mu\sigma} = -\kappa \bar{T}_{\mu\nu}\tag{10}$$

One important aspect of the field equations is that they can be shown to be invariant under a local gauge transformation of the type

$$h'_{\mu\nu} = h_{\mu\nu} + \partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu\tag{11}$$

involving an arbitrary gauge parameter $\epsilon_\mu(x)$. This invariance is therefore analogous to the local gauge invariance in QED, $A'_\mu = A_\mu + \partial_\mu \epsilon$. Furthermore, it suggests choosing a suitable gauge

(analogous to the familiar Lorentz gauge $\partial^\mu A_\mu = 0$) in order to simplify the field equations, for example

$$\partial^\sigma \bar{h}_{\mu\sigma} = 0 \quad (12)$$

which is usually referred to as the harmonic gauge condition. Then the field equations in this gauge become simply

$$\square h_{\mu\nu} = -\kappa \bar{T}_{\mu\nu} \quad (13)$$

These can then be easily solved in momentum space ($\square \rightarrow -k^2$) to give

$$h_{\mu\nu} = \kappa \frac{1}{k^2} \bar{T}_{\mu\nu} \quad (14)$$

or, in terms of the original $T_{\mu\nu}$,

$$h_{\mu\nu} = \kappa \frac{1}{k^2} (T_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} T_\sigma{}^\sigma) \quad (15)$$

It should be clear that this gauge is particularly convenient for practical calculations, since then graviton propagation is given simply by a factor of $1/k^2$; later on gauge choices will be introduced where this is no longer the case.

Next one can compute the amplitude for the interaction of two gravitational sources characterized by energy-momentum tensors T and T' . From Eqs. (1) and (14) one has

$$\frac{1}{2} \kappa T'_{\mu\nu} h^{\mu\nu} = \frac{1}{2} \kappa^2 T'_{\mu\nu} \frac{1}{k^2} \bar{T}^{\mu\nu} \quad (16)$$

which can be compared to the electromagnetism result $j'_\mu \frac{1}{k^2} j^\mu$.

To fix the value of the parameter κ it is easiest to look at the static case, for which the only non-vanishing component of $T_{\mu\nu}$ is T_{00} . Then

$$\frac{1}{2} \kappa^2 T'_{00} \frac{1}{k^2} \bar{T}^{00} = \frac{1}{2} \kappa^2 T'_{00} \frac{1}{k^2} (T_{00} - \frac{1}{2} \eta_{00} T_0{}^0) \quad (17)$$

For two bodies of mass M and M' the static instantaneous amplitude (by inverse Fourier transform, thus replacing $\frac{4\pi}{k^2} \rightarrow \frac{1}{r}$) then becomes

$$- \frac{1}{2} \kappa^2 \frac{1}{2} M' \frac{1}{4\pi r} M' \quad (18)$$

which, by comparison to the expected Newtonian potential energy $-GMM'/r$, gives the desired identification $\kappa = \sqrt{16\pi G}$.

The pure gravity part of the action in Eq. (7) only propagates transverse traceless modes (shear waves). These correspond quantum mechanically to a particle of zero mass and spin two, with two helicity states $h = \pm 2$, as shown for example in (Weinberg, 1973) by looking at the nature of plane wave solutions $h_{\mu\nu}(x) = e_{\mu\nu} e^{ik \cdot x}$ to the wave equation in the harmonic gauge. Helicity 0 and ± 1 appear initially, but can be made to vanish by a suitable choice of coordinates.

One would expect the gravitational field $h_{\mu\nu}$ to carry energy and momentum, which would be described by a tensor $\tau_{\mu\nu}(h)$. As in the case of electromagnetism, where one has

$$T_{\alpha\beta}^{(em)} = F_{\alpha\gamma} F_{\beta}{}^{\gamma} - \frac{1}{4} \eta_{\alpha\beta} F_{\gamma\delta} F^{\gamma\delta} \quad , \quad (19)$$

one would also expect such a tensor to be quadratic in the gravitational field $h_{\mu\nu}$. A suitable candidate for the energy-momentum tensor of the gravitational field is

$$\tau_{\mu\nu} = \frac{1}{8\pi G} \left(-\frac{1}{4} h_{\mu\nu} \partial^\lambda \partial_\lambda h^\sigma{}_\sigma + \dots \right) \quad (20)$$

where the dots indicate 37 possible additional terms, involving schematically, either terms of the type $h \partial^2 h$, or of the type $(\partial h)^2$. Such a $\tau_{\mu\nu}$ term would have to be added on the r.h.s. of the field equations in Eq. (10), and would therefore act as an additional source for the gravitational field (see Fig. 1). But the resulting field equations would then no longer invariant under Eq. (11), and one would have to change therefore the gauge transformation law by suitable terms of order h^2 , so as to ensure that the new field equations would still satisfy a local gauge invariance. In other words, all these complications arise because the gravitational field carries energy and momentum, and therefore gravitates.

Ultimately, a complete and satisfactory answer to these recursive attempts at constructing a consistent, locally gauge invariant, theory of the $h_{\mu\nu}$ field is found in Einstein's non-linear General Relativity theory, as shown in (Feynman, 1962; Boulware and Deser, 1969). The full theory is derived from the Einstein-Hilbert action

$$I_E = \frac{1}{16\pi G} \int dx \sqrt{g(x)} R(x) \quad (21)$$

which generalized Eq. (7) beyond the weak field limit. Here \sqrt{g} is the square root of the determinant of the metric field $g_{\mu\nu}(x)$, with $g = -\det g_{\mu\nu}$, and R the scalar curvature. The latter is related to the Ricci tensor $R_{\mu\nu}$ and the Riemann tensor $R_{\mu\nu\lambda\sigma}$ by

$$\begin{aligned} R_{\mu\nu} &= g^{\lambda\sigma} R_{\lambda\mu\sigma\nu} \\ R &= g^{\mu\nu} g^{\lambda\sigma} R_{\mu\lambda\nu\sigma} \end{aligned} \quad (22)$$

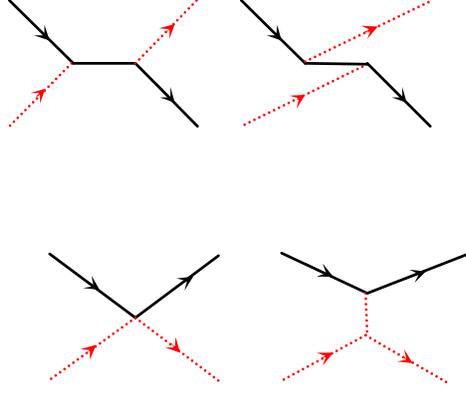


FIG. 1 Lowest order diagrams illustrating the gravitational analogue to Compton scattering. Continuous lines indicate a matter particle, short dashed lines a graviton. Consistency of the theory requires that the two bottom diagrams be added to the two on the top.

where $g^{\mu\nu}$ is the matrix inverse of $g_{\mu\nu}$,

$$g^{\mu\lambda} g_{\lambda\nu} = \delta^\mu_\nu \quad (23)$$

In terms of the affine connection $\Gamma^\lambda_{\mu\nu}$, the Riemann tensor $R^\lambda_{\mu\nu\sigma}(x)$ is given by

$$R^\lambda_{\mu\nu\sigma} = \partial_\nu \Gamma^\lambda_{\mu\sigma} - \partial_\sigma \Gamma^\lambda_{\mu\nu} + \Gamma^\eta_{\mu\sigma} \Gamma^\lambda_{\nu\eta} - \Gamma^\eta_{\mu\nu} \Gamma^\lambda_{\sigma\eta} \quad (24)$$

and therefore

$$R_{\mu\nu} = \partial_\sigma \Gamma^\sigma_{\mu\nu} - \partial_\nu \Gamma^\sigma_{\mu\sigma} + \Gamma^\lambda_{\sigma\lambda} \Gamma^\sigma_{\mu\nu} - \Gamma^\lambda_{\sigma\nu} \Gamma^\sigma_{\mu\lambda} \quad , \quad (25)$$

with the affine connection $\Gamma^\lambda_{\mu\nu}(x)$ in turn constructed from components of the metric field $g_{\mu\nu}$

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\sigma} \left(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu} \right) \quad (26)$$

The following algebraic symmetry properties of the Riemann tensor will be of use later

$$R_{\mu\nu\lambda\sigma} = -R_{\nu\mu\lambda\sigma} = -R_{\mu\nu\sigma\lambda} = R_{\nu\mu\sigma\lambda} \quad (27)$$

$$R_{\mu\nu\lambda\sigma} = R_{\lambda\sigma\mu\nu} \quad (28)$$

$$R_{\mu\nu\lambda\sigma} + R_{\mu\lambda\sigma\nu} + R_{\mu\sigma\nu\lambda} = 0 \quad (29)$$

In addition, the components of the Riemann tensor satisfy the differential Bianchi identities

$$\nabla_\alpha R_{\mu\nu\beta\gamma} + \nabla_\beta R_{\mu\nu\gamma\alpha} + \nabla_\gamma R_{\mu\nu\alpha\beta} = 0 \quad (30)$$

with ∇_μ the covariant derivative. It is known that these, in their contracted form, ensure the consistency of the field equations. From the expansion of the Einstein-Hilbert gravitational action in powers of the deviation of the metric from the flat metric $\eta_{\mu\nu}$, using

$$\begin{aligned} R_{\mu\nu} &= \frac{1}{2}(\partial^2 h_{\mu\nu} - \partial_\alpha \partial_\mu h^\alpha_\nu - \partial_\alpha \partial_\nu h^\alpha_\mu + \partial_\mu \partial_\nu h^\alpha_\alpha) + O(h^2) \\ R &= \partial^2 h^\mu_\mu - \partial_\alpha \partial_\mu h^{\alpha\mu} + O(h^2) \end{aligned} \quad (31)$$

one has for the action contribution

$$\begin{aligned} \sqrt{g} R &= -\frac{1}{4} \partial_\sigma h_{\mu\nu} \partial^\sigma h^{\mu\nu} + \frac{1}{2} \partial^\nu h_{\mu\nu} \partial_\sigma h^{\mu\sigma} \\ &\quad - \frac{1}{2} \partial^\nu h_{\mu\nu} \partial^\mu h_\sigma{}^\sigma + \frac{1}{4} \partial^\mu h_\nu{}^\nu \partial_\mu h_\sigma{}^\sigma + O(h^3) \end{aligned} \quad (32)$$

again up to total derivatives. This last expression is in fact the same as Eq. (7). The correct relationship between the original graviton field $h_{\mu\nu}$ and the metric field $g_{\mu\nu}$ is

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + \kappa h_{\mu\nu}(x) \quad (33)$$

If, as is often customary, one rescales $h_{\mu\nu}$ in such a way that the κ factor does not appear on the r.h.s., then both the g and h fields are dimensionless.

The weak field invariance properties of the gravitational action of Eq. (11) are replaced in the general theory by general coordinate transformations $x^\mu \rightarrow x'^\mu$, under which the metric transforms as a covariant second rank tensor

$$g'_{\mu\nu}(x') = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma}(x) \quad (34)$$

which leaves the infinitesimal proper time interval $d\tau$ with

$$d\tau^2 = -g_{\mu\nu} dx^\mu dx^\nu \quad (35)$$

invariant. In their infinitesimal form, coordinate transformations are written as

$$x'^\mu = x^\mu + \epsilon^\mu(x) \quad (36)$$

under which the metric at the same point x then transforms as

$$\delta g_{\mu\nu}(x) = -g_{\lambda\nu}(x) \partial_\mu \epsilon^\lambda(x) - g_{\lambda\mu}(x) \partial_\nu \epsilon^\lambda(x) - \epsilon^\lambda(x) \partial_\lambda g_{\mu\nu}(x) \quad (37)$$

and which is usually referred to as the Lie derivative of g . The latter generalizes the weak field gauge invariance property of Eq. (11) to all orders in $h_{\mu\nu}$.

For infinitesimal coordinate transformations, one can gain some additional physical insight by decomposing the derivative of the small coordinate change ϵ_μ in Eq. (36) as

$$\frac{\partial \epsilon_\mu}{\partial x^\nu} = s_{\mu\nu} + a_{\mu\nu} + t_{\mu\nu} \quad (38)$$

with

$$\begin{aligned} s_{\mu\nu} &= \frac{1}{d} \eta_{\mu\nu} \partial \cdot \epsilon \\ a_{\mu\nu} &= \frac{1}{2} (\partial_\mu \epsilon_\nu - \partial_\nu \epsilon_\mu) \\ t_{\mu\nu} &= \frac{1}{2} (\partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu - \frac{2}{d} \eta_{\mu\nu} \partial \cdot \epsilon) \end{aligned} \quad (39)$$

Then $s_{\mu\nu}(x)$ can be thought of describing local scale transformations, $a_{\mu\nu}(x)$ is written in terms of an antisymmetric tensor and therefore describes local rotations, while $t_{\mu\nu}(x)$ contains a traceless symmetric tensor and describes local shears.

Since both the scalar curvature $R(x)$ and the volume element $dx\sqrt{g(x)}$ are separately invariant under the general coordinate transformations of Eqs. (34) and (36), both of the following action contributions are acceptable

$$\begin{aligned} \int dx \sqrt{g(x)} \\ \int dx \sqrt{g(x)} R(x) \end{aligned} \quad (40)$$

the first being known as the cosmological constant contribution (as it represents the total space-time volume). In the weak field limit, the first, cosmological constant term involves

$$\sqrt{g} = 1 + \frac{1}{2} h_\mu{}^\mu + \frac{1}{8} h_\mu{}^\mu h_\nu{}^\nu - \frac{1}{4} h_{\mu\nu} h^{\mu\nu} + O(h^3) \quad (41)$$

which is easily obtained from the matrix formula

$$\sqrt{\det g} = \exp\left(\frac{1}{2} \text{tr} \ln g\right) = \exp\left[\frac{1}{2} \text{tr} \ln(\eta + h)\right] \quad (42)$$

after expanding out the exponential in powers of $h_{\mu\nu}$. We have also reverted here to the more traditional way of performing the weak field expansion (i.e. without factors of κ),

$$\begin{aligned} g_{\mu\nu} &= \eta_{\mu\nu} + h_{\mu\nu} \\ g^{\mu\nu} &= \eta^{\mu\nu} - h_{\mu\nu} + h_\mu{}^\alpha h_{\alpha\nu} + \dots \end{aligned} \quad (43)$$

with $\eta_{\mu\nu}$ the flat metric. The reason why such a \sqrt{g} cosmological constant term was not originally included in the construction of the Lagrangian of Eq. (7) is that it does not contain derivatives of the $h_{\mu\nu}$ field. It is in a sense analogous to a mass term, without giving rise to any breaking of local gauge invariance.

In the general theory, the energy-momentum tensor for matter $T_{\mu\nu}$ is most suitably defined in terms of the variation of the matter action I_{matter} ,

$$\delta I_{\text{matter}} = \frac{1}{2} \int dx \sqrt{g} \delta g_{\mu\nu} T^{\mu\nu} \quad (44)$$

and is conserved if the matter action is a scalar,

$$\nabla_{\mu} T^{\mu\nu} = 0 \quad (45)$$

Variation of the gravitational Einstein-Hilbert action of Eq. (21), with the matter part added, then leads to the field equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \quad (46)$$

Here we have also added a cosmological constant term, with a scaled cosmological constant $\lambda = \lambda_0/16\pi G$, which follows from adding to the gravitational action a term $\lambda_0 \int \sqrt{g}$ ¹.

One can exploit the freedom under general coordinate transformations $x'^{\mu} = f(x^{\mu})$ to impose a suitable coordinate condition, such as

$$\Gamma^{\lambda} \equiv g^{\mu\nu} \Gamma_{\mu\nu}^{\lambda} = 0 \quad (47)$$

which is seen to be equivalent to the following gauge condition on the metric

$$\partial_{\mu}(\sqrt{g} g^{\mu\nu}) = 0 \quad (48)$$

and therefore equivalent, in the weak field limit, to the harmonic gauge condition introduced previously in Eq. (12).

3. Feynman Rules

The Feynman rules represent the standard way to do perturbative calculations in quantum gravity. To this end one first expands again the action out in powers of the field $h_{\mu\nu}$ and separates

¹ The present experimental value for Newton's constant is $\hbar G/c^3 = (1.61624(12) \times 10^{-33} \text{ cm})^2$. Recent observational evidence [reviewed in (Damour,2007)] suggests a non-vanishing positive cosmological constant λ , corresponding to a vacuum density $\rho_{vac} \approx (2.3 \times 10^{-3} \text{ eV})^4$ with ρ_{vac} related to λ by $\lambda = 8\pi G \rho_{vac}/c^4$. As can be seen from the field equations, λ has the same dimensions as a curvature. One has from observation $\lambda \sim 1/(10^{28} \text{ cm})^2$, so this new curvature length scale is comparable to the size of the visible universe $\sim 4.4 \times 10^{28} \text{ cm}$.

out the quadratic part, which gives the graviton propagator, from the rest of the Lagrangian which gives the $O(h^3), O(h^4) \dots$ vertices. To define the graviton propagator one also requires the addition of a gauge fixing term and the associated Faddeev-Popov ghost contribution (Feynman, 1962; Faddeev and Popov, 1968). Since the diagrammatic calculations are performed using dimensional regularization, one first needs to define the theory in d dimensions; at the end of the calculations one will be interested in the limit $d \rightarrow 4$.

So first one expands around the d -dimensional flat Minkowski space-time metric, with signature given by $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1, \dots)$. The Einstein-Hilbert action in d dimensions is given by a generalization of Eq. (21)

$$I_E = \frac{1}{16\pi G} \int d^d x \sqrt{g(x)} R(x) \quad , \quad (49)$$

with again $g(x) = -\det(g_{\mu\nu})$ and R the scalar curvature; in the following it will be assumed, at least initially, that the bare cosmological constant λ_0 is zero. The simplest form of matter coupled in an invariant way to gravity is a set of spinless scalar particles of mass m , with action

$$I_m = \frac{1}{2} \int d^d x \sqrt{g(x)} \left[-g^{\mu\nu}(x) \partial_\mu \phi(x) \partial_\nu \phi(x) - m^2 \phi^2(x) \right] \quad . \quad (50)$$

In Feynman diagram perturbation theory the metric $g_{\mu\nu}(x)$ is expanded around the flat metric $\eta_{\mu\nu}$, by writing again

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + \sqrt{16\pi G} h_{\mu\nu}(x) \quad . \quad (51)$$

The quadratic part of the Lagrangian [see Eq. (7)] is then

$$\mathcal{L} = -\frac{1}{4} \partial_\mu h_{\alpha\beta} \partial^\mu h^{\alpha\beta} + \frac{1}{8} (\partial_\mu h^\alpha{}_\alpha)^2 + \frac{1}{2} C_\mu^2 + \frac{1}{2} \kappa h_{\mu\nu} T^{\mu\nu} + \mathcal{L}_{\text{gf}} + \dots \quad (52)$$

where the dots indicate terms that are either total derivatives, or higher order in h . A suitable gauge fixing term C_μ is given by

$$C_\mu \equiv \partial_\alpha h^\alpha{}_\mu - \frac{1}{2} \partial_\mu h^\alpha{}_\alpha \quad (53)$$

Without such a term the quadratic part of the gravitational Lagrangian of Eq. (7) would contain a zero mode $h_{\mu\nu} \sim \partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu$, due to the gauge invariance of Eq. (11), which would make the graviton propagator ill defined.

The gauge fixing contribution \mathcal{L}_{gf} itself will be written as the sum of two terms,

$$\mathcal{L}_{\text{gf}} = -\frac{1}{2} C_\mu^2 + \mathcal{L}_{\text{ghost}} \quad (54)$$

with the first term engineered so as to conveniently cancel the $+\frac{1}{2}C_\mu^2$ in Eq. (52) and thus give a well defined graviton propagator. Note incidentally that this gauge is *not* the harmonic gauge condition of Eq. (12), and is usually referred to instead as the DeDonder gauge. The second term is determined as usual from the variation of the gauge condition under an infinitesimal gauge transformation of the type in Eq. (11)

$$\delta C_\mu = \partial^2 \epsilon_\mu + O(\epsilon^2) \quad (55)$$

which leads to the lowest order ghost Lagrangian

$$\mathcal{L}_{\text{ghost}} = -\partial_\mu \bar{\eta}_\alpha \partial^\mu \eta^\alpha + O(\hbar^2) \quad (56)$$

where η_α is the spin-one anticommuting ghost field, with propagator

$$D_{\mu\nu}^{(\eta)}(k) = \frac{\eta_{\mu\nu}}{k^2} \quad (57)$$

In this gauge the graviton propagator is finally determined from the surviving quadratic part of the pure gravity Lagrangian, which is

$$\mathcal{L}_0 = -\frac{1}{4} \partial_\mu h_{\alpha\beta} \partial^\mu h^{\alpha\beta} + \frac{1}{8} (\partial_\mu h^\alpha{}_\alpha)^2 \quad (58)$$

The latter can be conveniently re-written in terms of a matrix V

$$\mathcal{L}_0 = -\frac{1}{2} \partial_\lambda h_{\alpha\beta} V^{\alpha\beta\mu\nu} \partial^\lambda h_{\mu\nu} \quad (59)$$

with

$$V_{\alpha\beta\mu\nu} = \frac{1}{2} \eta_{\alpha\mu} \eta_{\beta\nu} - \frac{1}{4} \eta_{\alpha\beta} \eta_{\mu\nu} \quad (60)$$

The matrix V can easily be inverted, for example by re-labelling rows and columns via the correspondence

$$11 \rightarrow 1, \quad 22 \rightarrow 2, \quad 33 \rightarrow 3, \quad \dots \quad 12 \rightarrow 5, \quad 13 \rightarrow 6, \quad 14 \rightarrow 7 \quad \dots \quad (61)$$

and the graviton Feynman propagator in d dimensions is then found to be of the form

$$D_{\mu\nu\alpha\beta}(k) = \frac{\eta_{\mu\alpha} \eta_{\nu\beta} + \eta_{\mu\beta} \eta_{\nu\alpha} - \frac{2}{d-2} \eta_{\mu\nu} \eta_{\alpha\beta}}{k^2} \quad (62)$$

with a suitable $i\epsilon$ prescription to correctly integrate around poles in the complex k space. Equivalently the whole procedure could have been performed from the start with an Euclidean metric $\eta_{\mu\nu} \rightarrow \delta_{\mu\nu}$ and a complex time coordinate $t = -i\tau$ with hardly any changes of substance. The

simple pole in the graviton propagator at $d = 2$ serves as a reminder of the fact that, due to the Gauss-Bonnet identity, the gravitational Einstein-Hilbert action of Eq. (49) becomes a topological invariant in two dimensions.

Higher order correction in h to the Lagrangian for pure gravity then determine to order h^3 the three-graviton vertex, to order h^4 the four-graviton vertex, and so on. Because of the \sqrt{g} and $g^{\mu\nu}$ terms in the action, there are an infinite number of vertices in h .

Had one included a cosmological constant term as in Eq. (41), which can also be expressed in terms of the matrix V as

$$\sqrt{g} = 1 + \frac{1}{2} h_{\mu\mu} - \frac{1}{2} h_{\alpha\beta} V^{\alpha\beta\mu\nu} h_{\mu\nu} + O(h^3), \quad (63)$$

then the the expression in Eq. (59) would have read

$$\mathcal{L}_0 = \lambda_0(1 + \kappa \frac{1}{2} h^\alpha{}_\alpha) + \frac{1}{2} h_{\alpha\beta} V^{\alpha\beta\mu\nu} (\partial^2 + \lambda_0 \kappa^2) h_{\mu\nu} \quad (64)$$

with $\kappa^2 = 16\pi G$. Then the graviton propagator would have been remained the same, except for the replacement $k^2 \rightarrow k^2 + \lambda_0 \kappa^2$. In this gauge it would correspond to the exchange of a particle of mass $\mu^2 = \lambda_0 \kappa^2$. The term linear in h can be interpreted as a uniform constant source for the gravitational field. But one needs to be quite careful, since for non-vanishing cosmological constant flat space $g_{\mu\nu} \sim \eta_{\mu\nu}$ is no longer a solution of the vacuum field equations and the problem becomes a bit more subtle: one needs to expand around the correct vacuum solutions in the presence of a λ -term, which are no longer constant.

Another point needs to be made here. One peculiar aspect of perturbative gravity is that there is no unique way of doing the weak field expansions, and one can have therefore different sets of Feynman rules, even apart from the choice of gauge condition, depending on how one chooses to do the expansion for the metric.

For example, the structure of the scalar field action of Eq. (50) suggests to define instead the small fluctuation graviton field $h_{\mu\nu}(x)$ via

$$\tilde{g}^{\mu\nu}(x) \equiv g^{\mu\nu}(x) \sqrt{g(x)} = \eta^{\mu\nu} + K h^{\mu\nu}(x) . \quad (65)$$

with $K^2 = 32\pi G$ (Faddeev and Popov, 1973; Capper et al, 1973). Here it is $h^{\mu\nu}(x)$ that should be referred to as "the graviton field". The change of variables from the $g_{\mu\nu}$'s to the $g^{\mu\nu}(x) \sqrt{g(x)}$'s involves a Jacobian, which can be taken to be one in dimensional regularization. There is one obvious advantage of this expansion over the previous one, namely that it leads to considerably simpler Feynman rules, both for the graviton vertices and for the scalar-graviton vertices, which

can be advantageous when computing one-loop scattering amplitudes of scalar particles (Hamber and Liu, 1985). Even the original gravitational action has a simpler form in terms of the variables of Eq. (65) as shown originally in (Goldberg, 1958).

Again, when performing Feynman diagram perturbation theory a gauge fixing term needs to be added in order to define the propagator, for example of the form

$$\frac{1}{K^2} (\partial_\mu \sqrt{g} g^{\mu\nu})^2 , \quad (66)$$

In this new framework the bare graviton propagator is given simply by

$$D_{\mu\nu\alpha\beta}(k) = \frac{\eta_{\mu\alpha}\eta_{\nu\beta} + \eta_{\mu\beta}\eta_{\nu\alpha} - \eta_{\mu\nu}\eta_{\alpha\beta}}{2k^2} \quad (67)$$

which should be compared to Eq. (62) (the extra factor of one half here is just due to the convention in the choice of K). One notices that now there are no factors of $1/(d-2)$ for the graviton propagator in d dimensions. But such factors appear instead in the expression for the Feynman rules for the graviton vertices, and such $(d-2)^{-1}$ pole terms appear therefore regardless of the choice of expansion field. For the three-graviton and two ghost-graviton vertex the relevant expressions are quite complicated. The three-graviton vertex is given by

$$\begin{aligned} U(q_1, q_2, q_3)_{\alpha_1\beta_1, \alpha_2\beta_2, \alpha_3\beta_3} = & \\ -\frac{K}{2} & \left[q_{(\alpha_1}^2 q_{\beta_1)}^3 \left(2\eta_{\alpha_2(\alpha_3}\eta_{\beta_3)\beta_2} - \frac{2}{d-2}\eta_{\alpha_2\beta_2}\eta_{\alpha_3\beta_3} \right) \right. \\ & + q_{(\alpha_2}^1 q_{\beta_2)}^3 \left(2\eta_{\alpha_1(\alpha_3}\eta_{\beta_3)\beta_1} - \frac{2}{d-2}\eta_{\alpha_1\beta_1}\eta_{\alpha_3\beta_3} \right) \\ & + q_{(\alpha_3}^1 q_{\beta_3)}^2 \left(2\eta_{\alpha_1(\alpha_2}\eta_{\beta_2)\beta_1} - \frac{2}{d-2}\eta_{\alpha_1\beta_1}\eta_{\alpha_2\beta_2} \right) \\ & + 2q_{(\alpha_2}^3 \eta_{\beta_2)(\alpha_1}\eta_{\beta_1)(\alpha_3} q_{\beta_3)}^2 + 2q_{(\alpha_3}^1 \eta_{\beta_3)(\alpha_2}\eta_{\beta_2)(\alpha_1} q_{\beta_1)}^3 + 2q_{(\alpha_1}^2 \eta_{\beta_1)(\alpha_3}\eta_{\beta_3)(\alpha_2} q_{\beta_2)}^1 \\ & + q^2 \cdot q^3 \left(\frac{2}{d-2}\eta_{\alpha_1(\alpha_2}\eta_{\beta_2)\beta_1}\eta_{\alpha_3\beta_3} + \frac{2}{d-2}\eta_{\alpha_1(\alpha_3}\eta_{\beta_3)\beta_1}\eta_{\alpha_2\beta_2} - 2\eta_{\alpha_1(\alpha_2}\eta_{\beta_2)(\alpha_3}\eta_{\beta_3)\beta_1} \right) \\ & + q^1 \cdot q^3 \left(\frac{2}{d-2}\eta_{\alpha_2(\alpha_1}\eta_{\beta_1)\beta_2}\eta_{\alpha_3\beta_3} + \frac{2}{d-2}\eta_{\alpha_2(\alpha_3}\eta_{\beta_3)\beta_2}\eta_{\alpha_1\beta_1} - 2\eta_{\alpha_2(\alpha_1}\eta_{\beta_1)(\alpha_3}\eta_{\beta_3)\beta_2} \right) \\ & \left. + q^1 \cdot q^2 \left(\frac{2}{d-2}\eta_{\alpha_3(\alpha_1}\eta_{\beta_1)\beta_3}\eta_{\alpha_2\beta_2} + \frac{2}{d-2}\eta_{\alpha_3(\alpha_2}\eta_{\beta_2)\beta_3}\eta_{\alpha_1\beta_1} - 2\eta_{\alpha_3(\alpha_1}\eta_{\beta_1)(\alpha_2}\eta_{\beta_2)\beta_3} \right) \right] . \end{aligned} \quad (68)$$

The ghost-graviton vertex is given by

$$V(k_1, k_2, k_3)_{\alpha\beta, l\mu} = K \left[-\eta_{l(\alpha} k_{1\beta)} k_{2\mu} + \eta_{l\mu} k_{2(\alpha} k_{3\beta)} \right] , \quad (69)$$

and the two scalar-one graviton vertex is given by

$$\frac{K}{2} \left(p_{1\mu} p_{2\nu} + p_{1\nu} p_{2\mu} - \frac{2}{d-2} m^2 \eta_{\mu\nu} \right) , \quad (70)$$

where the p_1, p_2 denote the four-momenta of the incoming and outgoing scalar field, respectively.

Finally the two scalar-two graviton vertex is given by

$$\frac{K^2 m^2}{2(d-2)} \left(\eta_{\mu\lambda} \eta_{\nu\sigma} + \eta_{\mu\sigma} \eta_{\nu\lambda} - \frac{2}{d-2} \eta_{\mu\nu} \eta_{\lambda\sigma} \right), \quad (71)$$

where one pair of indices (μ, ν) is associated with one graviton line, and the other pair (λ, σ) is associated with the other graviton line. These rules follow readily from the expansion of the gravitational action to order $G^{3/2}$ (K^3), and of the scalar field action to order G (K^2), as shown in detail in (Capper et al, 1973). Note that the poles in $1/(d-2)$ have disappeared from the propagator, but have moved to the vertex functions. As mentioned before, they reflect the kinematic singularities that arise in the theory as $d \rightarrow 2$ due to the Gauss-Bonnet identity. As an illustration, Fig. 2 shows the lowest order diagrams contributing to the static potential between two massive spinless sources (Hamber and Liu, 1995).

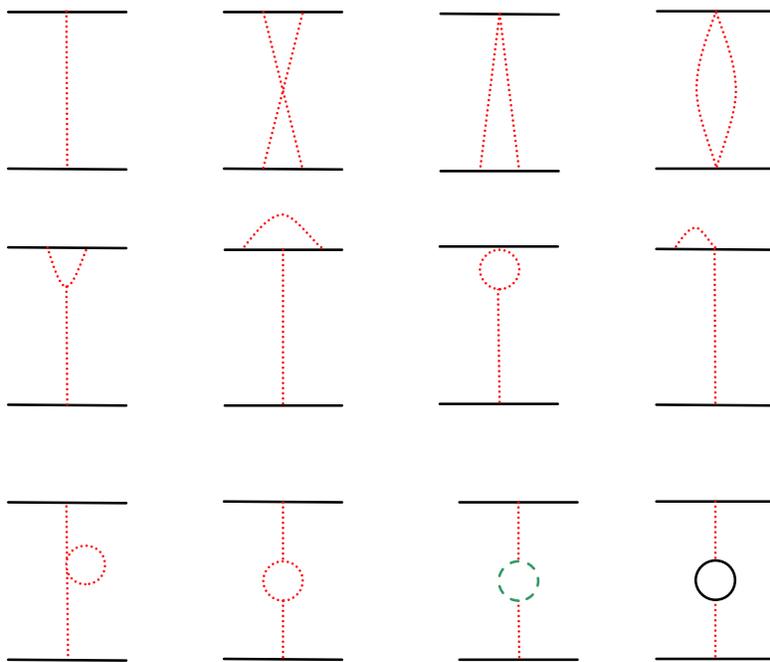


FIG. 2 Lowest order diagrams illustrating modifications to the classical gravitational potential due to graviton exchange. Continuous lines denote a spinless heavy matter particle, short dashed lines a graviton and the long dashed line the ghost loop. The last diagram shows the scalar matter loop contribution.

4. One-Loop Divergences

Once the propagators and vertices have been defined, one can then proceed as in QED and Yang-Mills theories and evaluate the quantum mechanical one loop corrections. In a renormalizable theory with a dimensionless coupling, such as QED and Yang-Mills theories, one has that the radiative corrections lead to charge, mass and field re-definitions. In particular, for the pure $SU(N)$ gauge action one finds

$$I_{YM} = -\frac{1}{4g^2 N} \int dx \operatorname{tr} F_{\mu\nu}^2 \rightarrow -\frac{1}{4g_R^2 N} \int dx \operatorname{tr} F_R^2{}_{\mu\nu} \quad (72)$$

so that the form of the action is preserved by the renormalization procedure: no new interaction terms such as $(D_\mu F^{\mu\nu})^2$ need to be introduced in order to re-absorb the divergences.

In gravity the coupling is dimensionful, $G \sim \mu^{2-d}$, and one expects trouble already on purely dimensional grounds, with divergent one loop corrections proportional to $G\Lambda^{d-2}$ where Λ is an ultraviolet cutoff². Equivalently, one expects to lowest order bad ultraviolet behavior for the running Newton's constant at large momenta,

$$\frac{G(k^2)}{G} \sim 1 + \operatorname{const.} G k^{d-2} + O(G^2) \quad (73)$$

These considerations also suggest that perhaps ordinary Einstein gravity is perturbatively renormalizable in the traditional sense in two dimensions, an issue to which we will return later in Sect. II.C.4.

A more general argument goes as follows. The gravitational action contains the scalar curvature R which involves two derivatives of the metric. Thus the graviton propagator in momentum space will go like $1/k^2$, and the vertex functions like k^2 . In d dimensions each loop integral will involve a momentum integration $d^d k$, so that the superficial degree of divergence \mathcal{D} of a Feynman diagram with V vertices, I internal lines and L loops will be given by

$$\mathcal{D} = dL + 2V - 2I \quad (74)$$

The topological relation involving V , I and L

$$L = 1 + I - V \quad (75)$$

² Indeed it was noticed very early on in the development of renormalization theory that perturbatively non-renormalizable theories would involve couplings with negative mass dimensions, and for which reaction rates would grow rapidly with energy (Sakata, Umezawa and Kamefuchi, 1952). It had originally been suggested by Heisenberg (Heisenberg, 1938) that the relevant mass scale appearing in such interactions with dimensionful coupling constants should be used to set an upper energy limit on the physical applicability of such theories.

is true for any diagram, and yields

$$\mathcal{D} = 2 + (d - 2) L \tag{76}$$

which is independent of the number of external lines. One concludes therefore that for $d > 2$ the degree of divergence increases with increasing loop order L .

The most convenient tool to determine the structure of the divergent one-loop corrections to Einstein gravity is the background field method (DeWitt, 1967; 't Hooft and Veltman, 1974) combined with dimensional regularization, wherein ultraviolet divergences appear as poles in $\epsilon = d - 4$ ³. In non-Abelian gauge theories the background field method greatly simplifies the calculation of renormalization factors, while at the same time maintaining explicit gauge invariance.

The essence of the method is easy to describe: one replaces the original field appearing in the classical action by $A + Q$, where A is a classical background field and Q the quantum fluctuation. A suitable gauge condition is chosen (the background gauge), such that manifest gauge invariance is preserved for the background A field. After expanding out the action to quadratic order in the Q field, the functional integration over Q is performed, leading to an effective action for the background A field. From the structure of this effective action the renormalization of the couplings, as well as possible additional counterterms, can then be read off. In the case of gravity it is in fact sufficient to look at the structure of those terms appearing in the effective action which are quadratic in the background field A . A very readable introduction to the background field method as applied to gauge theories can be found in (Abbot, 1981).

Unfortunately perturbative calculations in gravity are rather cumbersome due to the large number of indices and contractions, so the rest of this section is only intended more as a general outline, with the scope of hopefully providing some of the flavor of the original calculations. The first step consists in the replacement

$$g_{\mu\nu} \rightarrow \bar{g}_{\mu\nu} = g_{\mu\nu} + h_{\mu\nu} \tag{77}$$

where now $g_{\mu\nu}(x)$ is the classical background field and $h_{\mu\nu}$ the quantum field, to be integrated over. To determine the structure of one loop divergences it will often be sufficient to consider at the very end just the case of a flat background metric, $g_{\mu\nu} = \eta_{\mu\nu}$, or a small deviation from it.

After a somewhat tedious calculation one finds for the bare action

$$\mathcal{L} = \sqrt{g} [c_0 + c_1 R] \tag{78}$$

³ The second reference uses a complex time (Euclidean) $x_0 = ict$ notation that differs from the one used here.

expanded out to quadratic order in h

$$\begin{aligned}
\mathcal{L} = & \sqrt{g} [c_0 \{1 + \frac{1}{2}h^\alpha_\alpha - \frac{1}{4}h^\alpha_\beta h^\beta_\alpha + \frac{1}{8}h^\alpha_\alpha h^\beta_\beta\} \\
& + c_1 \{R - \frac{1}{2}h^\alpha_\alpha R + h^\alpha_\beta R^\beta_\alpha - \frac{1}{8}R h^\alpha_\alpha h^\beta_\beta + \frac{1}{4}R h^\alpha_\beta h^\beta_\alpha \\
& - h^\nu_\beta h^\beta_\alpha R^\alpha_\nu + \frac{1}{2}h^\alpha_\alpha h^\nu_\beta R^\beta_\nu - \frac{1}{4}\nabla_\nu h^\alpha_\beta \nabla^\nu h^\beta_\alpha \\
& + \nabla_\nu h^\alpha_\alpha \nabla^\nu h^\beta_\beta - \frac{1}{2}\nabla_\beta h^\alpha_\alpha \nabla^\mu h^\beta_\mu + \frac{1}{2}\nabla^\alpha h^\nu_\beta \nabla_\nu h^\beta_\alpha\}]
\end{aligned} \tag{79}$$

up to total derivatives. Here ∇_μ denotes a covariant derivative with respect to the metric $g_{\mu\nu}$. For $g_{\mu\nu} = \eta_{\mu\nu}$ the above expression coincides with the weak field Lagrangian contained in Eqs. (7) and (52), with a cosmological constant term added, as given in Eq. (41).

To this expression one needs to add the gauge fixing and ghost contributions, as was done in Eq. (52). The background gauge fixing term used is

$$-\frac{1}{2}C_\mu^2 = -\frac{1}{2}\sqrt{g} (\nabla_\alpha h^\alpha_\mu - \frac{1}{2}\nabla_\mu h^\alpha_\alpha)(\nabla_\beta h^{\beta\mu} - \frac{1}{2}\nabla^\mu h^\beta_\beta) \tag{80}$$

with a corresponding ghost Lagrangian

$$\mathcal{L}_{\text{ghost}} = \sqrt{g} \bar{\eta}_\mu (\partial_\alpha \partial^\alpha \eta^\mu - R^\mu_\alpha \eta^\alpha) \tag{81}$$

The integration over the $h_{\mu\nu}$ field can then be performed with the aid of the standard Gaussian integral formula

$$\ln \int [dh_{\mu\nu}] \exp\{-\frac{1}{2}h \cdot M(g) \cdot h - N(g) \cdot h\} = \frac{1}{2}N(g) \cdot M^{-1}(g) \cdot N(g) - \frac{1}{2} \text{tr} \ln M(g) + \text{const.} \tag{82}$$

leading to an effective action for the $g_{\mu\nu}$ field. In practice one is only interested in the divergent part, which can be shown to be local. Specific details of the functional measure over metrics $[dg_{\mu\nu}]$ are not deemed to be essential at this stage, as in perturbation theory one is only doing Gaussian integrals, with $h_{\mu\nu}$ ranging from $-\infty$ to $+\infty$. In particular when using dimensional regularization one uses the formal rule

$$\int d^d k = (2\pi)^d \delta^{(d)}(0) = 0 \tag{83}$$

which leads to some technical simplifications but obscures the role of the measure.

In the flat background field case case $g_{\mu\nu} = \eta_{\mu\nu}$, the functional integration over the $h_{\mu\nu}$ fields would have been particularly simple, since then one would be using

$$h_{\mu\nu}(x) h_{\alpha\beta}(x') \rightarrow \langle h_{\mu\nu}(x) h_{\alpha\beta}(x') \rangle = G_{\mu\nu\alpha\beta}(x, x') , \tag{84}$$

with the graviton propagator $G(k)$ given in Eq. (62). In practice, one can use the expected generally covariant structure of the one-loop divergent part

$$\Delta\mathcal{L}_g \propto \sqrt{g} \left(\alpha R^2 + \beta R_{\mu\nu} R^{\mu\nu} \right) , \quad (85)$$

with α and β some real parameters, as well as its weak field form, obtained from

$$\begin{aligned} R^2 &= \partial^2 h^\mu{}_\mu \partial^2 h^\alpha{}_\alpha - 2\partial^2 h^\mu{}_\mu \partial_\alpha \partial_\beta h^{\alpha\beta} + \partial_\alpha \partial_\beta h^{\alpha\beta} \partial_\mu \partial_\nu h^{\mu\nu} \\ R_{\alpha\beta} R^{\alpha\beta} &= \frac{1}{4} (\partial^2 h^\mu{}_\mu \partial^2 h^\alpha{}_\alpha + \partial^2 h_{\mu\alpha} \partial^2 h^{\mu\alpha} - 2\partial^2 h^\mu{}_\mu \partial_\alpha \partial_\beta h^{\alpha\beta} \\ &\quad - 2\partial_\alpha \partial_\nu h^\nu{}_\mu \partial_\alpha \partial_\beta h^{\mu\beta} + 2\partial_\mu \partial_\nu h^{\mu\nu} \partial_\alpha \partial_\beta h^{\alpha\beta} \end{aligned} \quad (86)$$

[compare with Eq. (31)], combined with some suitable special choices for the background metric, such as $g_{\mu\nu}(x) = \eta_{\mu\nu} f(x)$, to further simplify the calculation. This eventually determines the required one-loop counterterm for pure gravity to be

$$\Delta\mathcal{L}_g = \frac{\sqrt{g}}{8\pi^2(d-4)} \left(\frac{1}{120} R^2 + \frac{7}{20} R_{\mu\nu} R^{\mu\nu} \right) \quad (87)$$

For the simpler case of classical gravity coupled invariantly to a single real quantum scalar field one finds

$$\Delta\mathcal{L}_g = \frac{\sqrt{g}}{8\pi^2(d-4)} \frac{1}{120} \left(\frac{1}{2} R^2 + R_{\mu\nu} R^{\mu\nu} \right) \quad (88)$$

The complete set of one-loop divergences, computed using the alternate method of the heat kernel expansion and zeta function regularization ⁴ close to four dimensions, can be found in the comprehensive review (Hawking, 1977) and further references therein. In any case one is led to conclude that pure quantum gravity in four dimensions is not perturbatively renormalizable: the one-loop divergent part contains local operators which were not present in the original Lagrangian. It would seem therefore that these operators would have to be added to the bare \mathcal{L} , so that a consistent perturbative renormalization program can be developed in four dimensions.

There are two interesting, and interrelated, aspects of the result of Eq. (87). The first one is that for pure gravity the divergent part vanishes when one imposes the tree-level equations of motion $R_{\mu\nu} = 0$: the one-loop divergence vanishes on-shell. The second interesting aspect is that

⁴ The zeta-function regularization (Hawking, 1977) involves studying the behavior of the function $\zeta(s) = \sum_{n=0}^{\infty} (\lambda_n)^{-s}$, where the λ_n 's are the eigenvalues of the second order differential operator M in question. The series will converge for $s > 2$, and can be used for an analytic continuation to $s = 0$, which then leads to the formal result $\log(\det M) = \log \prod_{n=0}^{\infty} \lambda_n = -\zeta'(0)$.

the specific structure of the one-loop divergence is such that its effect can actually be re-absorbed into a field redefinition,

$$g_{\mu\nu} \rightarrow g_{\mu\nu} + \delta g_{\mu\nu}$$

$$\delta g_{\mu\nu} \propto \frac{7}{20} R_{\mu\nu} - \frac{11}{60} R g_{\mu\nu} \quad (89)$$

which renders the one-loop amplitudes finite for pure gravity. Unfortunately this hoped-for mechanism does not seem to work to two loops, and no additional miraculous cancellations seem to occur there. At two loops one expects on general grounds terms of the type $\nabla^4 R$, $R\nabla^2 R$ and R^3 . It can be shown that the first class of terms reduce to total derivatives, and that the second class of terms can also be made to vanish on shell by using the Bianchi identity. Out of the last set of terms, the R^3 ones, one can show ('t Hooft, 2002) that there are potentially 20 distinct contributions, of which 19 vanish on shell (i.e. by using the tree level field equations $R_{\mu\nu} = 0$). An explicit calculation then shows that a new non-removable on-shell R^3 -type divergence arises in pure gravity at two loops (Goroff and Sagnotti, 1985; van de Ven, 1992) from the only possible surviving non-vanishing counterterm, namely

$$\Delta\mathcal{L}^{(2)} = \frac{\sqrt{g}}{(16\pi^2)^2(4-d)} \frac{209}{2880} R_{\mu\nu}{}^{\rho\sigma} R_{\rho\sigma}{}^{\kappa\lambda} R_{\kappa\lambda}{}^{\mu\nu} . \quad (90)$$

To summarize, radiative corrections to pure Einstein gravity without a cosmological constant term induce one-loop R^2 -type divergences of the form

$$\Gamma_{div}^{(1)} = \frac{1}{d-4} \frac{\hbar}{16\pi^2} \int d^4x \sqrt{g} \left(\frac{7}{20} R_{\mu\nu} R^{\mu\nu} + \frac{1}{120} R^2 \right) , \quad (91)$$

and a two-loop non-removable on-shell R^3 -type divergence of the type

$$\Gamma_{div}^{(2)} = \frac{1}{d-4} \frac{209}{2880} \frac{\hbar^2 G}{(16\pi^2)^2} \int d^4x \sqrt{g} R_{\mu\nu}{}^{\rho\sigma} R_{\rho\sigma}{}^{\kappa\lambda} R_{\kappa\lambda}{}^{\mu\nu} . \quad (92)$$

which present an almost insurmountable obstacle to the traditional perturbative renormalization procedure in four dimensions⁵. One can therefore attempt to summarize the situation so far as follows:

- In principle perturbation theory in G provides a clear, covariant framework in which radiative corrections to gravity can be computed in a systematic loop expansion. The effects

⁵ At three loops a non-vanishing invariant exists as well (Deser Kay and Stelle, 1977)

$$\int d^4x \sqrt{g} R_{\mu\nu\alpha\beta} R^{\alpha\beta\rho\sigma} R_{\rho\sigma\kappa\lambda} R^{\kappa\lambda\mu\nu} \quad (93)$$

Again on-shell all other invariants can be shown to be proportional to this one.

of a possibly non-trivial gravitational measure do not show up at any order in the weak field expansion, and radiative corrections affecting the renormalization of the cosmological constant, proportional to $\delta^d(0)$, are set to zero in dimensional regularization.

- At the same time at every order in the loop expansion new invariant terms involving higher derivatives of the metric are generated, whose effects cannot be simply re-absorbed into a re-definition of the original couplings. As expected on the basis of power-counting arguments, the theory is not perturbatively renormalizable in the traditional sense in four dimensions (although it seems to fail this test by a small measure in lowest order perturbation theory).
- The standard approach based on a perturbative expansion of the pure Einstein theory in four dimensions is therefore not convergent (it is in fact badly divergent), and represents therefore a temporary dead end.

5. Higher Derivative Terms

In the previous section it was shown that quantum corrections to the Einstein theory generate in perturbation theory R^2 -type terms in four dimensions. It seems therefore that, for the consistency of the perturbative renormalization group approach in four dimension, these terms would have to be included from the start, at the level of the bare microscopic action. Thus the main motivation for studying gravity with higher derivative terms is that it might cure the problem of ordinary quantum gravity, namely its perturbative non-renormalizability in four dimensions. This is indeed the case, in fact one can prove that higher derivative gravity (to be defined below) is perturbatively renormalizable to all orders in four dimensions.

At the same time new issues arise, which will be detailed below. The first set of problems has to do with the fact that, quite generally higher derivative theories with terms of the type $\phi \partial^4 \phi$ suffer from potential unitarity problems, which can lead to physically unacceptable negative probabilities. But since these are genuinely dynamical issues, it will be difficult to answer them satisfactorily in perturbation theory. In non-Abelian gauge theories one can use higher derivative terms, instead of the more traditional dimensional continuation, to regulate ultraviolet divergences (Slavnov, 1973), and higher derivative terms have been used successfully for some time in lattice regulated field theories (Symanzik, 1983). In these approaches the coefficient of the higher derivative terms is taken to zero at the end. The second set of issues is connected with the fact that the theory is asymptotically free in the higher derivative couplings, implying an infrared growth which renders

the perturbative estimates unreliable at low energies, in the regime of perhaps greatest physical interest. Note that higher derivative terms arise in string theory as well (Förger, Ovrut, Theisen and Waldram, 1996).

Let us first discuss the general formulation. In four dimensions possible terms quadratic in the curvature are

$$\begin{aligned}
& \int d^4x \sqrt{g} R^2 \\
& \int d^4x \sqrt{g} R_{\mu\nu} R^{\mu\nu} \\
& \int d^4x \sqrt{g} R_{\mu\nu\lambda\sigma} R^{\mu\nu\lambda\sigma} \\
& \int d^4x \sqrt{g} C_{\mu\nu\lambda\sigma} C^{\mu\nu\lambda\sigma} \\
& \int d^4x \sqrt{g} \epsilon^{\mu\nu\kappa\lambda} \epsilon^{\rho\sigma\omega\tau} R_{\mu\nu\rho\sigma} R^{\kappa\lambda\omega\tau} = 128\pi^2 \chi \\
& \int d^4x \sqrt{g} \epsilon^{\rho\sigma\kappa\lambda} R_{\mu\nu\rho\sigma} R^{\mu\nu}_{\kappa\lambda} = 96\pi^2 \tau
\end{aligned} \tag{94}$$

where χ is the Euler characteristic and τ the Hirzebruch signature. It will be shown below that these quantities are not all independent. The Weyl conformal tensor is defined in d dimensions as

$$\begin{aligned}
C_{\mu\nu\lambda\sigma} &= R_{\mu\nu\lambda\sigma} - \frac{2}{d-2} (g_{\mu[\lambda} R_{\sigma]\nu} - g_{\nu[\lambda} R_{\sigma]\mu}) \\
&\quad + \frac{2}{(d-1)(d-2)} R g_{\mu[\lambda} g_{\sigma]\nu}
\end{aligned} \tag{95}$$

where square brackets denote antisymmetrization. It is called conformal because it can be shown to be invariant under conformal transformations of the metric, $g_{\mu\nu}(x) \rightarrow \Omega^2(x) g_{\mu\nu}(x)$. In four dimensions one has

$$C_{\mu\nu\lambda\sigma} = R_{\mu\nu\lambda\sigma} - R_{\lambda[\mu} g_{\nu]\sigma} - R_{\sigma[\mu} g_{\nu]\lambda} + \frac{1}{3} R g_{\lambda[\mu} g_{\nu]\sigma} \tag{96}$$

The Weyl tensor can be regarded as the traceless part of the Riemann curvature tensor,

$$g^{\lambda\sigma} C_{\lambda\mu\sigma\nu} = g^{\mu\nu} g^{\lambda\sigma} C_{\mu\lambda\nu\sigma} = 0 \ . \tag{97}$$

and on-shell the Riemann tensor in fact coincides with the Weyl tensor. From the definition of the Weyl tensor one infers in four dimensions the following curvature-squared identity

$$R_{\mu\nu\lambda\sigma} R^{\mu\nu\lambda\sigma} = C_{\mu\nu\lambda\sigma} C^{\mu\nu\lambda\sigma} + 2 R_{\mu\nu} R^{\mu\nu} - \frac{1}{3} R^2 \tag{98}$$

Some of these results are specific to four dimensions. For example, in three dimensions the Weyl tensor vanishes identically and one has

$$R_{\mu\nu\lambda\sigma} R^{\mu\nu\lambda\sigma} - 4 R_{\mu\nu} R^{\mu\nu} - 3 R^2 = 0 \quad C_{\mu\nu\lambda\sigma} C^{\mu\nu\lambda\sigma} = 0, \tag{99}$$

In four dimensions the expression for the Euler characteristic can be written equivalently as

$$\chi = \frac{1}{32\pi^2} \int d^4x \sqrt{g} \left[R_{\mu\nu\lambda\sigma} R^{\mu\nu\lambda\sigma} - 4R_{\mu\nu} R^{\mu\nu} + R^2 \right] \quad (100)$$

The last result is the four-dimensional analogue of the two-dimensional Gauss-Bonnet formula

$$\chi = \frac{1}{2\pi} \int d^2x \sqrt{g} R \quad (101)$$

where $\chi = 2(g - 1)$ and g is the genus of the surface (the number of handles). For a manifold of fixed topology one can therefore use in four dimensions

$$R_{\mu\nu\lambda\sigma} R^{\mu\nu\lambda\sigma} = 4 R_{\mu\nu} R^{\mu\nu} - R^2 + \text{const.} \quad (102)$$

and

$$C_{\mu\nu\lambda\sigma} C^{\mu\nu\lambda\sigma} = 2 (R_{\mu\nu} R^{\mu\nu} - \frac{1}{3} R^2) + \text{const.} \quad (103)$$

Thus only *two* curvature squared terms for the gravitational action are independent in four dimensions (Lanczos, 1938), which can be chosen, for example, to be R^2 and $R_{\mu\nu}^2$. Consequently the most general curvature squared action in four dimensions can be written as

$$I = \int d^4x \sqrt{g} \left[\lambda_0 + k R + a R_{\mu\nu} R^{\mu\nu} - \frac{1}{3} (b + a) R^2 \right] \quad (104)$$

with $k = 1/16\pi G$, and up to boundary terms. The case $b = 0$ corresponds, by virtue of Eq. (103), to the conformally invariant, pure Weyl-squared case. If $b < 0$ then around flat space one encounters a tachyon at tree level (Stelle, 1977). It will also be of some interest later that in the Euclidean case (signature + + + +) the full gravitational action of Eq. (104) is positive for $a > 0$, $b < 0$ and $\lambda_0 > -3/4b(16\pi G)^2$.

Curvature squared actions for classical gravity were originally considered in (Weyl, 1922) and (Pauli, 1956). In the sixties it was argued that the higher derivative action of Eq. (104) should be power counting renormalizable (Utiyama and DeWitt, 1961). Later it was proven to be renormalizable to all orders in perturbation theory (Stelle, 1977). Some special cases of higher derivative theories have been shown to be classically equivalent to scalar-tensor theories (Whitt 1984).

One way to investigate physical properties of higher derivative theories is again via the weak field expansion. In analyzing the particle content it is useful to introduce a set of spin projection operators (Arnowitt, Deser and Misner, 1958; van Nievenhuizen, 1973), quite analogous to what is used in describing transverse-traceless (TT) modes in classical gravity (Misner, Thorne and Wheeler, 1973). These projection operators then show explicitly the unique decomposition of the

continuum gravitational action for linearized gravity into spin two (transverse-traceless) and spin zero (conformal mode) parts. The spin-two projection operator $P^{(2)}$ is defined in k -space as

$$\begin{aligned}
P_{\mu\nu\alpha\beta}^{(2)} = & \frac{1}{3k^2} (k_\mu k_\nu \eta_{\alpha\beta} + k_\alpha k_\beta \eta_{\mu\nu}) \\
& - \frac{1}{2k^2} (k_\mu k_\alpha \eta_{\nu\beta} + k_\mu k_\beta \eta_{\nu\alpha} + k_\nu k_\alpha \eta_{\mu\beta} + k_\nu k_\beta \eta_{\mu\alpha}) \\
& + \frac{2}{3k^4} k_\mu k_\nu k_\alpha k_\beta + \frac{1}{2} (\eta_{\mu\alpha} \eta_{\nu\beta} + \eta_{\mu\beta} \eta_{\nu\alpha}) - \frac{1}{3} \eta_{\mu\nu} \eta_{\alpha\beta} \ , \tag{105}
\end{aligned}$$

the spin-one projection operator $P^{(1)}$ as

$$\begin{aligned}
P_{\mu\nu\alpha\beta}^{(1)} = & \frac{1}{2k^2} (k_\mu k_\alpha \eta_{\nu\beta} + k_\mu k_\beta \eta_{\nu\alpha} + k_\nu k_\alpha \eta_{\mu\beta} + k_\nu k_\beta \eta_{\mu\alpha}) \\
& - \frac{1}{k^4} k_\mu k_\nu k_\alpha k_\beta \tag{106}
\end{aligned}$$

and the spin-zero projection operator $P^{(0)}$ as

$$\begin{aligned}
P_{\mu\nu\alpha\beta}^{(0)} = & - \frac{1}{3k^2} (k_\mu k_\nu \eta_{\alpha\beta} + k_\alpha k_\beta \eta_{\mu\nu}) \\
& + \frac{1}{3} \eta_{\mu\nu} \eta_{\alpha\beta} + \frac{1}{3k^4} k_\mu k_\nu k_\alpha k_\beta \ . \tag{107}
\end{aligned}$$

It is easy to check that the sum of the three spin projection operators adds up to unity

$$P_{\mu\nu\alpha\beta}^{(2)} + P_{\mu\nu\alpha\beta}^{(1)} + P_{\mu\nu\alpha\beta}^{(0)} = \frac{1}{2} (\eta_{\mu\alpha} \eta_{\nu\beta} + \eta_{\mu\beta} \eta_{\nu\alpha}) \ . \tag{108}$$

These projection operators then allow a decomposition of the gravitational field $h_{\mu\nu}$ into three independent modes. The spin two or transverse-traceless part

$$h_{\mu\nu}^{TT} = P_{\mu}^{\alpha} P_{\nu}^{\beta} h_{\alpha\beta} - \frac{1}{3} P_{\mu\nu} P^{\alpha\beta} h_{\beta\alpha} \tag{109}$$

the spin one or longitudinal part

$$h_{\mu\nu}^L = h_{\mu\nu} - P_{\mu}^{\alpha} P_{\nu}^{\beta} h_{\alpha\beta} \tag{110}$$

and the spin zero or trace part

$$h_{\mu\nu}^T = \frac{1}{3} P_{\mu\nu} P^{\alpha\beta} h_{\alpha\beta} \tag{111}$$

are such that their sum gives the original field h

$$h = h^{TT} + h^L + h^T \ , \tag{112}$$

with the quantity $P_{\mu\nu}$ defined as

$$P_{\mu\nu} = \eta_{\mu\nu} - \frac{1}{\partial^2} \partial_\mu \partial_\nu \tag{113}$$

or, equivalently, in k -space $P_{\mu\nu} = \eta_{\mu\nu} - k_\mu k_\nu / k^2$.

One can learn a number of useful aspects of the theory by looking at the linearized form of the equations of motion. As before, the linearized form of the action is obtained by setting $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ and expanding in h . Besides the expressions given in Eq. (31), one needs

$$\begin{aligned}\sqrt{g} R^2 &= (\partial^2 h^\lambda{}_\lambda - \partial_\lambda \partial_\kappa h^{\lambda\kappa})^2 + O(h^3) \\ \sqrt{g} R_{\lambda\mu\nu\kappa} R^{\lambda\mu\nu\kappa} &= \frac{1}{4} (\partial_\mu \partial_\kappa h_{\nu\lambda} + \partial_\lambda \partial_\nu h_{\mu\kappa} - \partial_\lambda \partial_\kappa h_{\mu\nu} - \partial_\mu \partial_\nu h_{\kappa\lambda})^2 + O(h^3),\end{aligned}\quad (114)$$

from which one can then obtain, for example from Eq. (102), an expression for $\sqrt{g} (R_{\mu\nu})^2$,

$$\begin{aligned}\sqrt{g} R_{\alpha\beta} R^{\alpha\beta} &= \frac{1}{4} (\partial^2 h^\mu{}_\mu \partial^2 h^\alpha{}_\alpha + \partial^2 h_{\mu\alpha} \partial^2 h^{\mu\alpha} - 2\partial^2 h^\mu{}_\mu \partial_\alpha \partial_\beta h^{\alpha\beta} \\ &\quad - 2\partial_\alpha \partial_\nu h^\nu{}_\mu \partial_\alpha \partial_\beta h^{\mu\beta} + 2\partial_\mu \partial_\nu h^{\mu\nu} \partial_\alpha \partial_\beta h^{\alpha\beta}) + O(h^3)\end{aligned}\quad (115)$$

Using the three spin projection operators defined previously, the action for linearized gravity without a cosmological constant term, Eq. (7), can then be re-expressed as

$$I_{\text{lin}} = \frac{1}{4} k \int dx h^{\mu\nu} [P^{(2)} - 2P^{(0)}]_{\mu\nu\alpha\beta} \partial^2 h^{\alpha\beta}\quad (116)$$

Only the $P^{(2)}$ and $P^{(0)}$ projection operators for the spin-two and spin-zero modes, respectively, appear in the action for the linearized gravitational field; the spin-one gauge mode does not enter the linearized action. Note also that the spin-zero mode enters with the wrong sign (in the linearized action it appears as a ghost contribution), but to this order it can be removed by a suitable choice of gauge in which the trace mode is made to vanish, as can be seen, for example, from Eq. (13).

It is often stated that higher derivative theories suffer from unitarity problems. This is seen as follows. When the higher derivative terms are included, the corresponding linearized expression for the gravitational action becomes

$$\begin{aligned}I_{\text{lin}} &= \frac{1}{2} \int dx \left\{ h^{\mu\nu} \left[\frac{1}{2} k + \frac{1}{2} a (-\partial^2) \right] (-\partial^2) P_{\mu\nu\rho\sigma}^{(2)} h^{\rho\sigma} \right. \\ &\quad \left. + h^{\mu\nu} \left[-k - 2b (-\partial^2) \right] (-\partial^2) P_{\mu\nu\rho\sigma}^{(0)} h^{\rho\sigma} \right\}\end{aligned}\quad (117)$$

Then the potential problems with unitarity and ghosts at ultrahigh energies, say comparable to the Planck mass $q \sim 1/G$, can be seen by examining the graviton propagator (Salam and Strathdee, 1978). In momentum space the free graviton propagator for higher derivative gravity and $\lambda_0 = 0$ can be written as

$$k \langle h_{\mu\nu}(q) h_{\rho\sigma}(-q) \rangle = \frac{2P_{\mu\nu\rho\sigma}^{(2)}}{q^2 + \frac{a}{k} q^4} + \frac{P_{\mu\nu\rho\sigma}^{(0)}}{-q^2 - \frac{2b}{k} q^4} + \text{gauge terms}\quad (118)$$

The first two terms on the r.h.s. can be decomposed as

$$2P_{\mu\nu\rho\sigma}^{(2)} \left[\frac{1}{q^2} - \frac{1}{q^2 + \frac{k}{a}} \right] + P_{\mu\nu\rho\sigma}^{(0)} \left[-\frac{1}{q^2} + \frac{1}{q^2 + \frac{k}{2b}} \right] \quad (119)$$

One can see that, on the one hand, the higher derivative terms improve the ultraviolet behavior of the theory, since the propagator now falls off as $1/q^4$ for large q^2 . At the same time, the theory appears to contain a spin-two ghost of mass $m_2 = \mu/\sqrt{a}$ and a spin-zero particle of mass $m_0 = \mu/\sqrt{2b}$. Here we have set $\mu = 1/\sqrt{(16\pi G)}$, which is of the order of the Planck mass ($1/\sqrt{G/\hbar c} = 1.2209 \times 10^{19} \text{ GeV}/c^2$). For $b < 0$ one finds a tachyon pole, which seems, for the time being, to justify the original choice of $b > 0$ in Eq. (104).

Higher derivative gravity theories also lead to modifications to the standard Newtonian potential, even though such deviations only become visible at very short distances, comparable to the Planck length $l_P = \sqrt{\hbar G/c^3} = 1.61624 \times 10^{-33} \text{ cm}$. In some special cases they can be shown to be classically equivalent to scalar-tensor theories without higher derivative terms (Whitt, 1984). The presence of massive states in the tree level graviton propagator indicates short distance deviations from the static Newtonian potential of the form

$$h_{00} \sim \frac{1}{r} - \frac{4}{3} \frac{e^{-m_2 r}}{r} + \frac{1}{3} \frac{e^{-m_0 r}}{r} \quad (120)$$

Moreover in the extreme case corresponding to the absence of the Einstein term ($k = 0$) the potential is linear in r ; but in this limit the theory is strongly infrared divergent, and it is not at all clear whether weak coupling perturbation theory is of any relevance.

In the quantum theory perturbation theory is usually performed around flat space, which requires $\lambda_0 = 0$, or around some fixed classical background. One sets again $g_{\mu\nu} \rightarrow \bar{g}_{\mu\nu} = g_{\mu\nu} + h_{\mu\nu}$ and expands the higher derivative action in powers of $h_{\mu\nu}$. If λ_0 is nonzero, one has to expand around a solution of the classical equations of motion for higher derivative gravity with a λ -term (Barth and Christensen, 1983), and the solution will no longer be constant over space-time. The above expansion is consistent with the assumption that the two higher derivative couplings a and b are large, since in such a limit one is close to flat space. One-loop radiative corrections then show that the theory is asymptotically free in the higher derivative couplings a and b (Julve and Tonin, 1978; Fradkin and Tseytlin, 1981; Avramidy and Barvinsky, 1985).

The calculation of one-loop quantum fluctuation effects proceeds in a way that is similar to the pure Einstein gravity case. One first decomposes the metric field as a classical background part $g_{\mu\nu}(x)$ and a quantum fluctuation part $h_{\mu\nu}(x)$ as in Eq. (77), and then expands the classical action to quadratic order in $h_{\mu\nu}$, with gauge fixing and ghost contributions added, similar to those in

Eqs. (80) and (81), respectively. The first order variation of the action of Eq. (104) gives the field equations for higher derivative gravity in the absence of sources,

$$\begin{aligned}
\frac{\partial I}{\partial g^{\mu\nu}} &= \frac{1}{\kappa^2} \sqrt{g} (R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R) + \frac{1}{2} \lambda_0 \sqrt{g} g^{\mu\nu} \\
&+ a \sqrt{g} \left[\frac{2}{3} (1 + \omega) R (R^{\mu\nu} - \frac{1}{4} g^{\mu\nu} R) \right. \\
&+ \frac{1}{2} g^{\mu\nu} R_{\alpha\beta} R^{\alpha\beta} - 2 R^{\mu\alpha\nu\beta} R_{\alpha\beta} + \frac{1}{3} (1 - 2\omega) \nabla^\mu \nabla^\nu R \\
&\left. - \square R^{\mu\nu} + \frac{1}{6} (1 + 4\omega) g^{\mu\nu} \square R \right] = 0
\end{aligned} \tag{121}$$

where we have set for the ratio of the two higher derivative couplings $\omega = b/a$.

The second order variation is done similarly. It then allows the Gaussian integral over the quantum fields to be performed using the formula of Eq. (82). One then finds that the one-loop effective action, which depends on $g_{\mu\nu}$ only, can be expressed as

$$\Gamma = \frac{1}{2} \text{tr} \ln F_{mn} - \text{tr} \ln Q_{\alpha\beta} - \frac{1}{2} \text{tr} \ln c^{\alpha\beta} \tag{122}$$

with the quantities F_{nm} and $Q_{\alpha\beta}$ defined by

$$\begin{aligned}
F_{nm} &= \frac{\delta^2 I}{\delta g^m \delta g^n} + \frac{\delta \chi_\alpha}{\delta g^m} c^{\alpha\beta} \frac{\delta \chi_\beta}{\delta g^n} \\
Q_{\alpha\beta} &= \frac{\delta \chi_\alpha}{\delta g^m} \nabla_\beta^m
\end{aligned} \tag{123}$$

A shorthand notation is used here, where spacetime and internal indices are grouped together so that $g^m = g_{\mu\nu}(x)$. χ_α are a set of gauge conditions, $c^{\alpha\beta}$ is a nonsingular functional matrix fixing the gauge, and the ∇_α^i are the local generators of the group of general coordinate transformations, $\partial_\alpha^i f^\alpha = 2g_{\alpha(\mu} \nabla_{\nu)} f^\alpha(x)$.

Ultimately one is only interested in the divergent part of the effective one-loop action. The method of extracting the divergent part out of the determinant (or trace) expression in Eq. (122) is similar to what is done, for example, in QED to evaluate the contribution of the fermion vacuum polarization loop to the effective action. There, after integrating out the fermions, one obtains a functional determinant of the massless Dirac operator $\mathcal{D}(A)$ in an external A_μ field,

$$\text{tr} \ln \mathcal{D}(A) - \text{tr} \ln \not{\partial} = \frac{c}{\epsilon} \int d^4x A^\mu (\eta_{\mu\nu} \partial^2 - \partial_\mu \partial_\nu) A^\nu + \dots \tag{124}$$

with c a calculable numerical constant. The trace needs to be regulated, and one way of doing it is via the integral representation

$$\frac{1}{2} \text{tr} \ln \mathcal{D}^2(A) = -\frac{1}{2} \int_\eta^\infty \frac{dt}{t} \text{tr} \exp[-t \mathcal{D}^2(A)] \tag{125}$$

with η a cutoff that is sent to zero at the end of the calculation. For gauge theories a detailed discussion can be found for example in (Rothe, 1985), and references therein. In the gravity case further discussions and more results can be found in (De Witt 1965; 't Hooft and Veltman 1974; Gilkey 1975; Christensen and Duff 1979) and references therein.

In the end, by a calculation similar to the one done in the pure Einstein gravity case, one finds that the one-loop contribution to the effective action contains for $d \rightarrow 4$ a divergent term of the form

$$\Delta\mathcal{L} = \frac{\sqrt{g}}{16\pi^2(4-d)} \left\{ \beta_2 (R_{\mu\nu}^2 - \frac{1}{3}R^2) + \beta_3 R^2 + \beta_4 R + \beta_5 \right\} \quad (126)$$

with the coefficients for the divergent parts given by

$$\begin{aligned} \beta_2 &= \frac{133}{10} \\ \beta_3 &= \frac{10}{9} \omega^2 + \frac{5}{3} \omega + \frac{5}{36} \\ \beta_4 &= \frac{1}{a\kappa^2} \left(\frac{10}{3} \omega - \frac{13}{6} - \frac{1}{4\omega} \right) \\ \beta_5 &= \frac{1}{a^2\kappa^4} \left(\frac{5}{2} + \frac{1}{8\omega^2} \right) + \frac{\tilde{\lambda}}{a\kappa^4} \left(\frac{56}{3} + \frac{2}{3\omega} \right) \end{aligned} \quad (127)$$

Here $\omega \equiv b/a$ and $\tilde{\lambda}$ is the dimensionless combination of the cosmological and Newton's constant $\tilde{\lambda} \equiv \frac{1}{2}\lambda_0\kappa^4$ with $\kappa^2 = 16\pi G$. A divergence proportional to the topological invariant χ with coefficient β_1 has not been included, as it only adds a field-independent constant to the action for a manifold of fixed topology. Also $\delta^{(4)}(0)$ -type divergences possibly originating from a non-trivial functional measure over the $g_{\mu\nu}$'s have been set to zero.

The structure of the ultraviolet divergences (which for an explicit momentum cutoff Λ would have appeared as $1/\epsilon \leftrightarrow \ln \Lambda$) allow one to read off immediately the renormalization group β -functions for the various couplings. To this order, the renormalization group equations for the two higher derivative couplings a and b and the dimensionless ratio of cosmological and Newton's constant $\tilde{\lambda}$ are

$$\begin{aligned} \frac{\partial a}{\partial t} &= \beta_2 + \dots \\ \frac{\partial \omega}{\partial t} &= -\frac{1}{a} (3\beta_3 + \omega\beta_2) + \dots \\ \frac{\partial \tilde{\lambda}}{\partial t} &= \frac{1}{2}\kappa^4 \beta_5 + 2\tilde{\lambda}\kappa^2 \beta_4 + \dots \end{aligned} \quad (128)$$

with the dots indicating higher loop corrections. Here t is the logarithm of the relevant energy scale, $t = (4\pi)^{-2} \ln(\mu/\mu_0)$, with μ a momentum scale $q^2 \approx \mu^2$, and μ_0 some fixed reference scale. It

is argued furthermore by the quoted authors that only the quantities β_2 , β_3 and the combination $\kappa^4\beta_5 + 4\tilde{\lambda}\kappa^2\beta_4$ are gauge independent, the latter combination appearing in the renormalization group equation for $\tilde{\lambda}(t)$ (this is a point to which we shall return later, as it follows quite generally from the properties of the gravitational action, and therefore from the gravitational functional integral, under a field rescaling, see Sect. II.C.4).

The perturbative scale dependence of the couplings $a(\mu)$, $b(\mu)$ and $\tilde{\lambda}(\mu)$ follows from integrating the three differential equations in Eq. (128). The first renormalization group equation is easily integrated, and shows the existence of an ultraviolet fixed point at $a^{-1} = 0$; the one-loop result for the running coupling a is simply given by $a(t) = a(0) + \beta_2 t$, or

$$a^{-1}(\mu) \underset{\mu \rightarrow \infty}{\sim} \frac{16\pi^2}{\beta_2 \ln(\mu/\mu_0)} \quad (129)$$

with μ_0 a reference scale. It suggests that the effective higher derivative coupling $a(\mu)$ increases at short distances, but decreases in the infrared regime $\mu \rightarrow 0$. But one should keep in mind that the one loop results are reliable at best only at very short distances, or large energy scales, $t \rightarrow \infty$. At the same time these results seem physically reasonable, as one would expect curvature squared terms to play less of a role at larger distances, as in the classical theory.

The scale dependence of the other couplings is a bit more complicated. The equation for $\omega(t)$ exhibits two fixed points at $\omega_{uv} \approx -0.0229$ and $\omega_{ir} \approx -5.4671$; in either case this would correspond to a higher derivative action with a positive R^2 term. It would also give rise to rapid short distance oscillations in the static potential, as can be seen for example from Eq. (120) and the definition of $m_0 = \mu/\sqrt{2b}$. The equation for $\tilde{\lambda}(t)$ gives a solution to one-loop order $\tilde{\lambda}(t) \sim \text{const. } t^q$ with $q \approx 0.91$, suggesting that the effective gravitational constant, in units of the cosmological constant, decreases at large distances. The experimental value for Newton's constant $\hbar G/c^3 = (1.61624 \times 10^{-33} \text{ cm})^2$ and for the scaled cosmological constant $G\lambda_0 \sim 1/(10^{28} \text{ cm})^2$ is such that the observed dimensionless ratio between the two is very small, $G^2\lambda_0 \sim 10^{-120}$. In the present model it seems entirely unclear how such a small ratio could arise from perturbation theory alone.

At short distances the dimensionless coupling $\tilde{\lambda} \sim \lambda_0 G^2$ seems to increase rapidly, thus partially invalidating the conclusions of a weak field expansion around flat space, which are based generally on the assumption of small G and λ_0 . At the same time, the fact that the higher derivative coupling a grows more rapidly in the ultraviolet than the coupling $\tilde{\lambda}$ can be used retroactively at least as a partial justification for the flat space expansion, in which the cosmological and Einstein terms are treated perturbatively. Ultimately the resolution of such delicate and complex issues would presumably require the development of the perturbative expansion not around flat space, but more

appropriately around the de Sitter metric, for which $R = 2\lambda_0/\kappa^2$. Even then one would have to confront such genuinely non-perturbative issues, such as what happens to the spin-zero ghost mass, whether the ghost poles gets shifted away from the real axis by quantum effects, and what the true ground state of the theory looks like in the long distance, strong fluctuation regime not accessible by perturbation theory.

What is also a bit surprising is that higher derivative gravity, to one-loop order, does not exhibit a nontrivial ultraviolet point in G , even though such a fixed point is clearly present in the $2 + \epsilon$ expansion (to be discussed later) at the one- and two-loop order, as well as in the lattice regularized theory in four dimensions (also to be discussed later). But this could just reflect a limitation of the one-loop calculation; to properly estimate the uncertainties of the perturbative results in higher derivative gravity and their potential physical implications a two-loop calculation is needed, which hopefully will be performed in the near future.

To summarize, higher derivative gravity theories based on R^2 -type terms are perturbatively renormalizable, but exhibit some short-distance oddities in the tree-level spectrum, associated with either ghosts or tachyons. Their perturbative (weak field) treatment suggest that the higher derivative couplings are only relevant at short distances, comparable to the Planck length, but the general evolution of the couplings away from a regime where perturbation theory is reliable remains an open question, which perhaps will never be answered satisfactorily in perturbation theory, if non-Abelian gauge theories, which are also asymptotically free, are taken as a guide.

6. Supergravity

An alternative approach to the vexing problem of ultraviolet divergences in perturbative quantum gravity (and for that matter, in any field theory) is to build in some additional degree of symmetry, such that loop effects acquire reduced divergence properties, or even become finite. One such approach, based on the invariance under local supersymmetry transformation, adds to the Einstein gravity Lagrangian a spin-3/2 gravitino field, whose purpose is to exactly cancel the loop divergences in the Einstein contributions. The enhanced symmetry is built in to ensure that such a cancellation does not just occur at one loop order, but propagates to every order of the loop expansion. The intent of this section is more to provide the general flavor of such an approach, and illustrate supergravity theories by a few specific examples of suitable actions. The reader is then referred to the vast literature on the subject for further examples, as well as contemporary leading candidate theories.

In the simplest scenario, one adds to gravity a spin- $\frac{3}{2}$ fermion field with suitable symmetry properties. A generally covariant action describing the interaction of vierbein fields $e_\mu^a(x)$ (with the metric field given by $g_{\mu\nu} = e_\mu^a e_{a\nu}$) and Rarita-Schwinger spin- $\frac{3}{2}$ fields $\psi_\mu(x)$, subject to the Majorana constraints $\psi_\rho = C\bar{\psi}_\rho^T$, was originally given in (Ferrara, Freedman and van Nieuwenhuizen, 1976). In the second order formulation it contains three contributions

$$I = \int d^4x (\mathcal{L}_2 + \mathcal{L}_{3/2} + \mathcal{L}_4) \quad (130)$$

with the usual Einstein term

$$\mathcal{L}_2 = \frac{1}{4\kappa^2} \sqrt{g} R, \quad (131)$$

the gravitino contribution

$$\mathcal{L}_{3/2} = -\frac{1}{2} \epsilon^{\lambda\rho\mu\nu} \bar{\psi}_\lambda \gamma_5 \gamma_\mu D_\nu \psi_\rho, \quad (132)$$

and a quartic fermion self-interaction

$$\begin{aligned} \mathcal{L}_4 = & - \frac{1}{32\kappa^2 \sqrt{g}} (\epsilon^{\tau\alpha\beta\nu} \epsilon_\tau^{\gamma\delta\mu} + \epsilon^{\tau\alpha\mu\nu} \epsilon_\tau^{\gamma\delta\beta} - \epsilon^{\tau\beta\mu\nu} \epsilon_\tau^{\gamma\delta\alpha}) \\ & \times (\bar{\psi}_\alpha \gamma_\mu \psi_\beta) (\bar{\psi}_\gamma \gamma_\nu \psi_\delta) \end{aligned} \quad (133)$$

The covariant derivative defined as

$$D_\nu \psi_\rho = \partial_\nu \psi_\rho - \Gamma_{\nu\rho}^\sigma \psi_\sigma + \frac{1}{2} \omega_{\nu ab} \sigma^{ab} \psi_\rho \quad (134)$$

involves the standard affine connection $\Gamma_{\nu\rho}^\sigma$, as well as the vierbein connection

$$\begin{aligned} \omega_{\nu ab} = & \frac{1}{2} [e_a^\mu (\partial_\nu e_{b\mu} - \partial_\mu e_{b\nu}) + e_a^\rho e_b^\sigma (\partial_\sigma e_{c\rho}) e^c{}_\nu] \\ & - (a \leftrightarrow b) \end{aligned} \quad (135)$$

with Dirac spin matrices

$$\sigma_{ab} = \frac{1}{4} [\gamma_a, \gamma_b]. \quad (136)$$

One can show that the combined Lagrangian is invariant, up to terms of order $(\psi)^5$, under the simultaneous transformations

$$\begin{aligned} \delta e_\mu^a(x) &= i\kappa \bar{\epsilon}(x) \gamma^a \psi_\mu(x) \\ \delta g_{\mu\nu} &= i\kappa \bar{\epsilon}(x) [\gamma_\mu \psi_\nu(x) + \gamma_\nu \psi_\mu] \\ \delta \psi_\mu(x) &= \kappa^{-1} D_\mu \epsilon(x) + \frac{1}{4} i\kappa (2 \bar{\psi}_\mu \gamma_a \psi_b + \bar{\psi}_a \gamma_\mu \psi_b) \sigma^{ab} \epsilon(x) \end{aligned} \quad (137)$$

where $\epsilon(x)$ is an arbitrary Majorana spinor.

The action of Eq. (130) can be written equivalently in first order form (Deser and Zumino, 1976) as

$$I = \int d^4x \left(\frac{1}{4} \kappa^{-2} e R - \frac{1}{2} \epsilon^{\lambda\rho\mu\nu} \bar{\psi}_\lambda \gamma_5 \gamma_\mu D_\nu \psi_\rho \right) \quad (138)$$

with $e_{a\mu}$ the vierbein with $e_{a\mu} e_\nu^a = g_{\mu\nu}$, and

$$e = \det e_{a\mu}, \quad R = e_a^\mu e_b^\nu R_{\mu\nu}{}^{ab}. \quad (139)$$

The covariant derivative D_μ on ψ_ν is defined in terms of its spin- $\frac{1}{2}$ part only

$$D_\mu = \partial_\mu - \frac{1}{2} \omega_{\mu ab} \sigma^{ab}. \quad (140)$$

and is related to the curvature tensor via the commutator identity

$$[D_\mu, D_\nu] = -\frac{1}{2} R_{\mu\nu ab} \sigma^{ab} \quad (141)$$

The first order action in Eq. (138) is invariant under

$$\begin{aligned} \delta e_\mu^a(x) &= i\kappa \bar{\epsilon}(x) \gamma^a \psi_\mu(x) \\ \delta \psi_\mu(x) &= \kappa^{-1} D_\mu \epsilon(x) \\ \delta \omega_\mu{}^{ab} &= B_\mu{}^{ab} - \frac{1}{2} e_\mu{}^b B_c{}^{ab} + \frac{1}{2} e_\mu{}^a B_c{}^{bc} \end{aligned} \quad (142)$$

with the quantity B defined as

$$B_a^{\lambda\mu} = i \epsilon^{\lambda\mu\nu\rho} \bar{\epsilon} \gamma_5 \gamma_a D_\nu \psi_\rho \quad (143)$$

and $\epsilon(x)$ is an arbitrary local Majorana spinor. In the first order formulation the vierbeins $e_{a\mu}(x)$, the connections $\omega_\mu{}^{ab}(x)$ and the Majorana vector-spinors $\psi_\mu(x)$ are supposed to be varied independently.

In (D'Eath, 1984; 1994) the $\mathcal{N} = 1$ supergravity action is written as

$$I = \int d^4x (\mathcal{L}_2 + \mathcal{L}_{3/2}) \quad (144)$$

with

$$\mathcal{L}_2 = \frac{1}{8\kappa^2} \epsilon^{\mu\nu\rho\sigma} \epsilon_{abcd} e_\mu^a e_\nu^b R_{\rho\sigma}{}^{cd} \quad (145)$$

$$\mathcal{L}_{3/2} = -\frac{1}{2} \epsilon^{\mu\nu\rho\sigma} (\bar{\psi}_\mu e_\nu^a \bar{\sigma}_a D_\rho \psi_\sigma - D_\rho \bar{\psi}_\mu e_\nu^a \bar{\sigma}_a \psi_\sigma) \quad (146)$$

in terms of the Weyl spinor gravitino fields $\psi_{A\mu}$ and $\bar{\psi}_{A'\mu}$ and the vierbein field $e^a{}_\mu$. The ϵ 's are Levi-Civita symbols with curved (up) and flat (down) indices respectively. The quantities $\bar{\sigma}_a$ represent a curved space generalization of the Pauli matrices discussed in (Carroll et al, 1994).

The original motivation for the supergravity action of Eqs. (130) or (138) was that, just like ordinary source-free gravity is ultraviolet finite on-shell because of the identity relating the invariant $(R_{\mu\nu\rho\sigma})^2$ to $(R_{\mu\nu})^2$ and R^2 , identities among invariants constructed out of ψ_μ and the strong constraints of supersymmetry would ensure one-loop, and higher, renormalizability of supergravity. There are reasons to believe that the triviality results found originally in globally supersymmetric theories (Nicolai, 1984) will not carry over into theories with local supersymmetry.

It was shown originally in (Grisaru, van Nieuwenhuizen and Vermaseren, 1976) and (Grisaru, 1976) that the original supergravity theory is finite to at least two loops. But most likely it fails to be finite at three loops (Deser, Kay and Stelle, 1977). As a consequence, more complex theories were devised to avoid the three-loop catastrophe. A new formulation, $\mathcal{N} = 4$ extended supergravity based on an $SO(4)$ symmetry, was suggested in (Das 1977; Cremmer and Scherk, 1977; Nicolai and Townsend, 1981). This theory now contains vector, spinor and scalar particle in addition to the gravitino and the graviton. Specifically, the theory contains a vierbein field $e_{a\mu}$, four spin- $\frac{3}{2}$ Majorana fields ψ_μ^i , four spin- $\frac{1}{2}$ Majorana fields ξ^i , six vector fields A_μ^{ij} , a scalar field A and a pseudoscalar field B , all massless, for a grand total of 53 independent terms in the Lagrangian. Subsequently $\mathcal{N} = 8$ supergravity was proposed, based on the even larger group $SO(8)$ (Cremmer and Julia, 1978). The enlarged theory now contains one graviton, 8 gravitinos, 28 vector fields, 56 Majorana spin- $\frac{1}{2}$ fields and 70 scalar fields, all massless. In general, $SO(\mathcal{N})$ supergravity contains \mathcal{N} gravitinos, $\frac{1}{2}\mathcal{N}(\mathcal{N} - 1)$ gauge fields, as well as several spin- $\frac{1}{2}$ Majorana fermions and complex scalars. The $SO(\mathcal{N})$ symmetry here is one which rotates, for example, the \mathcal{N} gravitinos into each other. In (Christensen, Duff, Gibbons and Roček, 1980) it was shown that in general such theories are finite at one loop order for $\mathcal{N} > 4$. For $\mathcal{N} > 8$ these theories become less viable since one then has more than one graviton, which leads to paradoxes, as well as particles with spin $j > 2$.

It is beyond our scope here to go any more deeply in the issue of the origin of such intriguing ultraviolet cancellations. But, as perhaps the simplest and most elementary motivation, one can use the Nielsen-Hughes formula (Nielsen, 1980; Hughes, 1981) for the one-loop β -function contribution from a particle of spin s

$$\beta_0 = -(-1)^{2s} [(2s)^2 - \frac{1}{3}] \quad (147)$$

to verify, by virtue of the particle multiplicities given above, that for example for $\mathcal{N} = 4$ the lowest

order divergences cancel

$$\beta_0 = -\frac{47}{3} \cdot 1 + \frac{26}{3} \cdot 4 - \frac{11}{3} \cdot 6 + \frac{2}{3} \cdot 4 + \frac{1}{3} \cdot 1 = 0 \quad (148)$$

For $\mathcal{N} = 8$ one has a similar complete cancellation

$$\beta_0 = -\frac{47}{3} \cdot 1 + \frac{26}{3} \cdot 8 - \frac{11}{3} \cdot 28 + \frac{2}{3} \cdot 56 + \frac{1}{3} \cdot 35 = 0 \quad (149)$$

Still, the issue of perturbative ultraviolet finiteness of these theories remains largely an open question, in part due to the daunting complexity of higher loop calculations, even though one believes that the high level of symmetry should ensure the cancellation of ultraviolet divergences to a very high order (at least up to seven loops). Recently it was suggested, based on the correspondence between $\mathcal{N} = 8$ supergravity and $\mathcal{N} = 4$ super Yang-Mills theory and the cancellations which arise at one and higher loops, that supergravity theories might be finite to all orders in the loop expansion (Bern, Dixon and Roiban, 2006).

One undoubtedly very attractive feature of supergravity theories is that they lead naturally to a small, or even vanishing, renormalized cosmological constant λ_0 . Due to the high level of symmetry, quartic and quadratic divergences in this quantity are expected to cancel exactly between bosonic and fermionic contributions, leaving a finite or even zero result. The hope is that some of these desirable features will survive supersymmetry breaking, a mechanism eventually required in order to remove, or shift to a high mass, the so far unobserved supersymmetric partners of the standard model particles.

B. Feynman Path Integral Formulation

So far the discussion of quantum gravity has focused almost entirely on perturbative aspects, where the gravitational coupling G is assumed to be weak, and the weak field expansion based on $\bar{g}_{\mu\nu} = g_{\mu\nu} + h_{\mu\nu}$ can be performed with some degree of reliability. At every order in the loop expansion the problem then reduces to the systematic evaluation of an increasingly complex sequence of Gaussian integrals over the (small) quantum fluctuation $h_{\mu\nu}$.

But there are reasons to expect that non-perturbative effects play an important role in quantum gravity. Then an improved formulation of the quantum theory is required, which does not rely exclusively on the framework of a perturbative expansion. Indeed already classically a black hole solution can hardly be considered a small perturbation of flat space. Furthermore, the fluctuating metric field $g_{\mu\nu}$ is dimensionless and carries therefore no natural scale. For the simpler cases of

a scalar field and non-Abelian gauge theories a consistent non-perturbative formulation based on the Feynman path integral has been known for some time and is well developed. Combined with the lattice approach, it provides an effective and powerful tool for systematically investigating non-trivial strong coupling behavior, such as confinement and chiral symmetry breaking. These phenomena are known to be generally inaccessible in weak coupling perturbation theory. In addition, the Feynman path integral approach provides a manifestly covariant formulation of the quantum theory, without the need for an artificial $3 + 1$ split required by the more traditional canonical approach, and the ambiguities that may follow from it. In fact, as will be seen later, in its non-perturbative lattice formulation no gauge fixing is required.

In a nutshell, the Feynman path integral formulation for pure quantum gravitation can be expressed in the functional integral formula

$$Z = \int_{\text{geometries}} e^{\frac{i}{\hbar} I_{\text{geometry}}} , \quad (150)$$

(for an illustration see Fig. 3), just like the Feynman path integral for a non-relativistic quantum mechanical particle (Feynman, 1951; Feynman and Hibbs, 1962) expresses quantum-mechanical amplitudes in terms of sums over paths

$$A(i \rightarrow f) = \int_{\text{paths}} e^{\frac{i}{\hbar} I_{\text{path}}} . \quad (151)$$

What is the precise meaning of the expression in Eq. (150)? The remainder of this section will be devoted to discussing attempts at a proper definition of the gravitational path integral of Eq. (150). A modern rigorous discussion of path integrals in quantum mechanics and (Euclidean) quantum field theory can be found, for example, in (Albeverio and Hoegh-Krohn, 1976), (Glimm and Jaffe, 1981), and (Zinn-Justin, 2002).

1. Sum over Paths

Already for a non-relativistic particle the path integral needs to be defined quite carefully by discretizing the time and introducing a short distance cutoff. The standard procedure starts from the quantum-mechanical transition amplitude

$$A(q_i, t_i \rightarrow q_f, t_f) = \langle q_f | e^{-\frac{i}{\hbar} H(t_f - t_i)} | q_i \rangle \quad (152)$$

and subdivides the time interval into $n + 1$ segments of size ϵ with $t_f = (n + 1)\epsilon + t_i$. Using completeness of the coordinate basis $|q_j \rangle$ at all intermediate times, one obtains the textbook

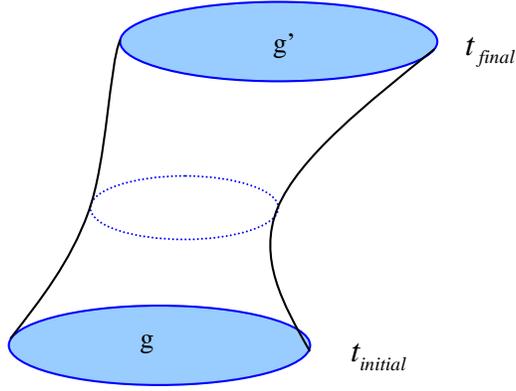


FIG. 3 Quantum mechanical amplitude of transitioning from an initial three-geometry described by g at time $t_{initial}$ to a final three-geometry described by g' at a later time t_{final} . The full amplitude is a sum over all intervening metrics connecting the two bounding three-surfaces, weighted by $\exp(iI/\hbar)$ where I is a suitably defined gravitational action.

result, here for a non-relativistic particle described by a Hamiltonian $H(p, q) = p^2/(2m) + V(q)$,

$$A(q_i, t_i \rightarrow q_f, t_f) = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \prod_{j=1}^n \frac{dq_j}{\sqrt{2\pi i \hbar \epsilon / m}} \times \exp \left\{ \frac{i}{\hbar} \sum_{j=1}^{n+1} \epsilon \left[\frac{1}{2} m \left(\frac{q_j - q_{j-1}}{\epsilon} \right)^2 - V \left(\frac{q_j + q_{j-1}}{2} \right) \right] \right\} \quad (153)$$

The expression in the exponent is recognized as a discretized form of the classical action. The above quantum-mechanical amplitude A is usually written in shorthand as

$$A(q_i, t_i \rightarrow q_f, t_f) = \int_{q_i(t_i)}^{q_f(t_f)} [dq] \exp \left\{ \frac{i}{\hbar} \int_{t_i}^{t_f} dt L(q, \dot{q}) \right\} \quad (154)$$

with $L = \frac{1}{2} m \dot{q}^2 - V(q)$ the Lagrangian for the particle. Thus what appears in the exponent is the classical action

$$I = \int_{t_i}^{t_f} dt L(q, \dot{q}) \quad (155)$$

associated with a given trajectory $q(t)$, connecting the initial coordinate $q_i(t_i)$ with the final one $q_f(t_f)$. Then $[dq]$ is the functional measure over paths $q(t)$, as spelled out explicitly in the precise lattice definition of Eq. (153). One advantage associated with having the classical action appear in the quantum mechanical amplitude is that all the symmetries of the theory are manifest in the Lagrangian form. The symmetries of the Lagrangian then have direct implications for the study

of quantum mechanical amplitudes. A stationary phase approximation to the path integral, valid in the limit $\hbar \rightarrow 0$, leads to the least action principle of classical mechanics

$$\delta I = 0 \tag{156}$$

In the above derivation it is not necessary to use a uniform lattice spacing ϵ ; one could have used as well a non-uniform spacing $\epsilon_i = t_i - t_{i-1}$ but the result would have been the same in the limit $n \rightarrow \infty$ (in analogy with the definition of the Riemann sum for ordinary integrals). Since quantum mechanical paths have a zig-zag nature and are nowhere differentiable, the mathematically correct definition should be taken from the finite sum in Eq. (153). In fact it can be shown that differentiable paths have zero measure in the Feynman path integral: already for the non-relativistic particle most of the contributions to the path integral come from paths that are far from smooth on all scales (Feynman and Hibbs, 1963), the so-called Wiener paths in turn related to Brownian motion. In particular, the derivative $\dot{q}(t)$ is not always defined, and the correct definition for the path integral is the one given in Eq. (153). A very complete and contemporary reference to the many applications of path integrals to non-relativistic quantum systems and statistical physics can be found in two recent books (Zinn-Justin, 2003; Kleinert, 2006).

As a next step, one can generalize the Feynman path integral construction to N particles with coordinates $q_i(t)$ ($i = 1, N$), and finally to the limiting case of continuous fields $\phi(x)$. If the field theory is defined from the start on a lattice, then the quantum fields are defined on suitable lattice points as ϕ_i .

2. Euclidean Rotation

In the case of quantum fields, one is generally interested in the vacuum-to-vacuum amplitude, which requires $t_i \rightarrow -\infty$ and $t_f \rightarrow +\infty$. Then the functional integral with sources is of the form

$$Z[J] = \int [d\phi] \exp \left\{ i \int d^4x [\mathcal{L}(x) + J(x)\phi(x)] \right\} \tag{157}$$

where $[d\phi] = \prod_x d\phi(x)$, and \mathcal{L} the usual Lagrangian density for the scalar field,

$$\mathcal{L} = -\frac{1}{2} [(\partial_\mu \phi)^2 - \mu^2 \phi^2 - i\epsilon \phi^2] - V(\phi) \tag{158}$$

However even with an underlying lattice discretization, the integral in Eq. (157) is in general ill-defined without a damping factor, due to the i in the exponent (Zinn-Justin, 2003).

Advances in axiomatic field theory (Osterwalder and Schrader, 1973; Glimm and Jaffe 1974; Glimm and Jaffe, 1981) indicate that if one is able to construct a well defined field theory in

Euclidean space $x = (\mathbf{x}, \tau)$ obeying certain axioms, then there is a corresponding field theory in Minkowski space (\mathbf{x}, t) with

$$t = -i\tau \tag{159}$$

defined as an analytic continuation of the Euclidean theory, such that it obeys the Wightman axioms (Streater and Wightman, 2000). The latter is known as the *Euclidicity Postulate*, which states that the Minkowski Green's functions are obtained by analytic continuation of the Green's function derived from the Euclidean functional. One of the earliest discussion of the connection between Euclidean and Minkowski field theory can be found in (Symanzik, 1969). In cases where the Minkowski theory appears pathological, the situation generally does not improve by rotating to Euclidean space. Conversely, if the Euclidean theory is pathological, the problems are generally not removed by considering the Lorentzian case. From a constructive field theory point of view it seems difficult for example to make sense, for either signature, out of one of the simplest cases: a scalar field theory where the kinetic term has the wrong sign (Gallavotti, 1985).

Then the Euclidean functional integral with sources is defined as

$$Z_E[J] = \int [d\phi] \exp \left\{ - \int d^4x [\mathcal{L}_E(x) + J(x)\phi(x)] \right\} \tag{160}$$

with $\int \mathcal{L}_E$ the Euclidean action, and

$$\mathcal{L}_E = \frac{1}{2}(\partial_\mu\phi)^2 + \frac{1}{2}\mu^2\phi^2 + V(\phi) \tag{161}$$

with now $(\partial_\mu\phi)^2 = (\nabla\phi)^2 + (\partial\phi/\partial\tau)^2$. If the potential $V(\phi)$ is bounded from below, then the integral in Eq. (160) is expected to be convergent. In addition, the Euclidicity Postulate determines the correct boundary conditions to be imposed on the propagator (the Feynman $i\epsilon$ prescription). Euclidean field theory has a close and deep connection with statistical field theory and critical phenomena, whose foundations are surveyed for example in the monographs of (Parisi, 1982) and (Cardy, 1996).

Turning to the case of gravity, it should be clear that to all orders in the weak field expansion there is really no difference of substance between the Lorentzian (or pseudo-Riemannian) and the Euclidean (or Riemannian) formulation. Indeed most, if not all, of the perturbative calculations in the preceding sections could have been carried out with the Riemannian weak field expansion about flat Euclidean space

$$g_{\mu\nu} = \delta_{\mu\nu} + h_{\mu\nu} \tag{162}$$

with signature + + + +, or about some suitable classical Riemannian background manifold, without any change of substance in the results. The structure of the divergences would have been identical, and the renormalization group properties of the coupling the same (up to the trivial replacement of say the Minkowski momentum q^2 by its Euclidean expression $q^2 = q_0^2 + \mathbf{q}^2$ etc.). Starting from the Euclidean result, the analytic continuation of results such as Eq. (128) to the pseudo-Riemannian case would have been trivial.

3. Gravitational Functional Measure

It is still true in function space that one needs a metric before one can define a volume element. Therefore, following De Witt (De Witt 1962), one needs first to define an invariant norm for metric deformations

$$\|\delta g\|^2 = \int d^d x \delta g_{\mu\nu}(x) G^{\mu\nu,\alpha\beta}(g(x)) \delta g_{\alpha\beta}(x) , \quad (163)$$

with the inverse of the super-metric G given by the ultra-local expression

$$G^{\mu\nu,\alpha\beta}(g(x)) = \frac{1}{2} \sqrt{g(x)} \left[g^{\mu\alpha}(x) g^{\nu\beta}(x) + g^{\mu\beta}(x) g^{\nu\alpha}(x) + \lambda g^{\mu\nu}(x) g^{\alpha\beta}(x) \right] \quad (164)$$

with λ an arbitrary real parameter. The De Witt supermetric then defines a suitable volume element \sqrt{G} in function space, such that the functional measure over the $g_{\mu\nu}$'s taken on the form

$$\int [d g_{\mu\nu}] \equiv \int \prod_x \left[\det G(g(x)) \right]^{1/2} \prod_{\mu \geq \nu} d g_{\mu\nu}(x) . \quad (165)$$

The assumed locality of the super-metric $G^{\mu\nu,\alpha\beta}(g(x))$ implies that its determinant is a local function of x as well. By a scaling argument given below one finds that, up to an inessential multiplicative constant, the determinant of the supermetric is given by

$$\det G(g(x)) \propto \left(1 + \frac{1}{2} d \lambda\right) [g(x)]^{(d-4)(d+1)/4} . \quad (166)$$

which shows that one needs to impose the condition $\lambda \neq -2/d$ in order to avoid the vanishing of $\det G$. Thus the local measure for the Feynman path integral for pure gravity is given by

$$\int \prod_x [g(x)]^{(d-4)(d+1)/8} \prod_{\mu \geq \nu} d g_{\mu\nu}(x) \quad (167)$$

In four dimensions this becomes simply

$$\int [d g_{\mu\nu}] = \int \prod_x \prod_{\mu \geq \nu} d g_{\mu\nu}(x) \quad (168)$$

However it is not obvious that the above construction is unique. One could have defined, instead of Eq. (164), G to be almost the same, but without the \sqrt{g} factor in front,

$$G^{\mu\nu,\alpha\beta}(g(x)) = \frac{1}{2} \left[g^{\mu\alpha}(x)g^{\nu\beta}(x) + g^{\mu\beta}(x)g^{\nu\alpha}(x) + \lambda g^{\mu\nu}(x)g^{\alpha\beta}(x) \right] \quad (169)$$

Then one would have obtained

$$\det G(g(x)) \propto (1 + \frac{1}{2} d \lambda) [g(x)]^{-(d+1)} \quad , \quad (170)$$

and the local measure for the path integral for gravity would have been given now by

$$\int \prod_x [g(x)]^{-(d+1)/2} \prod_{\mu \geq \nu} dg_{\mu\nu}(x) \quad . \quad (171)$$

In four dimensions this becomes

$$\int [d g_{\mu\nu}] = \int \prod_x [g(x)]^{-5/2} \prod_{\mu \geq \nu} dg_{\mu\nu}(x) \quad (172)$$

which was originally suggested in (Misner, 1957).

One can find in the original reference an argument suggesting that the last measure is unique, provided the product \prod_x is interpreted over 'physical' points, and invariance is imposed at one and the same 'physical' point. Furthermore since there are $d(d+1)/2$ independent components of the metric in d dimensions, the Misner measure is seen to be invariant under a re-scaling $g_{\mu\nu} \rightarrow \Omega^2 g_{\mu\nu}$ of the metric for any d , but as a result is also found to be singular at small g . Indeed the De Witt measure of Eq. (167) and the Misner scale invariant measure of Eqs. (171) and (172) could be just as well regarded as two special cases of a slightly more general supermetric G with prefactor $\sqrt{g}^{(1-\omega)}$, with $\omega = 0$ and $\omega = 1$ corresponding to the original De Witt and Misner measures, respectively.

The power in Eqs. (166) and (167) can be found for example as follows. In the Misner case, Eq. (171), the scale invariance of the functional measure follows directly from the original form of the supermetric $G(g)$ in Eq. (169), and the fact that the metric $g_{\mu\nu}$ has $\frac{1}{2}d(d+1)$ independent components in d dimensions. In the DeWitt case one rescales the matrix $G(g)$ by a factor \sqrt{g} . Since $G(g)$ is a $\frac{1}{2}d(d+1) \times \frac{1}{2}d(d+1)$ matrix, its determinant is modified by an overall factor of $g^{d(d+1)/4}$. So the required power in the functional measure is $-\frac{1}{2}(d+1) + \frac{1}{8}d(d+1) = \frac{1}{8}(d-4)(d+1)$, in agreement with Eq. (167).

Furthermore, one can show that if one introduces an n -component scalar field $\phi(x)$ in the functional integral, it leads to further changes in the gravitational measure. First, in complete

analogy to the gravitational case, one has for the scalar field deformation

$$\|\delta\phi\|^2 = \int d^d x \sqrt{g(x)} (\delta\phi(x))^2 , \quad (173)$$

and therefore for the functional measure over ϕ one has the expression

$$\int [d\phi] = \int \prod_x [\sqrt{g(x)}]^{n/2} \prod_x d\phi(x) . \quad (174)$$

The first factor clearly represents an additional contribution to the gravitational measure. One can indeed verify that one just followed the correct procedure, by evaluating for example the scalar functional integral in the large mass limit,

$$\int \prod_x [\sqrt{g(x)}]^{n/2} \prod_x d\phi(x) \exp\left(-\frac{1}{2}m^2 \int \sqrt{g} \phi^2\right) = \left(\frac{2\pi}{m^2}\right)^{nV/2} = \text{const.} \quad (175)$$

so that, as expected, for a large scalar mass m the field ϕ completely decouples, leaving the dynamics of pure gravity unaffected.

These arguments would lead one to suspect that the volume factor $g^{\sigma/2}$, when included in a slightly more general gravitational functional measure of the form

$$\int [dg_{\mu\nu}] = \prod_x [g(x)]^{\sigma/2} \prod_{\mu \geq \nu} dg_{\mu\nu}(x) , \quad (176)$$

perhaps does not play much of a role after all, at least as far as physical properties are concerned. Furthermore, in d dimensions the \sqrt{g} volume factors are entirely absent ($\sigma = 0$) if one chooses $\omega = 1 - 4/d$, which would certainly seem the simplest choice from a practical point of view.

When considering a Hamiltonian approach to quantum gravity, one finds a rather different form for the functional measure (Leutwyler, 1964), which now includes non-covariant terms. This is not entirely surprising, as the introduction of a Hamiltonian requires the definition of a time variable and therefore a preferred direction, and a specific choice of gauge. The full invariance properties of the original action are no longer manifest in this approach, which is further reflected in the use of a rigid lattice to properly define and regulate the Hamiltonian path integral, allowing subsequent formal manipulations to have a well defined meaning. In the covariant approach one can regard formally the measure contribution as effectively a modification of the Lagrangian, leading to an L_{eff} . The additional terms, if treated consistently will result in a modification of the Hamiltonian, which therefore in general will not be of the form one would have naively guessed from the canonical rules (Abers, 2004). One can see therefore that the possible original measure ambiguity found in the covariant approach is still present in the canonical formulation. One new aspect of the Hamiltonian approach is though that conservation of probability, which implies the unitarity of

the scattering matrix, can further restrict the form of the measure, if such a requirement is pushed down all the way to the cutoff scale (in a simplicial lattice context, the latter would be equivalent to the requirement of Osterwalder-Schrader reflection positivity at the cutoff scale). Whether such a requirement is physical and meaningful in a geometry that is strongly fluctuating at short distances, and for which a notion of time and orthogonal space-like hypersurfaces is not necessarily well defined, remains an open question, and perhaps mainly an academic one. When an ultraviolet cutoff is introduced (without which the theory would not be well defined), one is after all concerned in the end only with distance scales which are much larger than this short distance cutoff.

Along these lines, the following argument supporting the possible irrelevance of the measure parameter σ can be given (Faddeev and Popov, 1973; Fradkin and Vilkovisky, 1973). Namely, one can show that the gravitational functional measure of Eq. (176) is invariant under infinitesimal general coordinate transformations, irrespective of the value of σ . Under an infinitesimal change of coordinates $x'^{\mu} = x^{\mu} + \epsilon^{\mu}(x)$ one has

$$\prod_x [g(x)]^{\sigma/2} \prod_{\mu \geq \nu} dg_{\mu\nu}(x) \rightarrow \prod_x \left(\det \frac{\partial x'^{\beta}}{\partial x^{\alpha}} \right)^{\gamma} [g(x)]^{\sigma/2} \prod_{\mu \geq \nu} dg_{\mu\nu}(x) . \quad (177)$$

with γ a power that depends on σ and the dimension. But for an infinitesimal coordinate transformations the additional factor is equal to one,

$$\prod_x \left(\det \frac{\partial x'^{\beta}}{\partial x^{\alpha}} \right)^{\gamma} = \prod_x [\det(\delta_{\alpha}^{\beta} + \partial_{\alpha}\epsilon^{\beta})]^{\gamma} = \exp \left\{ \gamma \delta^d(0) \int d^d x \partial_{\alpha}\epsilon^{\alpha} \right\} = 1 . \quad (178)$$

and we have used

$$a^d \sum_x \rightarrow \int d^d x \quad (179)$$

with lattice spacing $a = \pi/\Lambda$ and momentum cutoff Λ [see Eq. (83)]. So in some respects it appears that σ can be compared to a gauge parameter.

In conclusion, there is no clear a priori way of deciding between the various choices for σ , and the evidence so far suggests that it may very well turn out to be an irrelevant parameter. The only constraint seems that the regularized gravitational path integral should be well defined, which would seem to rule out singular measures, which need additional regularizations at small volumes. It is noteworthy though that the $g^{\sigma/2}$ volume term in the measure is completely local and contains no derivatives. Thus in perturbation theory it cannot affect the propagation properties of gravitons, and only contributes ultralocal $\delta^d(0)$ terms to the effective action, as can be seen from

$$\prod_x [g(x)]^{\sigma/2} = \exp \left\{ \frac{1}{2} \sigma \delta^d(0) \int d^d x \ln g(x) \right\} . \quad (180)$$

with

$$\ln g(x) = \frac{1}{2} h_{\mu}^{\mu} - \frac{1}{4} h_{\mu\nu} h^{\mu\nu} + O(h^3) \quad (181)$$

which follows from the general formal expansion formula for an operator $\mathbf{M} \equiv 1 + \mathbf{K}$

$$\text{tr} \ln(1 + \mathbf{K}) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \text{tr} \mathbf{K}^n \quad (182)$$

which is valid provided the traces of all powers of \mathbf{K} exist. On a spacetime lattice one can interpret the delta function as an ultraviolet cutoff term, $\delta^d(0) \approx \Lambda^d$. Then the first term shifts the vacuum solution and the second one modifies the bare cosmological constant. To some extent these type of contributions can be regarded as similar to the effects arising from a renormalization of the cosmological constant, ultimately affecting only the distribution of local volumes. So far numerical studies of the lattice models to be discussed later show no evidence of any sensitivity of the critical exponents to the measure parameter σ .

Later in this review (Sect. III.A.7) we will again return to the issue of the functional measure for gravity in possibly the only context where it can be posed, and to some extent answered, satisfactorily: in a lattice regularized version of quantum gravity, going back to the spirit of the original definition of Eq. (153).

In conclusion, the Euclidean Feynman path integral for pure Einstein gravity with a cosmological constant term is given by

$$Z_{cont} = \int [d g_{\mu\nu}] \exp\left\{-\lambda_0 \int dx \sqrt{g} + \frac{1}{16\pi G} \int dx \sqrt{g} R\right\} \quad (183)$$

It involves a functional integration over all metrics, with measure given by a suitably regularized form of

$$\int [d g_{\mu\nu}] = \int \prod_x (g(x))^{\sigma/2} \prod_{\mu \geq \nu} d g_{\mu\nu}(x) \quad (184)$$

as in Eqs. (167), (171) and (176). For geometries with boundaries, further terms should be added to the action, representing the effects of those boundaries. Then the path integral will depend in general on some specified initial and final three-geometry (Hartle and Hawking 1977).

4. Conformal Instability

Euclidean quantum gravity suffers potentially from a disastrous problem associated with the conformal instability: the presence of kinetic contributions to the linearized action entering with the wrong sign.

As was discussed previously in Sec. II.A.5, the action for linearized gravity without a cosmological constant term, Eq. (7), can be conveniently written using the three spin projection operators $P^{(0)}, P^{(1)}, P^{(2)}$ as

$$I_{\text{lin}} = \frac{k}{4} \int dx h^{\mu\nu} [P^{(2)} - 2P^{(0)}]_{\mu\nu\alpha\beta} \partial^2 h^{\alpha\beta} \quad (185)$$

so that the spin-zero mode enters with the wrong sign, or what is normally referred to as a ghost contribution. Actually to this order it can be removed by a suitable choice of gauge, in which the trace mode is made to vanish, as can be seen, for example, in Eq. (13). Still, if one were to integrate in the functional integral over the spin-zero mode, one would have to distort the integration contour to complex values, so as to render the functional integral convergent.

The problem is not removed by introducing higher derivative terms, as can be seen from the action for the linearized theory of Eq. (117),

$$I_{\text{lin}} = \frac{1}{2} \int dx \left\{ h^{\mu\nu} \left[\frac{1}{2} k + \frac{1}{2} a (-\partial^2) \right] (-\partial^2) P_{\mu\nu\rho\sigma}^{(2)} h^{\rho\sigma} \right. \\ \left. + h^{\mu\nu} \left[-k - 2b (-\partial^2) \right] (-\partial^2) P_{\mu\nu\rho\sigma}^{(0)} h^{\rho\sigma} \right\} \quad (186)$$

as the instability reappears for small momenta, where the higher derivative terms can be ignored (see for example Eq. (119)). There is a slight improvement, as the instability is cured for large momenta, but it is not for small ones. If the perturbative quantum calculations can be used as a guide, then at the fixed points one has $b < 0$, corresponding to a tachyon pole in the spin-zero sector, which would indicate further perturbative instabilities. Of course in perturbation theory there never is a real problem, with or without higher derivatives, as one can just define Gaussian integrals by a suitable analytic continuation.

But the instability seen in the weak field limit is not an artifact of the weak field expansion. If one attempts to write down a path integral for pure gravity of the form

$$Z = \int [dg_{\mu\nu}] e^{-I_E} \quad (187)$$

with an Euclidean action

$$I_E = \lambda_0 \int dx \sqrt{g} - \frac{1}{16\pi G} \int dx \sqrt{g} R \quad (188)$$

one realizes that it too appears ill defined due to the fact that the scalar curvature can become arbitrarily positive, or negative. In turn this can be seen as a direct consequence of the fact that while gravitational radiation has positive energy, gravitational potential energy is negative because gravity is attractive. To see more clearly that the gravitational action can be made arbitrarily

negative consider the conformal transformation $\tilde{g}_{\mu\nu} = \Omega^2 g_{\mu\nu}$ where Ω is some positive function. Then the Einstein action transforms into

$$I_E(\tilde{g}) = -\frac{1}{16\pi G} \int d^4x \sqrt{\tilde{g}} (\Omega^2 R + 6 g^{\mu\nu} \partial_\mu \Omega \partial_\nu \Omega) . \quad (189)$$

which can be made arbitrarily negative by choosing a rapidly varying conformal factor Ω . Indeed in the simplest case of a metric $g_{\mu\nu} = \Omega^2 \eta_{\mu\nu}$ one has

$$\sqrt{\tilde{g}} (R - 2\lambda) = 6 g^{\mu\nu} \partial_\mu \Omega \partial_\nu \Omega - 2\lambda \Omega^4 \quad (190)$$

which looks like a $\lambda\phi^4$ theory but with the wrong sign for the kinetic term. The problem is referred to as the conformal instability of the classical Euclidean gravitational action (Hawking 1977). The gravitational action is unbounded from below, and the functional integral is possibly divergent, depending on the detailed nature of the gravitational measure contribution $[dg_{\mu\nu}]$, more specifically its behavior in the regime of strong fields and rapidly varying conformal factors.

A possible solution to the unboundedness problem has been described by Hawking, who suggests performing the integration over all metrics by first integrating over conformal factors by distorting the integration contour in the complex plane to avoid the unboundedness problem, followed by an integration over conformal equivalence classes of metrics (Gibbons and Hawking, 1977; Hawking, 1978; Gibbons, Hawking and Perry, 1978; Gibbons and Perry, 1978). Explicit examples have been given where manifestly convergent Euclidean functional integrals have been formulated in terms of physical (transverse-traceless) degrees of freedom, where the weighting can be shown to arise from a manifestly positive action (Schleich, 1985; Schleich 1987). A similar convergent procedure seem obtainable for some so-called minisuperspace models, where the full functional integration over the fluctuating metric is replace by a finite dimensional integral over a set of parameters characterizing the reduced subspace of the metric in question, see for example (Barvinsky, 2007). But it is unclear how this procedure can be applied outside perturbation theory, where it not obvious how such a split for the metric should be performed.

An alternate possibility is that the unboundedness of the classical Euclidean gravitational action (which in the general case is certainly physical, and cannot therefore be simply removed by a suitable choice of gauge) is not necessarily an obstacle to defining the quantum theory. The quantum mechanical attractive Coulomb well problem has, for zero orbital angular momentum or in the one-dimensional case, a similar type of instability, since the action there is also unbounded from below. The way the quantum mechanical treatment ultimately evades the problem is that the particle has a vanishingly small probability amplitude to fall into the infinitely deep well. In

other words, the effect of quantum mechanical fluctuations in the paths (their zig-zag motion) is just as important as the fact that the action is unbounded. Not unexpectedly, the Feynman path integral solution of the Coulomb problem requires again first the introduction of a lattice, and then a very careful treatment of the behavior close to the singularity (Kleinert, 2006). For this particular problem one is of course aided by the fact that the exact solution is known from the Schrödinger theory.

In quantum gravity the question regarding the conformal instability can then be rephrased in a similar way: Will the quantum fluctuations in the metric be strong enough so that physical excitations will not fall into the conformal well? Phrased differently, what is the role of a non-trivial gravitational measure, giving rise to a density of states $n(E)$

$$Z \propto \int_0^\infty dE n(E) e^{-E}, \quad (191)$$

regarding the issue of ultimate convergence (or divergence) of the Euclidean path integral. Of course to answer such questions satisfactorily one needs a formulation which is not restricted to small fluctuations and to the weak field limit. Ultimately in the lattice theory the answer is yes, for sufficiently strong coupling G (Hamber and Williams, 1984; Berg 1985).

C. Gravity in $2 + \epsilon$ Dimensions

In the previous sections it was shown that pure Einstein gravity is not perturbatively renormalizable in the traditional sense in four dimensions. To one-loop order higher derivative terms are generated, which, when included in the bare action, lead to potential unitarity problems, whose proper treatment most likely lies outside the perturbative regime. The natural question then arises: Are there any other field theories where the standard perturbative treatment fails, yet for which one can find alternative methods and from them develop consistent predictions? The answer seems unequivocally yes. Outside of gravity, there are two notable examples of field theories, the non-linear sigma model and the self-coupled fermion model, which are not perturbatively renormalizable for $d > 2$, and yet lead to consistent and in some instances testable predictions above $d = 2$.

The key ingredient to all of these results is, as originally recognized by Wilson, the existence of a non-trivial ultraviolet fixed point, a phase transition in the statistical field theory context, with non-trivial universal scaling dimensions (Wilson, 1971; Wilson and Fisher 1972; Wilson, 1973; Gross, 1976). Furthermore, three quite different theoretical approaches are available for comparing predictions: the $2 + \epsilon$ expansion, the large- N limit, and the lattice approach. Within the lattice

approach, several additional techniques are available: the strong coupling expansion, the weak coupling expansion and the numerically exact evaluation of the path integral. Finally, the results for the non-linear sigma model in the scaling regime around the non-trivial ultraviolet fixed point can be compared to high accuracy satellite experiments on three-dimensional systems, and the results agree in some cases to several decimals.

The next three sections will therefore discuss these models from the perspective of those results which will have some relevance later for the gravity case. Of particular interest are predictions for universal corrections to free field behavior, for the scale dependence of couplings, and the role of the non-perturbative correlation length which arises in the strong coupling regime.

Later sections will then discuss the $2 + \epsilon$ expansion for gravity, and what can be learned from it by comparing it to the analogous expansion in the non-linear sigma model. The similarity between the two models is such that they both exhibit a non-trivial ultraviolet fixed point, a two-phase structure, non-trivial exponents and scale-dependent couplings.

1. Perturbatively Non-renormalizable Theories: The Sigma Model

The $O(N)$ -symmetric non-linear σ -model provides an instructive and rich example of a theory which, above two dimensions, is not perturbatively renormalizable in the traditional sense, and yet can be studied in a controlled way in the context of Wilson's $2 + \epsilon$ expansion. Such framework provides a consistent way to calculate nontrivial scaling properties of the theory in those dimensions where it is not perturbatively renormalizable (e.g. $d = 3$ and $d = 4$), which can then be compared to non-perturbative results based on the lattice theory, as well as to experiments, since in $d = 3$ the model describes either a ferromagnet or superfluid helium in the vicinity of its critical point. In addition, the model can be solved exactly in the large N limit for any d , without any reliance on the $2 + \epsilon$ expansion. Remarkably, in all three approaches it exhibits a non-trivial ultraviolet fixed point at some coupling g_c (a phase transition in statistical mechanics language), separating a weak coupling massless ordered phase from a massive strong coupling phase.

The non-linear σ -model is described by an N -component field ϕ_a satisfying a unit constraint $\phi^2(x) = 1$, with functional integral given by

$$\begin{aligned}
 Z[J] = & \int [d\phi] \prod_x \delta[\phi(x) \cdot \phi(x) - 1] \\
 & \times \exp\left(-\frac{\Lambda^{d-2}}{g} S(\phi) + \int d^d x J(x) \cdot \phi(x)\right)
 \end{aligned}
 \tag{192}$$

The action is taken to be $O(N)$ -invariant

$$S(\phi) = \frac{1}{2} \int d^d x \partial_\mu \phi(x) \cdot \partial_\mu \phi(x) \quad (193)$$

Λ here is the ultraviolet cutoff and g the bare dimensionless coupling at the cutoff scale Λ ; in a statistical field theory context g plays the role of a temperature.

In perturbation theory one can eliminate one ϕ field by introducing a convenient parametrization for the unit sphere, $\phi(x) = \{\sigma(x), \pi(x)\}$ where π_a is an $N - 1$ -component field, and then solving locally for $\sigma(x)$

$$\sigma(x) = [1 - \pi^2(x)]^{1/2} \quad (194)$$

In the framework of perturbation theory in g the constraint $|\pi(x)| < 1$ is not important as one is restricting the fluctuations to be small. Nevertheless the π integrations will be extended from $-\infty$ to $+\infty$, which reduces the development of the perturbative expansion to a sequence of Gaussian integrals. Values of $\pi(x) \sim 1$ give exponentially small contributions of order $\exp(-\text{const.}/g)$ which are therefore negligible to any finite order in perturbation theory.

In term of the π field the original action S becomes

$$S(\pi) = \frac{1}{2} \int d^d x \left[(\partial_\mu \pi)^2 + \frac{(\pi \cdot \partial_\mu \pi)^2}{1 - \pi^2} \right] \quad (195)$$

The change of variables from $\phi(x)$ to $\pi(x)$ also gives rise to a Jacobian

$$\prod_x [1 - \pi^2]^{-1/2} \sim \exp \left[-\frac{1}{2} \delta^d(0) \int d^d x \ln(1 - \pi^2) \right] \quad (196)$$

which is necessary for the cancellation of spurious tadpole divergences. The combined functional integral for the unconstrained π field is then given by

$$Z[\mathbf{J}] = \int [d\pi] \exp \left(-\frac{\Lambda^{d-2}}{g} S_0(\pi) + \int d^d x \mathbf{J}(\mathbf{x}) \cdot \pi(x) \right) \quad (197)$$

with

$$S_0(\pi) = \frac{1}{2} \int d^d x \left[(\partial_\mu \pi)^2 + \frac{(\pi \cdot \partial_\mu \pi)^2}{1 - \pi^2} \right] + \frac{1}{2} \delta^d(0) \int d^d x \ln(1 - \pi^2) \quad (198)$$

In perturbation theory the above action is expanded out in powers of π . The propagator for the π field can be read off from the quadratic part of the action,

$$\Delta_{ab}(k^2) = \frac{\delta_{ab}}{k^2} \quad (199)$$

In the weak coupling limit the π fields correspond to the Goldstone modes of the spontaneously broken $O(N)$ symmetry, the latter being broken spontaneously by a non-vanishing vacuum expectation value $\langle \pi \rangle \neq 0$.

Since the π field has mass dimension $\frac{1}{2}(d-2)$, and the interactions $\partial^2 \pi^{2n}$ consequently has dimension $n(d-2) + 2$, one finds that the theory is renormalizable in $d = 2$ and perturbatively non-renormalizable above $d = 2$. Furthermore, in spite of the theory being non-polynomial, it can still be renormalized via the introduction of only two renormalization constant, the coupling renormalization being given by a constant Z_g and the wavefunction renormalization by a constant π by Z . Potential infrared problems due to massless propagators are handled by introducing an external h -field term for the original composite σ field, which then acts as a mass term for the π field,

$$\begin{aligned} h \int d^d x \sigma(x) &= h \int d^d x [1 - \pi^2(x)]^{1/2} \\ &= \int d^d x [h - \frac{1}{2} h \pi^2(x) + \dots] \end{aligned} \quad (200)$$

A proof can be found (David, 1982) that all $O(N)$ invariant Green's are infrared finite in the limit $h \rightarrow 0$.

One can write down the same field theory on a lattice, where it corresponds to the $O(N)$ -symmetric classical Heisenberg model at a finite temperature $T \sim g$. The simplest procedure is to introduce a hypercubic lattice of spacing a , with sites labeled by integers $\mathbf{n} = (n_1 \dots n_d)$, which introduces an ultraviolet cutoff $\Lambda \sim \pi/a$. On the lattice field derivatives are replaced by finite differences

$$\partial_\mu \phi(x) \rightarrow \Delta_\mu \phi(\mathbf{n}) = \frac{\phi(\mathbf{n} + \mu) - \phi(\mathbf{n})}{a} \quad (201)$$

and the discretized path integral then reads

$$\begin{aligned} Z[\mathbf{J}] &= \prod_{\mathbf{n}} d\phi(\mathbf{n}) \delta[\phi^2(\mathbf{n}) - 1] \\ &\times \exp \left[-\frac{a^{2-d}}{2g} \sum_{\mathbf{n}, \mu} (\Delta_\mu \phi(\mathbf{n}))^2 + \sum_{\mathbf{n}} \mathbf{J}(\mathbf{x}) \cdot \phi(x) \right] \end{aligned} \quad (202)$$

The above expression is recognized as the partition function for a ferromagnetic $O(N)$ -symmetric spin system at finite temperature. Besides ferromagnets, it can be used to describe systems which are related to it by universality, such as superconductors and superfluid helium transitions, whose

critical behavior is described by a complex phase, and which are therefore directly connected to the plane rotator $N = 2$, or $U(1)$, model.

In addition the lattice model of Eq. (202) provides an explicit regularization for the continuum theory, which makes expressions like the one in Eq. (196) acquire a well defined meaning. It is in fact the only regularization which allows a discussion of the role of the measure in perturbation theory (Zinn-Justin, 2002). At the same time it provides an ultraviolet regularization for perturbation theory, and allows for various non-perturbative calculations, such as power series expansions in three dimensions and explicit numerical integrations of the path integral via Monte Carlo methods.

In two dimensions one can compute the renormalization of the coupling g from the action of Eq. (197) and one finds after a short calculation (Polyakov 1975) for small g

$$\frac{1}{g(\mu)} = \frac{1}{g} + \frac{N-2}{8\pi} \ln \frac{\mu^2}{\Lambda^2} + \dots \quad (203)$$

where μ is an arbitrary momentum scale. Physically one can view the origin of the factor of $N - 2$ in the fact that there are $N - 2$ directions in which the spin can experience rapid small fluctuations perpendicular to its average slow motion on the unit sphere, and that only these fluctuations contribute to leading order (Kogut 1979).

In two dimensions the quantum correction (the second term on the r.h.s.) increases the value of the effective coupling at low momenta (large distances), unless $N = 2$ in which case the correction vanishes. In fact the quantum correction can be shown to vanish to all orders in this case; the vanishing of the β -function in two dimensions for the $O(2)$ model is true only in perturbation theory, for sufficiently strong coupling a phase transition appears, driven by the unbinding of vortex pairs (Kosterlitz and Thouless, 1973). For $N > 2$ as $g(\mu)$ flows toward increasingly strong coupling it eventually leaves the regime where perturbation theory can be considered reliable. But for bare $g \approx 0$ the quantum correction is negligible and the theory is scale invariant around the origin: the only fixed point of the renormalization group, at least in lowest order perturbation theory, is at $g = 0$. For fixed cutoff Λ , the theory is weakly coupled at short distances but strongly coupled at large distances. The results in two dimensions for $N > 2$ are qualitatively very similar to asymptotic freedom in four-dimensional $SU(N)$ Yang-Mills theories.

Above two dimensions, $d - 2 = \epsilon > 0$ and one can redo the same type of perturbative calculation to determine the coupling renormalization. The relevant diagrams are shown in Fig. 4. One finds for the effective coupling g_e , i.e. the coupling which includes the leading radiative correction (using dimensional regularization, which is more convenient than an explicit ultraviolet cutoff Λ

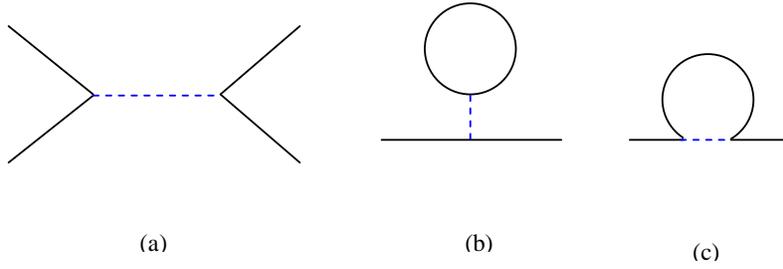


FIG. 4 One-loop diagrams giving rise to coupling and field renormalizations in the non-linear σ -model. Group theory indices a flow along the thick lines, dashed lines should be contracted to a point.

for performing actual perturbative calculations),

$$\frac{1}{g_e} = \frac{\Lambda^\epsilon}{g} \left[1 - \frac{1}{\epsilon} \frac{N-2}{2\pi} g + O(g^2) \right] \quad (204)$$

The requirement that the dimensionful effective coupling g_e be defined independently of the scale Λ is expressed as $\Lambda \frac{d}{d\Lambda} g_e = 0$, and gives for the Callan-Symanzik β -function (Callan, 1970; Symanzik, 1970) for g

$$\Lambda \frac{\partial g}{\partial \Lambda} = \beta(g) = \epsilon g - \frac{N-2}{2\pi} g^2 + O(g^3, \epsilon g^2) \quad (205)$$

The above β -function determines the scale dependence (at least in perturbation theory) of g for an arbitrary scale, which from now on will be denoted as μ . Then the differential equation $\mu \frac{\partial g}{\partial \mu} = \beta(g(\mu))$ uniquely determines how $g(\mu)$ flows as a function of momentum scale μ . The scale dependence of $g(\mu)$ is such that if the initial g is less than the ultraviolet fixed point value g_c , with

$$g_c = \frac{2\pi\epsilon}{N-2} + \dots \quad (206)$$

then the coupling will flow towards the Gaussian fixed point at $g = 0$. The new phase that appears when $\epsilon > 0$ and corresponds to a low temperature, spontaneously broken phase with finite order parameter. On the other hand if $g > g_c$ then the coupling $g(\mu)$ flows towards increasingly strong coupling, and eventually out of reach of perturbation theory. In two dimensions the β -function has no zero and only the strong coupling phase is present.

The one-loop running of g as a function of a sliding momentum scale $\mu = k$ and $\epsilon > 0$ can be obtained by integrating Eq. (205). One finds

$$g(k^2) = \frac{g_c}{1 \pm a_0 (m^2/k^2)^{(d-2)/2}} \quad (207)$$

with a_0 a positive constant and m a mass scale; the combination $a_0 m^{d-2}$ is just the integration constant for the differential equation, which we prefer to split here in a momentum scale and a dimensionless coefficient for reasons that will become clear later. The choice of $+$ or $-$ sign is determined from whether one is to the left ($+$), or to right ($-$) of g_c , in which case $g(k^2)$ decreases or, respectively, increases as one flows away from the ultraviolet fixed point. The renormalization group invariant mass scale $\sim m$ arises here as an arbitrary integration constant of the renormalization group equations, and cannot be determined from perturbative arguments alone. It should also be clear that multiplying both sides of Eq. (207) by the ultraviolet cutoff factor Λ^{2-d} to get back the original dimensionful coupling multiplying the action $S(\phi)$ in Eq. (192) does not change any of the conclusions.

Note that the result of Eq. (207) is quite different from the naive expectation based on straight perturbation theory in $d > 2$ dimensions (where the theory is not perturbatively renormalizable)

$$\frac{g(k^2)}{g} \sim 1 + \text{const. } g k^{d-2} + O(g^2) \quad (208)$$

which gives a much worse ultraviolet behavior. The existence of a non-trivial ultraviolet fixed point alters the naive picture and drastically improves the ultraviolet behavior.

For large g one can easily see, for example from the structure of the lattice action in Eq. (202), that correlation functions must decay exponentially at large separations. In the strong coupling limit, spins separated by a distance $|x|$ will fluctuate in an uncorrelated fashion, unless they are connected by a minimal number of link contributions from the action. One expects therefore for the lattice connected correlation function of two ϕ fields, separated by a lattice distance na ,

$$\langle \pi(na) \pi(0) \rangle_c \underset{n \rightarrow \infty}{\sim} \left(\frac{1}{g} \right)^n \quad (209)$$

which can be re-written in continuum language

$$\langle \pi(x) \pi(0) \rangle_c \underset{|x| \rightarrow \infty}{\sim} \exp(-|x|/\xi) \quad (210)$$

with $m = 1/\xi = \Lambda [\ln g + O(1/g)]$ and $\Lambda = 1/a$. From the requirement that the correlation length ξ be a physical quantity independent of scale, and consequently a renormalization group invariant,

$$\Lambda \frac{d}{d\Lambda} m(\Lambda, g(\Lambda)) = 0 \quad , \quad (211)$$

one obtains in the strong coupling limit

$$\beta(g) = -g \ln g + O(1/g) \quad . \quad (212)$$

If in Eq. (207) one sets the momentum scale k equal to the cutoff scale Λ and solves for m in the strong coupling phase one obtains

$$\frac{g_c}{g(\Lambda)} = 1 - a_0 \left(\frac{m^2}{\Lambda^2} \right)^{(d-2)/2} \quad (213)$$

and therefore for m in terms of the bare coupling $g \equiv g(\Lambda)$

$$m(g) = \Lambda \left(\frac{g_c}{a_0} \right)^{2/(d-2)} \left(\frac{1}{g_c} - \frac{1}{g} \right)^{1/(d-2)} \quad (214)$$

This last result shows the following important fact: if m is identified with the inverse of the correlation length ξ (which can be extracted non-perturbatively, for example from the exponential decay of correlation functions, and is therefore generally unambiguous), then the calculable constant relating m to g in Eq. (214) uniquely determines the coefficient a_0 in Eq. (207). For example, in the large N limit the value for a_0 will be given later in Eq. (245).

In general one can write down the complete renormalization group equations for the cutoff-dependent n -point functions $\Gamma^{(n)}(p_i, g, h, \Lambda)$ (Brezin and Zinn-Justin, 1976; Zinn-Justin, 2002). For this purpose one needs to define the renormalized truncated n -point function $\Gamma_r^{(n)}$,

$$\Gamma_r^{(n)}(p_i, g_r, h_r, \mu) = Z^{n/2}(\Lambda/\mu, g) \Gamma^{(n)}(p_i, g, h, \Lambda) \quad (215)$$

where μ is a renormalization scale, and the constants g_r , h_r and Z are defined by

$$\begin{aligned} g &= (\Lambda/\mu)^{d-2} Z_g g_r & \pi(x) &= Z^{1/2} \pi_r(x) \\ h &= Z_h h_r & Z_h &= Z_g / \sqrt{Z} \end{aligned} \quad (216)$$

The requirement that the renormalized n -point function $\Gamma_r^{(n)}$ be independent of the cutoff Λ then implies

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} - \frac{n}{2} \zeta(g) + \rho(g) h \frac{\partial}{\partial h} \right] \Gamma^{(n)}(p_i, g, h, \Lambda) = 0 \quad (217)$$

with the renormalization group functions $\beta(g)$, $\zeta(g)$ and $\rho(g)$ defined as

$$\begin{aligned} \Lambda \frac{\partial}{\partial \Lambda} \Big|_{\text{ren.fixed}} g &= \beta(g) \\ \Lambda \frac{\partial}{\partial \Lambda} \Big|_{\text{ren.fixed}} (-\ln Z) &= \zeta(g) \\ 2 - d + \frac{1}{2} \zeta(g) + \frac{\beta(g)}{g} &= \rho(g) . \end{aligned} \quad (218)$$

Here the derivatives of the bare coupling g , of the π -field wave function renormalization constant Z and of the external field h with respect to the cutoff Λ are evaluated at fixed renormalized (or effective) coupling, at the renormalization scale μ .

To determine the renormalization group functions $\beta(g)$, $\zeta(g)$ and $\rho(g)$ one can in fact follow a related but equivalent procedure, in which, instead of requiring the renormalized n -point functions $\Gamma_r^{(n)}$ to be independent of the cutoff Λ at fixed renormalization scale μ as in Eq. (218), one imposes that the *bare* n -point functions $\Gamma^{(n)}$ be independent of the renormalization scale μ at *fixed* cutoff Λ . One can show (Brezin, Le Guillou and Zinn-Justin, 1978) that the resulting renormalization group functions are identical to the previous ones, and that one can obtain the scale dependence of the couplings (i.e. $\beta(g)$) either way. Physically the latter way of thinking is perhaps more suited to a situation where one is dealing with a finite cutoff theory, where the ultraviolet cutoff Λ is fixed and one wants to investigate the scale (momentum) dependence of the couplings, for example $g(k^2)$.

For our purposes it will sufficient to look, in the zero-field case $h = 0$, at the β -function of Eq. (205) which incorporates, as should already be clear from the result of Eq. (218), a tremendous amount of information about the model. Herein lies the power of the renormalization group: the knowledge of a handful of functions $[\beta(g), \zeta(g)]$ is sufficient to completely determine the momentum dependence of all n -point functions $\Gamma^{(n)}(p_i, g, h, \Lambda)$.

One can integrate the β -function equation in Eq. (205) to obtain the renormalization group invariant quantity

$$\xi^{-1}(g) = m(g) = \text{const. } \Lambda \exp\left(-\int^g \frac{dg'}{\beta(g')}\right) \quad (219)$$

which is identified with the correlation length appearing, for example, in Eq. (210). The multiplicative constant in front of the expression on the right hand side arises as an integration constant, and cannot be determined from perturbation theory in g . Conversely, it is easy to verify that ξ is indeed a renormalization group invariant, $\Lambda \frac{d}{d\Lambda} \xi(\Lambda, g(\Lambda)) = 0$, as stated previously in Eq. (211).

In the vicinity of the fixed point at g_c one can do the integral in Eq. (219), using Eq. (206) and the resulting linearized expression for the β -function in the vicinity of the non-trivial ultraviolet fixed point,

$$\beta(g) \underset{g \rightarrow g_c}{\sim} \beta'(g_c)(g - g_c) + \dots \quad (220)$$

and one finds

$$\xi^{-1}(g) = m(g) \propto \Lambda |g - g_c|^\nu \quad (221)$$

with a correlation length exponent $\nu = -1/\beta'(g_c) \sim 1/(d - 2) + \dots$. Thus the correlation length $\xi(g)$ diverges as one approaches the fixed point at g_c .

In general the existence of a non-trivial ultraviolet fixed point implies that the large momentum behavior above two dimensions is not given by naive perturbation theory; it is given instead by

the critical behavior of the renormalized theory. In the weak coupling, small g phase the scale m can be regarded as a crossover scale between the free field behavior at large distance scales and the critical behavior which sets in at large momenta.

In the non-linear σ -model another quantity of physical interest is the function $M_0(g)$,

$$M_0(g) = \exp \left[-\frac{1}{2} \int_0^g dg' \frac{\zeta(g')}{\beta(g')} \right] \quad (222)$$

which is proportional to the order parameter (the magnetization) of the non-linear σ -model. To one-loop order one finds $\zeta(g) = \frac{1}{2\pi}(N-1)g + \dots$ and therefore

$$M_0(g) = \text{const.} (g_c - g)^\beta \quad (223)$$

with $\beta = \frac{1}{2}\nu(d-2+\eta)$ and $\eta = \zeta(g_c) - \epsilon$. To leading order in the ϵ expansion one has for the anomalous dimension of the π field $\eta = \epsilon/(N-2) + O(\epsilon^2)$. In gauge theories, including gravity, there is no local order parameter, so this quantity has no obvious generalization there.

In general the ϵ -expansion is only expected to be asymptotic. This is already seen from the expansion for ν which has recently been computed to four loops (Hikami and Brezin 1978, Bernreuther and Wegner 1986, Kleinert 2000)

$$\begin{aligned} \nu^{-1} = & \epsilon + \frac{\epsilon^2}{N-2} + \frac{\epsilon^3}{2(N-2)} \\ & - [30 - 14N + N^2 + (54 - 18N)\zeta(3)] \\ & \times \frac{\epsilon^4}{4(N-2)^3} + \dots \end{aligned} \quad (224)$$

which needs to be summed by Borel-Padé methods to obtain reliable results in three dimensions. For example, for $N=3$ one finds in three dimensions $\nu \approx 0.799$, which can be compared to the $4-\epsilon$ result for the $\lambda\phi^4$ theory to five loops $\nu \simeq 0.705$, to the seven-loop perturbative expansion for the $\lambda\phi^4$ theory directly in $3d$ which gives $\nu \simeq 0.707$, with the high temperature series result $\nu \simeq 0.717$ and the Monte Carlo estimates $\nu \simeq 0.718$, as compiled for example in a recent comprehensive review (Guida and Zinn-Justin, 1997).

There exist standard methods to deal with asymptotic series such as the one in Eq. (224). To this purpose one considers a general series

$$f(g) = \sum_{n=0}^{\infty} f_n g^n \quad (225)$$

and defines its Borel transform as

$$F(b) = \sum_{n=0}^{\infty} \frac{f_n}{n!} b^n \quad (226)$$

One can attempt to sum the series for $F(b)$ using Padé methods and conformal transformations. The original function $f(g)$ is then recovered by performing an integral over the Borel transform variable b

$$f(g) = \frac{1}{g} \int_0^\infty db e^{-b/g} F(b) \quad (227)$$

where the familiar formula

$$\int_0^\infty dz z^n e^{-z/g} = n! g^{n+1} \quad (228)$$

has been used. Bounds on the coefficients f_n suggest that in most cases $F(z)$ is analytic in a circle of radius a around the origin, and that the integral will converge for $|z|$ small enough, within a sector $|\arg z| < \alpha/2$ with typically $\alpha \geq \pi$ (Le Guillou and Zinn-Justin, 1990).

The first singularity along the positive real axis is generally referred to as an infrared renormalon, and is expected to be, in the $2d$ non-linear σ -model, at $b = 1/2\beta_0$ where $\beta_0 = (N - 2)/2\pi$, and gives rise to non-analytic corrections of order $\exp(-2\pi/(N - 2)g)$ (David, 1982). Such non-analytic contributions presumably account for the fact (Cardy and Hamber, 1980) that the $N = 2$ model has a vanishing β -function to all orders in $d = 2$, and yet has non-trivial finite exponents in $d = 3$, in spite of the result of Eq. (224). Indeed the $2 + \epsilon$ expansion is not particularly useful for the special case of the $N = 2$ σ -model. Then the action is simply given by

$$S(\theta) = \frac{\Lambda^{d-2}}{2g} \int d^d x [\partial_\mu \theta(x)]^2 \quad (229)$$

with $\phi_1(x) = \sin \theta(x)$ and $\phi_2(x) = \cos \theta(x)$, describing the fluctuations of a planar spin in d dimensions. The β -function of Eq. (205) then vanishes identically in $d = 2$, and the corrections to ν diverge for $d > 2$, as in Eq. (224)). Yet this appears to be more a pathology of the perturbative expansion in ϵ , since after all the lattice model of Eq. (202) is still well defined, and so is the $4 - \epsilon$ expansion for the continuum linear $O(N)$ σ -model. Thus, in spite of the model being again not perturbatively renormalizable in $d = 3$, one can still develop, for these models, the full machinery of the renormalization group and compute the relevant critical exponents.

Perhaps more importantly, a recent space shuttle experiment (Lipa et al 2003) has succeeded in measuring the specific heat exponent $\alpha = 2 - 3\nu$ of superfluid Helium (which is supposed to share the same universality class as the $N = 2$ non-linear σ -model, with the complex phase of the superfluid condensate acting as the order parameter) to very high accuracy

$$\alpha = -0.0127(3) \quad (230)$$

Previous theoretical predictions for the $N = 2$ model include the most recent four-loop $4 - \epsilon$ continuum result $\alpha = -0.01126(10)$ (Kleinert, 2000), a recent lattice Monte Carlo estimate $\alpha = -0.0146(8)$ (Camposstrini et al, 2001), and the lattice variational renormalization group prediction $\alpha = -0.0125(39)$ (Hamber, 1981).

Perhaps the message one gains from this rather lengthy discussion of the non-linear σ -model in $d > 2$ is that:

- The model provides a specific example of a theory which is not perturbatively renormalizable in the traditional sense, and for which the naive perturbative expansion in fixed dimension leads to uncontrollable divergences and inconsistent results;
- Yet the model can be constructed perturbatively in terms of a double expansion in g and $\epsilon = d - 2$. This new perturbative expansion, combined with the renormalization group, in the end provides explicit and detailed information about universal scaling properties of the theory in the vicinity of the non-trivial ultraviolet point at g_c ;
- The continuum field theory predictions obtained this way generally agree, for distances much larger than the cutoff scale, with lattice results, and, perhaps more importantly, with high precision experiments on systems belonging to the same universality class of the $O(N)$ model.

2. Non-linear Sigma Model in the Large-N Limit

A rather fortunate circumstance is represented by the fact that in the large N limit the non-linear σ -model can be solved exactly. This allows an independent verification of the correctness of the general ideas developed in the previous section, as well as a direct comparison of explicit results for universal quantities. The starting point is the functional integral of Eq. (192),

$$Z = \int [d\phi(x)] \prod_x \delta[\phi^2(x) - 1] \exp(-S(\phi)) \quad (231)$$

with

$$S(\phi) = \frac{1}{2T} \int d^d x \partial_\mu \phi(x) \cdot \partial_\mu \phi(x) \quad (232)$$

The constraint on the ϕ field can be implemented via an auxiliary Lagrange multiplier field $\alpha(x)$. One writes

$$Z = \int [d\phi(x)] [d\alpha(x)] \exp(-S(\phi, \alpha)) \quad (233)$$

with

$$S(\phi, \alpha) = \frac{1}{2T} \int d^d x \left[(\partial_\mu \phi(x))^2 + \alpha(x)(\phi^2(x) - 1) \right] \quad (234)$$

Since the action is now quadratic in $\phi(x)$ one can integrate over $N - 1$ ϕ -fields (denoted previously by π). The resulting determinant is then re-exponentiated, and one is left with a functional integral over the remaining first field $\phi_1(x) \equiv \sigma(x)$, as well as the Lagrange multiplier field $\alpha(x)$,

$$Z = \int [d\sigma(x) d\alpha(x)] \exp(-S_N(\phi, \alpha)) \quad (235)$$

with now

$$S_N(\phi, \alpha) = \frac{1}{2T} \int d^d x \left[(\partial_\mu \sigma)^2 + \alpha(\sigma^2 - 1) \right] + \frac{1}{2}(N - 1) \text{tr} \ln[-\partial^2 + \alpha] \quad (236)$$

In the large N limit one can neglect, to leading order, fluctuations in the α and σ fields. For a constant α field, $\langle \alpha(x) \rangle = m^2$, the last (trace) term can be written in momentum space as

$$\frac{1}{2}(N - 1) \int^\Lambda \frac{d^d k}{(2\pi)^d} \ln(k^2 + m^2) \quad (237)$$

which makes the evaluation of the trace straightforward. As should be clear from Eq. (234), the parameter m can be interpreted as the mass of the ϕ field. The functional integral in Eq. (235) can then be evaluated by the saddle point method. It is easy to see from Eq. (236) that the saddle point conditions are

$$\begin{aligned} \sigma^2 &= 1 - (N - 1)\Omega_d(m)T \\ m^2\sigma &= 0 \end{aligned} \quad (238)$$

with the function $\Omega_d(m)$ given by the integral

$$\Omega_d = \int^\Lambda \frac{d^d k}{(2\pi)^d} \frac{1}{k^2 + m^2} \quad (239)$$

The latter can be evaluated in terms of a hypergeometric function,

$$\Omega_d = \frac{1}{2^{d-1}\pi^{d/2}\Gamma(d/2)} \frac{\Lambda^d}{m^2 d} {}_2F_1 \left[1, \frac{d}{2}; 1 + \frac{d}{2}; -\frac{\Lambda^2}{m^2} \right] \quad (240)$$

but here one only really needs it in the large cutoff limit, $m \ll \Lambda$, in which case one finds the more tractable expression

$$\Omega_d(m) - \Omega_d(0) = m^2 [c_1 m^{d-4} + c_2 \Lambda^{d-4} + O(m^2 \Lambda^{d-6})] \quad (241)$$

with c_1 and c_2 some d -dependent coefficients.

From Eq. (238) one notices that at weak coupling and for $d > 2$ a non-vanishing σ -field expectation value implies that m , the mass of the π field, is zero. If one sets $(N - 1)\Omega_d(0) = 1/T_c$, one can then write the first expression in Eq. (238) as

$$\sigma(T) = \pm [1 - T/T_c]^{1/2} \quad (242)$$

which shows that T_c is the critical coupling at which the order parameter σ vanishes.

Above T_c the order parameter σ vanishes, and $m(T)$ is obtained, from Eq. (238), by the solution of the nonlinear gap equation

$$\frac{1}{T} = (N - 1) \int^\Lambda \frac{d^d k}{(2\pi)^d} \frac{1}{k^2 + m^2} \quad (243)$$

Using the definition of the critical coupling T_c , one can now write, for $2 < d < 4$, for the common mass of the σ and π fields

$$m(T) \underset{m \ll \Lambda}{\sim} \left(\frac{1}{T_c} - \frac{1}{T} \right)^{1/(d-2)} \quad (244)$$

which gives for the correlation length exponent the non-gaussian value $\nu = 1/(d - 2)$, with the gaussian value $\nu = 1/2$ being recovered as expected at $d = 4$ (Wilson and Fisher 1972). Note that in the large N limit the constant of proportionality in Eq. (244) is completely determined by the explicit expression for $\Omega_d(m)$.

Perhaps one of the most striking aspects of the non-linear sigma model above two dimensions is that all particles are massless in perturbation theory, yet they all become massive in the strong coupling phase $T > T_c$, with masses proportional to the non-perturbative scale m .

Again one can perform a renormalization group analysis as was done in the previous section in the context of the $2 + \epsilon$ expansion. To this end one defines dimensionless coupling constants $g = \Lambda^{d-2}T$ and $g_c = \Lambda^{d-2}T_c$ as was done in Eq. (192). Then the non-perturbative result of Eq. (244) becomes

$$m(g) \simeq c_d \cdot \Lambda \left(\frac{1}{g_c} - \frac{1}{g} \right)^{1/(d-2)} \quad (245)$$

with the numerical coefficient given by $c_d = [\frac{1}{2}(d-2)\pi | \csc(\frac{d\pi}{2}) |]^{-\frac{1}{d-2}}$. One welcome feature of this large- N result is the fact that it provides an explicit value for the coefficient in Eq. (214), namely

$$c_d = \left(\frac{g_c}{a_0} \right)^{1/(d-2)}, \quad (246)$$

and thereby for the numerical factor a_0 appearing in Eqs. (214) and (207).

Again the physical, dimensionful mass m in Eqs. (244) or (245) is required to be scale- and cutoff-independent as in Eq. (211)

$$\Lambda \frac{d}{d\Lambda} m(\Lambda, g(\Lambda)) = 0 \quad (247)$$

or, more explicitly, using the expression for m in Eq. (245),

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g} \right] \Lambda \left(\frac{1}{g_c} - \frac{1}{g} \right)^{1/(d-2)} = 0 \quad (248)$$

which implies for the $O(N)$ β -function in the large N limit the simple result

$$\beta(g) = (d-2)g(1-g/g_c) \quad (249)$$

The latter is valid again in the vicinity of the fixed point at g_c , due to the assumption, used in Eq. (244), of $m \ll \Lambda$. Note that it vanishes in $d = 2$, and for $g = 0$, in agreement with the $2 + \epsilon$ result of Eq. (205). Furthermore Eq. (249) gives the momentum dependence of the coupling at fixed cutoff. After integration, one finds for the momentum (μ) dependence of the coupling at fixed cutoff Λ

$$\frac{g(\mu)}{g_c} = \frac{1}{1 - c(\mu_0/\mu)^{d-2}} \approx 1 + c(\mu_0/\mu)^{d-2} + \dots \quad (250)$$

with $c\mu_0^{d-2}$ the integration constant. The sign of c then depends on whether one is on the right ($c > 0$) or on the left ($c < 0$) of the ultraviolet fixed point at g_c .

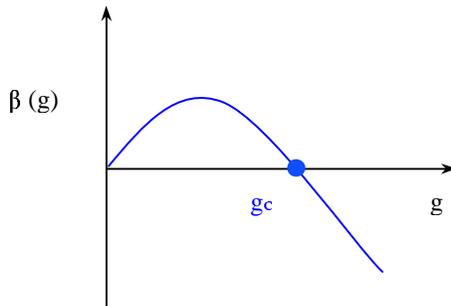


FIG. 5 The β -function for the non-linear σ -model in the large- N limit for $d > 2$.

One notices therefore again that the general shape of $\beta(g)$ is of the type shown in Fig. 5, with g_c a stable non-trivial UV fixed point, and $g = 0$ and $g = \infty$ two stable (trivial) IR fixed points. Once more, at the critical point g_c the β -function vanishes and the theory becomes scale invariant.

Furthermore one can check that again $\nu = -1/\beta'(g_c)$ where ν is the exponent in Eq. (244). As before one can re-write the physical mass m for $2 < d < 4$ as

$$\xi^{-1}(g) = m(g) \propto \Lambda \exp\left(-\int^g \frac{dg'}{\beta(g')}\right) \quad (251)$$

as was done previously in Eq. (219).

Another general lesson one learns is that that Eq. (247),

$$\left[\Lambda \frac{\partial}{\partial \Lambda} + \beta(g) \frac{\partial}{\partial g}\right] m(\Lambda, g(\Lambda)) = 0 \quad (252)$$

can be used to provide a non-perturbative definition for the β -function $\beta(g)$. If one sets $m = \Lambda F(g)$, with $F(g)$ a dimensionless function of g , then one has the simple result

$$\beta(g) = -\frac{F(g)}{F'(g)} \quad (253)$$

Thus the knowledge of the dependence of the mass gap m on the bare coupling g fixes the shape of the β function, at least in the vicinity of the fixed point. It should be clear then that the definition of the β -function per se, and therefore the scale dependence of $g(\mu)$ which follows from it [as determined from the solution of the differential equation $\mu \frac{\partial g}{\partial \mu} = \beta(g(\mu))$] is *not* necessarily tied to perturbation theory.

When N is large but finite, one can develop a systematic $1/N$ expansion in order to evaluate the corrections to the picture presented above (Zinn-Justin, 2002). Corrections to the exponents are known up to order $1/N^2$, but the expressions are rather complicated for arbitrary d and will not be reproduced here. In general it appears that the $1/N$ expansion is only asymptotic, and somewhat slowly convergent for useful values of N in three dimensions.

3. Self-coupled Fermion Model

Finally it would seem worthwhile to mention another example of a theory which naively is not perturbatively renormalizable in $d > 2$, and yet whose critical properties can be worked out both in the $2 + \epsilon$ expansion, and in the large N limit. It is described by an $U(N)$ -invariant action containing a set of N massless self-coupled Dirac fermions (Wilson, 1973; Gross and Neveu, 1974)

$$S(\psi, \bar{\psi}) = -\int d^d x [\bar{\psi} \cdot \not{\partial} \psi + \frac{1}{2} \Lambda^{d-2} u (\bar{\psi} \cdot \psi)^2] . \quad (254)$$

In even dimensions the discrete chiral symmetry $\psi \rightarrow \gamma_5 \psi$, $\bar{\psi} \rightarrow -\bar{\psi} \gamma_5$ prevents the appearance of a fermion mass term. Interest in the model resides in the fact that it exhibits a mechanism for dynamical mass generation and chiral symmetry breaking.

In two dimensions the fermion self-coupling constant is dimensionless, and after setting $d = 2 + \epsilon$ one is again ready to develop the full machinery of the perturbative expansion in u and ϵ , as was done for the non-linear σ -model, since the model is again believed to be multiplicatively renormalizable in the framework of the $2 + \epsilon$ expansion. For the β -function one finds to three loops

$$\beta(u) = \epsilon u - \frac{\bar{N} - 2}{2\pi} u^2 + \frac{\bar{N} - 2}{4\pi^2} u^3 + \frac{(\bar{N} - 2)(\bar{N} - 7)}{32\pi^3} u^4 + \dots \quad (255)$$

with the parameter $\bar{N} = N \text{tr } \mathbf{1}$, where the last quantity is the identity matrix in the γ -matrix algebra. In two dimensions $\bar{N} = 2N$ and the model is asymptotically free; for $\bar{N} = 2$ the interaction is proportional to the Thirring one and the β function vanishes identically.

As for the case of the non-linear σ -model, the solution of the renormalization group equations involves an invariant scale, which can be obtained (up to a constant which cannot be determined from perturbation theory alone) by integrating Eq. (255)

$$\xi^{-1}(u) = m(u) = \text{const. } \Lambda \exp \left[- \int^u \frac{du'}{\beta(u')} \right] \quad (256)$$

In two dimensions this scale is, to lowest order in u , proportional to

$$m(u) \underset{u \rightarrow 0}{\sim} \Lambda \exp \left[- \frac{2\pi}{(\bar{N} - 2)u} \right] \quad (257)$$

and thus non-analytic in the bare coupling u . Above two dimensions a non-trivial ultraviolet fixed point appears at

$$u_c = \frac{2\pi}{\bar{N} - 2} \epsilon + \frac{2\pi}{(\bar{N} - 2)^2} \epsilon^2 + \frac{(\bar{N} + 1)\pi}{2(\bar{N} - 2)^3} \epsilon^3 + \dots \quad (258)$$

In the weak coupling phase $u < u_c$ the fermions stay massless and chiral symmetry is unbroken, whereas in the strong coupling phase $u > u_c$ (which is the only phase present in $d = 2$) chiral symmetry is broken, a fermion condensate arises and a non-perturbative fermion mass is generated.

In the vicinity of the ultraviolet fixed point one has for the mass gap

$$m(u) \underset{u \rightarrow u_c}{\sim} \Lambda (u - u_c)^\nu \quad (259)$$

up to a constant of proportionality, with the exponent ν given by

$$\nu^{-1} \equiv -\beta'(u_c) = \epsilon - \frac{\epsilon^2}{\bar{N} - 2} - \frac{(\bar{N} - 3)\pi}{2(\bar{N} - 2)^2} \epsilon^3 + \dots \quad (260)$$

The rest of the analysis proceeds in a way that, at least formally, is virtually identical to the non-linear σ -model case. It need not be repeated here, as one can just take over the relevant formulas for the renormalization group behavior of n -point functions, for the running of the couplings etc.

The existence of a non-trivial ultraviolet fixed point implies that the large momentum behavior above two dimensions is not given by naive perturbation theory; it is given instead by the critical behavior of the renormalized theory, in accordance with Eq. (255). In the weak coupling, small u phase the scale m can be regarded as a crossover scale between the free field behavior at large distance scales and the critical behavior which sets in at large momenta.

Finally, the same model can be solved exactly in the large N limit. There too one can show that the model is characterized by two phases, a weak coupling phase where the fermions are massless and a strong coupling phase in which a chiral symmetry is spontaneously broken.

4. The Gravitational Case

In two dimensions the gravitational coupling becomes dimensionless, $G \sim \Lambda^{2-d}$, and the theory appears perturbatively renormalizable. In spite of the fact that the gravitational action reduces to a topological invariant in two dimensions, it would seem meaningful to try to construct, in analogy to what was suggested originally for scalar field theories (Wilson, 1973), the theory perturbatively as a double series in $\epsilon = d - 2$ and G .

One first notices though that in pure Einstein gravity, with Lagrangian density

$$\mathcal{L} = -\frac{1}{16\pi G_0} \sqrt{g} R , \quad (261)$$

the bare coupling G_0 can be completely reabsorbed by a field redefinition

$$g_{\mu\nu} = \omega g'_{\mu\nu} \quad (262)$$

with ω is a constant, and thus the renormalization properties of G_0 have no physical meaning for this theory. This simply follows from the fact that $\sqrt{g}R$ is homogeneous in $g_{\mu\nu}$, which is quite different from the Yang-Mills case. The situation changes though when one introduces a second dimensionful quantity to compare to. In the pure gravity case this contribution is naturally supplied by the cosmological constant term proportional to λ_0 ,

$$\mathcal{L} = -\frac{1}{16\pi G_0} \sqrt{g} R + \lambda_0 \sqrt{g} \quad (263)$$

Under a rescaling of the metric as in Eq. (262) one has

$$\mathcal{L} = -\frac{1}{16\pi G_0} \omega^{d/2-1} \sqrt{g'} R' + \lambda_0 \omega^{d/2} \sqrt{g'} \quad (264)$$

which is interpreted as a rescaling of the two bare couplings

$$G_0 \rightarrow \omega^{-d/2+1} G_0 , \quad \lambda_0 \rightarrow \lambda_0 \omega^{d/2} \quad (265)$$

leaving the dimensionless combination $G_0^d \lambda_0^{d-2}$ unchanged. Therefore only the latter combination has physical meaning in pure gravity. In particular, one can always choose the scale $\omega = \lambda_0^{-2/d}$ so as to adjust the volume term to have a unit coefficient. More importantly, it is physically meaningless to discuss separately the renormalization properties of G_0 and λ_0 , as they are individually gauge-dependent in the sense just illustrated. These arguments should clarify why in the following it will be sufficient at the end to just focus on the renormalization properties of one coupling, such as Newton's constant G_0 .

In general it is possible at least in principle to define quantum gravity in any $d > 2$. There are $d(d+1)/2$ independent components of the metric in d dimensions, and the same number of algebraically independent components of the Ricci tensor appearing in the field equations. The contracted Bianchi identities reduce the count by d , and so does general coordinate invariance, leaving $d(d-3)/2$ physical gravitational degrees of freedom in d dimensions. At the same time, four space-time dimensions is known to be the lowest dimension for which Ricci flatness does not imply the vanishing of the gravitational field, $R_{\mu\nu\lambda\sigma} = 0$, and therefore the first dimension to allow for gravitational waves and their quantum counterparts, gravitons.

In a general dimension the position space tree-level graviton propagator of the linearized theory, given in k -space in Eq. (62), can be obtained by Fourier transform and is proportional to

$$\int d^d k \frac{1}{k^2} e^{ik \cdot x} = \frac{\Gamma\left(\frac{d-2}{2}\right)}{4 \pi^{d/2} (x^2)^{d/2-1}}. \quad (266)$$

The static gravitational potential is then proportional to the spatial Fourier transform

$$V(r) \propto \int d^{d-1} \mathbf{k} \frac{e^{i\mathbf{k} \cdot \mathbf{x}}}{\mathbf{k}^2} \sim \frac{1}{r^{d-3}} \quad (267)$$

and can be shown to vanish in $d = 3$. To show this one needs to compute the analog of Eq. (17) in d dimensions, which is

$$-\frac{\kappa^2}{2} \int d^d x \left[T_{\mu\nu} \square^{-1} T^{\mu\nu} - (d-2)^{-1} T_\mu{}^\mu \square^{-1} T_\nu{}^\nu \right] \rightarrow -\frac{d-3}{d-2} \frac{\kappa^2}{2} \int d^{d-1} x T^{00} \mathcal{G} T^{00}, \quad (268)$$

where the Green's function \mathcal{G} is the static limit of $1/\square$, and $\kappa^2 = 16\pi G$. The above result then shows that there are no Newtonian forces in $d=2+1$ dimensions (Deser, Jackiw and 't Hooft, 1984; Deser and Jackiw, 1984). The only fluctuations left in $3d$ are associated with the scalar curvature (Deser, Jackiw and Templeton, 1982).

The $2 + \epsilon$ expansion for pure gravity then proceeds as follows. First the gravitational part of the action

$$\mathcal{L} = -\frac{\mu^\epsilon}{16\pi G} \sqrt{g} R, \quad (269)$$

with G dimensionless and μ an arbitrary momentum scale, is expanded by setting

$$g_{\mu\nu} \rightarrow \bar{g}_{\mu\nu} = g_{\mu\nu} + h_{\mu\nu} \quad (270)$$

where $g_{\mu\nu}$ is the classical background field and $h_{\mu\nu}$ the small quantum fluctuation. The quantity \mathcal{L} in Eq. (269) is naturally identified with the bare Lagrangian, and the scale μ with a microscopic ultraviolet cutoff Λ , the inverse lattice spacing in a lattice formulation. Since the resulting perturbative expansion is generally reduced to the evaluation of Gaussian integrals, the original constraint (in the Euclidean theory)

$$\det g_{\mu\nu}(x) > 0 \quad (271)$$

is no longer enforced (the same is not true in the lattice regulated theory, where it plays an important role, see the discussion following Eq. (400)).

A gauge fixing term needs to be added, in the form of a background harmonic gauge condition,

$$\mathcal{L}_{gf} = \frac{1}{2}\alpha\sqrt{g}g_{\nu\rho}\left(\nabla_\mu h^{\mu\nu} - \frac{1}{2}\beta g^{\mu\nu}\partial_\mu h\right)\left(\nabla_\lambda h^{\lambda\rho} - \frac{1}{2}\beta g^{\lambda\rho}\partial_\lambda h\right) \quad (272)$$

with $h^{\mu\nu} = g^{\mu\alpha}g^{\nu\beta}h_{\alpha\beta}$, $h = g^{\mu\nu}h_{\mu\nu}$ and ∇_μ the covariant derivative with respect to the background metric $g_{\mu\nu}$. The gauge fixing term also gives rise to a Faddeev-Popov ghost contribution \mathcal{L}_{ghost} containing the ghost field ψ_μ , so that the total Lagrangian becomes $\mathcal{L} + \mathcal{L}_{gf} + \mathcal{L}_{ghost}$.

In a flat background, $g^{\mu\nu} = \delta^{\mu\nu}$, one obtains from the quadratic part of the Lagrangian of Eqs. (269) and (272) the following expression for the graviton propagator

$$\begin{aligned} \langle h_{\mu\nu}(k)h_{\alpha\beta}(-k) \rangle = & \\ & \frac{1}{k^2}(\delta_{\mu\alpha}\delta_{\nu\beta} + \delta_{\mu\beta}\delta_{\nu\alpha}) - \frac{2}{d-2}\frac{1}{k^2}\delta_{\mu\nu}\delta_{\alpha\beta} \\ & - \left(1 - \frac{1}{\alpha}\right)\frac{1}{k^4}(\delta_{\mu\alpha}k_\nu k_\beta + \delta_{\nu\alpha}k_\mu k_\beta + \delta_{\mu\beta}k_\nu k_\alpha + \delta_{\nu\beta}k_\mu k_\alpha) \\ & + \frac{1}{d-2}\frac{4(\beta-1)}{\beta-2}\frac{1}{k^4}(\delta_{\mu\nu}k_\alpha k_\beta + \delta_{\alpha\beta}k_\mu k_\nu) \\ & + \frac{4(1-\beta)}{(\beta-2)^2}\left[2 - \frac{3-\beta}{\alpha} - \frac{2(1-\beta)}{d-2}\right]\frac{1}{k^6}k_\nu k_\mu k_\alpha k_\beta \end{aligned} \quad (273)$$

Normally it would be convenient to choose a gauge $\alpha=\beta=1$, in which case only the first two terms for the graviton propagator survive. But here it might be advantageous to leave the two gauge parameters unspecified, so that one can later show explicitly the gauge independence of the final result. In particular the gauge parameter β is related to the gauge freedom associated with

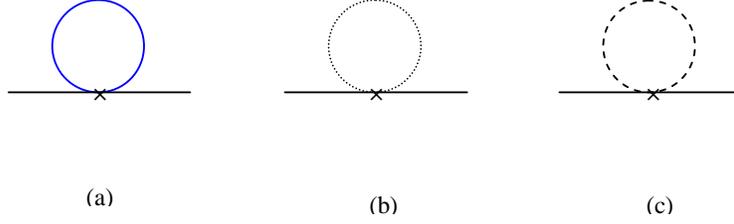


FIG. 6 One-loop diagrams giving rise to coupling renormalizations in gravity. From left to right, graviton loop, ghost loop and scalar matter loop.

the possibility, described above, of rescaling the metric $g_{\mu\nu}$. Note also the presence of kinematical poles in $\epsilon = d - 2$ in the second, fourth and fifth term for the graviton propagator.

To illustrate explicitly the mechanism of coupling renormalization, the cosmological term will be discussed first, since the procedure is a bit simpler. The cosmological term \sqrt{g} is first expanded by setting $g_{\mu\nu} = g_{\mu\nu} + h_{\mu\nu}$ with a flat background $g_{\mu\nu} = \delta_{\mu\nu}$. One has

$$\sqrt{g} = 1 + \frac{1}{2}h - \frac{1}{4}h_{\mu\nu}h^{\mu\nu} + \frac{1}{8}h^2 + O(h^3) \quad (274)$$

with $h = h^\mu{}_\mu$. Terms linear in the fluctuation $h_{\mu\nu}$ are dropped, since in a properly chosen background such terms are expected to be absent. The one-loop correction to the 1 term in the above expression is then given by the tadpole diagrams for the two quadratic terms,

$$-\frac{1}{4}h_{\mu\nu}h^{\mu\nu} + \frac{1}{8}h^2 \rightarrow -\frac{1}{4} \langle h_{\mu\nu}h^{\mu\nu} \rangle + \frac{1}{8} \langle h^2 \rangle \quad (275)$$

These are easily evaluated using the graviton propagator of Eq. (273). For the one loop divergences (see Fig. 6) associated with the \sqrt{g} term one then obtains

$$\lambda_0 \rightarrow \lambda_0 \left[1 - \left(\frac{a_1}{\epsilon} + \frac{a_2}{\epsilon^2} \right) G \right] \quad (276)$$

with coefficients

$$\begin{aligned} a_1 &= -\frac{8}{\alpha} + 8\frac{(\beta-1)^2}{(\beta-2)^2} + 4\frac{(\beta-1)(\beta-3)}{\alpha(\beta-2)^2} \\ a_2 &= 8\frac{(\beta-1)^2}{(\beta-2)^2} \end{aligned} \quad (277)$$

One notices that the kinematic singularities in the graviton propagator, proportional to $1/(d-2)$, can combine with the one loop ultraviolet divergent part of momentum integrals, as in

$$\frac{1}{\epsilon} \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2} \sim \frac{1}{\epsilon^2} \quad (278)$$

to give terms of order $1/\epsilon^2$ in Eq. (276). Generally it is better to separate the ultraviolet divergence from the infrared one, by using for example the following regulated integral

$$\int \frac{d^d k}{(2\pi)^d} \frac{1}{(k^2 + \mu^2)^a} = \frac{1}{(4\pi)^{d/2}} \frac{\Gamma(a - d/2)}{\Gamma(a)} (\mu^2)^{d/2 - a} \quad (279)$$

for $a = 1$ and $\mu \rightarrow 0$.

One can then follow the same procedure for the $\sqrt{g} R$ term. First one needs to expand the Einstein term to quadratic order in the quantum field $h_{\mu\nu}$

$$\begin{aligned} \sqrt{g} \bar{R} &= \sqrt{g} R \\ &+ \sqrt{g} \left\{ \frac{1}{4} \nabla_\rho h^\mu{}_\nu \nabla^\rho h^\nu{}_\mu - \frac{1}{2} \nabla_\nu h^\nu{}_\mu \nabla_\rho h^{\rho\mu} + \frac{1}{2} R^\sigma{}_{\rho\mu\nu} h^\rho{}_\sigma h^{\mu\nu} \right\} + \dots \end{aligned} \quad (280)$$

where ∇_μ denotes the covariant derivative with respect to the background metric $g_{\mu\nu}$. The complete expansion was given previously in Eq. (79). The same expansion then needs to be done for the gauge fixing term of Eq. (272) as well, and furthermore it is again convenient to choose as a background field the flat metric $g_{\mu\nu} = \delta_{\mu\nu}$. For the one loop divergences associated with the $\sqrt{g} R$ term one then finds

$$\frac{\mu^\epsilon}{16\pi G} \rightarrow \frac{\mu^\epsilon}{16\pi G} \left(1 - \frac{b}{\epsilon} G \right) \quad (281)$$

with the coefficient b given by (Gastmans et al 1977, Christensen et al 1978)

$$b = \frac{2}{3} \cdot 19 + \frac{4(\beta - 1)^2}{(\beta - 2)^2} \quad (282)$$

Thus the one-loop radiative corrections modify the total Lagrangian to

$$\mathcal{L} \rightarrow -\frac{\mu^\epsilon}{16\pi G} \left(1 - \frac{b}{\epsilon} G \right) \sqrt{g} R + \lambda_0 \left[1 - \left(\frac{a_1}{\epsilon} + \frac{a_2}{\epsilon^2} \right) G \right] \sqrt{g} \quad (283)$$

Next one can make use of the freedom to rescale the metric, by setting

$$\left[1 - \left(\frac{a_1}{\epsilon} + \frac{a_2}{\epsilon^2} \right) G \right] \sqrt{g} = \sqrt{g'} \quad (284)$$

which restores the original unit coefficient for the cosmological constant term. The rescaling is achieved by the following field redefinition

$$g_{\mu\nu} = \left[1 - \left(\frac{a_1}{\epsilon} + \frac{a_2}{\epsilon^2} \right) G \right]^{-2/d} g'_{\mu\nu} \quad (285)$$

Hence the cosmological term is brought back into the standard form $\lambda_0 \sqrt{g'}$, and one obtains for the complete Lagrangian to first order in G

$$\mathcal{L} \rightarrow -\frac{\mu^\epsilon}{16\pi G} \left[1 - \frac{1}{\epsilon} (b - \frac{1}{2} a_2) G \right] \sqrt{g'} R' + \lambda_0 \sqrt{g'} \quad (286)$$

where only terms singular in ϵ have been retained. From this last result one can finally read off the renormalization of Newton's constant

$$\frac{1}{G} \rightarrow \frac{1}{G} \left[1 - \frac{1}{\epsilon} (b - \frac{1}{2} a_2) G \right] \quad (287)$$

From Eqs. (277) and (282) one notices that the a_2 contribution cancels out the gauge-dependent part of b , giving for the remaining contribution $b - \frac{1}{2} a_2 = \frac{2}{3} \cdot 19$. Therefore the gauge dependence has, as one would have hoped on physical grounds, disappeared from the final answer. It is easy to see that the same result would have been obtained if the *scaled* cosmological constant $G\lambda_0$ had been held constant, instead of λ_0 as in Eq. (284). One important aspect of the result of Eq. (287) is that the quantum correction is negative, meaning that the strength of G is effectively increased by the lowest order radiative correction.

In the presence of an explicit renormalization scale parameter μ the β -function for pure gravity is obtained by requiring the independence of the quantity G_e (here identified as an effective coupling constant, with lowest order radiative corrections included) from the original renormalization scale μ ,

$$\begin{aligned} \mu \frac{d}{d\mu} G_e &= 0 \\ \frac{1}{G_e} &\equiv \frac{\mu^\epsilon}{G(\mu)} \left[1 - \frac{1}{\epsilon} (b - \frac{1}{2} a_2) G(\mu) \right] \end{aligned} \quad (288)$$

To zero-th order in G , the renormalization group β -function entering the renormalization group equation

$$\mu \frac{\partial}{\partial \mu} G = \beta(G) \quad (289)$$

is just given by

$$\beta(G) = \epsilon G + \dots \quad (290)$$

The above result just follows from the trivial scale dependence of the classical, dimensionful gravitational coupling: to achieve a fixed given G_e , the dimensionless quantity $G(\mu)$ itself has to scale like μ^ϵ . Next, to first order in G , one has from Eq. (288)

$$\mu \frac{\partial}{\partial \mu} G = \beta(G) = \epsilon G - \beta_0 G^2 + O(G^3, G^2 \epsilon) \quad (291)$$

with $\beta_0 = \frac{2}{3} \cdot 19$. From the procedure outlined above it is clear that G is the only coupling that is scale-dependent in pure gravity. As will be appreciated further below, the importance of the gravitational β -function $\beta(G)$ lies in the fact that it can be used either to determine the ultraviolet

cutoff dependence of the bare coupling needed to keep the effective coupling fixed (as in Eq. (288), or to determine the momentum dependence of the physical coupling $G(k)$ for a *fixed* cutoff.

Matter fields can be included as well. When N_S scalar fields and N_F Majorana fermion fields are added, the results of Eqs. (282), (287) and (288) are modified to

$$b \rightarrow b - \frac{2}{3}c \quad (292)$$

with $c = N_S + \frac{1}{2}N_F$ (the central charge of the Virasoro algebra in two dimensions), and therefore for the combined β -function of Eq. (291) to one-loop order one has $\beta_0 = \frac{2}{3}(19 - c)$. Of course one noteworthy aspect of the perturbative calculation is the appearance of a non-trivial ultraviolet fixed point at $G_c = (d - 2)/\beta_0$ for which $\beta(G_c) = 0$, whose physical significance will be discussed further below.

To check their consistency, the above one-loop calculations have been repeated by performing a number of natural variations. One modification consists in using the Thirring interaction

$$\mathcal{L} = -\frac{\mu^\epsilon}{16\pi G} \sqrt{g} R + e\bar{\psi}i\gamma^\mu D_\mu \psi - ke\bar{\psi}\gamma^a \psi \bar{\psi}\gamma_a \psi \quad (293)$$

instead of the cosmological constant term to set the scale for the metric. This results in a β -function still of the form of Eq. (291) but with $\beta_0 = \frac{2}{3}(25 - c)$. The slight discrepancy between the two results was initially attributed (Kawai and Ninomiya, 1990) to the well known problems related to the kinematic singularities of the graviton propagator in $2d$ discussed previously. To address this issue, a new perturbative expansion can be performed with the metric parametrized as

$$g_{\mu\nu} \rightarrow \bar{g}_{\mu\nu} = g_{\mu\rho} \left(e^h\right)^\rho{}_\nu e^{-\phi} \quad (294)$$

where the conformal mode ϕ (responsible for the kinematic singularity in the second term on the r.h.s. of Eq. (273)) is explicitly separated out, and $h^\mu{}_\nu$ is now taken to be traceless, $h^\mu{}_\mu = 0$. Furthermore the conformal mode is made massive by adding a cosmological constant term $\lambda_0\sqrt{g}$, which again acts as an infrared regulator. Repeating the calculation for the one loop divergences (Kawai, Kitazawa and Ninomiya, 1993) one now finds $\beta_0 = \frac{2}{3}(25 - c)$ which is consistent with the above quoted Thirring result ⁶.

⁶ For a while there was considerable uncertainty about the magnitude of the graviton contribution to β_0 , which was quoted originally as $38/3$ (Tsao 1977), later as $2/3$ (Gastman et al 1977, Christensen and Duff 1978; Weinberg, 1977), and more recently as $50/3$ (Kawai, Kitazawa and Ninomiya, 1992). As discussed in (Weinberg, 1977), the original expectation was that the graviton contribution should be $d(d - 3)/2 = -1$ times the scalar contribution close to $d = 2$, which would suggest for gravity the value $2/3$. Direct numerical estimates of the scaling exponent ν in the lattice theory for $d = 3$ (Hamber and Williams, 1993) give, using Eq. (308), a value $\beta_0 \approx 44/3$ and are therefore in much better agreement with the larger, more recent values.

In the meantime the calculations have been laboriously extended to two loops (Aida and Kitazawa, 1997), with the result

$$\mu \frac{\partial}{\partial \mu} G = \beta(G) = \epsilon G - \beta_0 G^2 - \beta_1 G^3 + O(G^4, G^3 \epsilon, G^2 \epsilon^2) \quad (295)$$

with $\beta_0 = \frac{2}{3} (25 - c)$ and $\beta_1 = \frac{20}{3} (25 - c)$.

5. Phases of Gravity in $2 + \epsilon$ Dimensions

The gravitational β -function of Eqs. (291) and (295) determines the scale dependence of Newton's constant G for d close to two. It has the general shape shown in Fig. 7. Because one is left, for the reasons described above, with a single coupling constant in the pure gravity case, the discussion becomes in fact quite similar to the non-linear σ -model case.

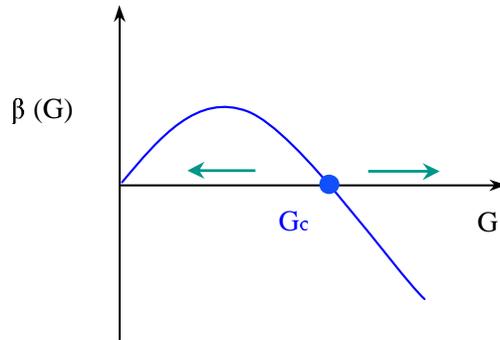


FIG. 7 The β -function for gravity in $2 + \epsilon$ dimensions. The arrows indicate the coupling constant flow as one approaches increasingly larger distance scales.

For a qualitative discussion of the physics it will be simpler in the following to just focus on the one loop result of Eq. (291); the inclusion of the two-loop correction does not alter the qualitative conclusions by much, as it has the same sign as the lower order, one-loop term. Depending on whether one is on the right ($G > G_c$) or on the left ($G < G_c$) of the non-trivial ultraviolet fixed point at

$$G_c = \frac{d-2}{\beta_0} + O((d-2)^2) \quad (296)$$

(with G_c positive provided one has $c < 25$) the coupling will either flow to increasingly larger values of G , or flow towards the Gaussian fixed point at $G = 0$, respectively. In the following we will refer to the two phases as the strong coupling and weak coupling phase, respectively. Perturbatively

one only has control on the small G regime. When one then sets $d = 2$ only the strong coupling phase survives, so two-dimensional gravity is always strongly coupled within this picture.

The running of G as a function of a sliding momentum scale $\mu = k$ in pure gravity can be obtained by integrating Eq. (291), and one finds

$$G(k^2) = \frac{G_c}{1 \pm a_0 (m^2/k^2)^{(d-2)/2}} \quad (297)$$

with a_0 a positive constant and m a mass scale. The choice of $+$ or $-$ sign is determined from whether one is to the left ($+$), or to right ($-$) of G_c , in which case the effective $G(k^2)$ decreases or, respectively, increases as one flows away from the ultraviolet fixed point towards lower momenta, or larger distances. Physically the two solutions represent a screening ($G < G_c$) and an anti-screening ($G > G_c$) situation. The renormalization group invariant mass scale $\sim m$ arises here as an arbitrary integration constant of the renormalization group equations (one could have absorbed the constant a_0 in m , but we will not do so here for reasons that will become clearer later).

While in the continuum perturbative calculation both phases, and therefore both signs, seem acceptable, the Euclidean lattice results on the other hand seem to rule out the weak coupling phase as pathological, in the sense that the lattice collapses into a two-dimensional branched polymer, as will be discussed later in this review in Sec. IV.A.

At the opposite end, at energies sufficiently high to become comparable to the ultraviolet cutoff (the inverse lattice spacing in a lattice discretization), the gravitational coupling G flows towards the ultraviolet fixed point,

$$G(k^2) \underset{k^2 \rightarrow \Lambda^2}{\sim} G(\Lambda) \quad (298)$$

where $G(\Lambda)$ is the coupling at the cutoff scale Λ , to be identified with the bare (or lattice) coupling. Note that it would seem meaningless to consider, within this framework, momenta which are larger than the ultraviolet cutoff Λ . At such energies higher dimension operators (such as higher derivative curvature terms) are expected to become important and should therefore be included in the microscopic action.

Note that the result of Eq. (297) is quite different from the naive expectation based on straight perturbation theory in $d > 2$ dimensions (where the theory is not perturbatively renormalizable)

$$\frac{G(k^2)}{G} \sim 1 + \text{const. } G k^{d-2} + O(G^2) \quad (299)$$

which gives a much worse ultraviolet behavior. The existence of a non-trivial ultraviolet fixed point alters the naive picture and drastically improves the ultraviolet behavior.

The k^2 -dependent contribution in the denominator of Eq. (297) is the quantum correction, which at least within a perturbative framework is assumed to be small. In the weak coupling phase $G < G_c$ the coupling then flows towards the origin corresponding to a gravitational screening solution, which sounds a bit odd as one would not expect gravity to be screened. On the other hand the infrared growth of the coupling in the strong coupling phase $G > G_c$ can be written equivalently as

$$G(k^2) \simeq G_c \left[1 + a_0 \left(\frac{m^2}{k^2} \right)^{(d-2)/2} + \dots \right] \quad (300)$$

where the dots indicate higher order radiative corrections, and which exhibits a number of interesting features. Firstly the fractional power suggests new non-trivial gravitational scaling dimensions, just as in the case of the non-linear σ -model. Furthermore, the quantum correction involves a new physical, renormalization group invariant scale $\xi = 1/m$ which cannot be fixed perturbatively, and whose size determines the scale for the quantum effects. In terms of the bare coupling $G(\Lambda)$, it is given by

$$m = A_m \cdot \Lambda \exp \left(- \int^{G(\Lambda)} \frac{dG'}{\beta(G')} \right) \quad (301)$$

which just follows from integrating $\mu \frac{\partial}{\partial \mu} G = \beta(G)$ and then setting as the arbitrary scale $\mu \rightarrow \Lambda$. Conversely, since m is an invariant, one has $\Lambda \frac{d}{d\Lambda} m = 0$; the running of $G(\mu)$ in accordance with the renormalization group equation of Eq. (289) ensures that the l.h.s. is indeed a renormalization group invariant. The constant A_m on the r.h.s. of Eq. (301) cannot be determined perturbatively, it needs to be computed by non-perturbative (lattice) methods, for example by evaluating invariant correlations at fixed geodesic distances. It is related to the constant a_0 in Eq. (300) by $a_0 = 1/(A_m^{1/\nu} G_c)$.

At the fixed point $G = G_c$ the theory is scale invariant by definition. In statistical field theory language the fixed point corresponds to a phase transition, where the correlation length $\xi = 1/m$ diverges and the theory becomes scale (conformally) invariant. In general in the vicinity of the fixed point, for which $\beta(G) = 0$, one can write

$$\beta(G) \underset{G \rightarrow G_c}{\sim} \beta'(G_c) (G - G_c) + O((G - G_c)^2) \quad (302)$$

If one then defines the exponent ν by

$$\beta'(G_c) = -1/\nu \quad (303)$$

then from Eq. (301) one has by integration in the vicinity of the fixed point

$$m \underset{G \rightarrow G_c}{\sim} \Lambda \cdot A_m |G(\Lambda) - G_c|^\nu . \quad (304)$$

which is why ν is often referred to as the mass gap exponent. Solving the above equation (with $\Lambda \rightarrow k$) for $G(k)$ one obtains back Eq. (300), with the constant a_0 there related to A_m in Eq. (304) by $a_0 = 1/(A_m^{1/\nu} G_c)$ and $\nu = 1/(d - 2)$.

That m is a renormalization group invariant is seen from

$$\mu \frac{d}{d\mu} m \equiv \mu \frac{d}{d\mu} [A_m \mu |G(\mu) - G_c|^\nu] = 0 \quad (305)$$

provided G runs in accordance with Eq. (300). To one-loop order one has from Eqs. (291) and (303) $\nu = 1/(d - 2)$. When the bare (lattice) coupling $G(\Lambda) = G_c$ one has achieved criticality, $m = 0$. How far the bare theory is from the critical point is determined by the choice of $G(\Lambda)$, the distance from criticality being measured by the deviation $\Delta G = G(\Lambda) - G_c$.

Furthermore Eq. (304) shows how the (lattice) continuum limit is to be taken. In order to reach the continuum limit $a \equiv 1/\Lambda \rightarrow 0$ for fixed physical correlation $\xi = 1/m$, the bare coupling $G(\Lambda)$ needs to be tuned so as to approach the ultraviolet fixed point at G_c ,

$$\Lambda \rightarrow \infty , \quad m \text{ fixed} , \quad G \rightarrow G_c . \quad (306)$$

The fixed point at G_c thus plays a central role in the cutoff theory: together with the universal scaling exponent ν it determines the correct unique quantum continuum limit in the presence of an ultraviolet cutoff Λ . Sometimes it can be convenient to measure all quantities in units of the cutoff and set $\Lambda = 1/a = 1$. In this case the quantity m measured in units of the cutoff (i.e. m/Λ) has to be tuned to zero in order to construct the lattice continuum limit: for a fixed lattice cutoff, the continuum limit is approached by tuning the bare lattice $G(\Lambda)$ to G_c . In other words, the lattice continuum limit has to be taken in the vicinity of the non-trivial ultraviolet point.

The discussion given above is not altered significantly, at least in its qualitative aspects, by the inclusion of the two-loop correction of Eq. (295). From the expression for the two-loop β -function

$$\mu \frac{\partial}{\partial \mu} G = \beta(G) = \epsilon G - \frac{2}{3} (25 - c) G^2 - \frac{20}{3} (25 - c) G^3 + \dots , \quad (307)$$

for c massless real scalar fields minimally coupled to gravity, one computes the roots $\beta(G_c) = 0$ to obtain the location of the ultraviolet fixed point, and from it on can then determine the universal exponent $\nu = -1/\beta'(G_c)$. One finds

$$\begin{aligned} G_c &= \frac{3}{2(25 - c)} \epsilon - \frac{45}{2(25 - c)^2} \epsilon^2 + \dots \\ \nu^{-1} &= \epsilon + \frac{15}{25 - c} \epsilon^2 + \dots \end{aligned} \quad (308)$$

which gives, for pure gravity without matter ($c = 0$) in four dimensions, to lowest order $\nu^{-1} = 2$, and $\nu^{-1} \approx 4.4$ at the next order.

Also, in general higher order corrections to the results of the linearized renormalization group equations of Eq. (302) are present, which affect the scaling away from the fixed point. Let us assume that close to the ultraviolet fixed point at G_c one can write for the β -function the following expansion

$$\beta(G) = -\frac{1}{\nu}(G - G_c) - \omega(G - G_c)^2 + \mathcal{O}((G - G_c)^3) , \quad (309)$$

After integrating $\mu \frac{\partial}{\partial \mu} G = \beta(G)$ as before, one finds for the structure of the correction to m [see for comparison Eq. (304)]

$$\left(\frac{m}{\Lambda}\right)^{1/\nu} = A_m \left[(G(\Lambda) - G_c) - \omega \nu (G(\Lambda) - G_c)^2 + \dots \right] . \quad (310)$$

The hope of course is that these corrections to scaling are small, ($\omega \ll 1$); in the vicinity of the fixed point the higher order term becomes unimportant when $|G - G_c| \ll 1/(\omega \nu)$. For the effective running coupling one then has

$$\frac{G(\mu)}{G_c} = 1 + a_0 \left(\frac{m}{\mu}\right)^{1/\nu} + a_0 \omega \nu \left(\frac{m}{\mu}\right)^{2/\nu} + \mathcal{O}\left(\left(\frac{m}{\mu}\right)^{3/\nu}\right) . \quad (311)$$

which gives an estimate for the size of the modifications to Eq. (300).

Finally, as a word of caution, one should mention that in general the convergence properties of the $2+\epsilon$ expansion are not well understood. The poor convergence found in some better known cases is usually ascribed to the suspected existence of infrared renormalon-type singularities $\sim e^{-c/G}$ close to two dimensions, and which could possibly arise in gravity as well. At the quantitative level, the results of the $2 + \epsilon$ expansion for gravity therefore remain somewhat limited, and obtaining the three- or four-loop term still represents a daunting task. Nevertheless they provide, through Eqs. (300) and (304), an analytical insight into the scaling properties of quantum gravity close and above two dimensions, including the suggestion of a non-trivial phase structure and an estimate for the non-trivial universal scaling exponents (Eq. (308)). The key question raised by the perturbative calculations is therefore: what remains of the above phase transition in four dimensions, how are the two phases of gravity characterized there non-perturbatively, and what is the value of the exponent ν determining the running of G in the vicinity of the fixed point *in four dimensions*.

Finally we should mention that there are other continuum renormalization group methods which can be used to estimate the scaling exponents. An approach which is closely related to the $2 + \epsilon$ expansion for gravity is the derivation of approximate flow equations from the changes of the

Legendre effective action with respect to a suitably introduced infrared cutoff μ . The method can be regarded as a variation on Wilson's original momentum slicing technique for obtaining approximate renormalization group equations for lattice couplings. In the simplest case of a scalar field theory (Morris, 1994) one starts from the partition function

$$\exp(W[J]) = \int [d\phi] \exp \left\{ -\frac{1}{2} \phi \cdot C^{-1} \cdot \phi - I_\Lambda[\phi] + J \cdot \phi \right\} \quad (312)$$

The $C \equiv C(k, \mu)$ term is taken to be an 'additive infrared cutoff term'. For it to be an infrared cutoff it needs to be small for $k < \mu$, ideally tending to zero as $k \rightarrow 0$, and such that $k^2 C(k, \mu)$ is large when $k > \mu$. Since the method is only ultimately applied to the vicinity of the fixed point, for which all physical relevant scales are much smaller than the ultraviolet cutoff Λ , it is argued that the specific nature of this cutoff is not really relevant in the following. Taking a derivative of $W[J]$ with respect to the scale μ gives

$$\frac{\partial W[J]}{\partial \mu} = -\frac{1}{2} \left[\frac{\delta W}{\delta J} \cdot \frac{\partial C^{-1}}{\partial \mu} \cdot \frac{\delta W}{\delta J} + \text{tr} \left(\frac{\partial C^{-1}}{\partial \mu} \frac{\delta^2 W}{\delta J \delta J} \right) \right] \quad (313)$$

which can be re-written in terms of the Legendre transform $\Gamma[\phi] = -W[J] - \frac{1}{2} \phi \cdot C^{-1} \cdot \phi + J \cdot \phi$ as

$$\frac{\partial \Gamma[\phi]}{\partial \mu} = -\frac{1}{2} \text{tr} \left[\frac{1}{C} \frac{\partial C}{\partial \mu} \cdot \left(1 + C \cdot \frac{\delta^2 \Gamma}{\delta \phi \delta \phi} \right)^{-1} \right] \quad (314)$$

where now $\phi = \delta W / \delta J$ is regarded as the classical field. The traces can then be simplified by writing them in momentum space. What remains to be done is first settle on a suitable cutoff function $C(k, \mu)$, and subsequently compute the effective action $\Gamma[\phi]$ in a derivative expansion, thus involving terms of the type $\partial^n \phi^m$, with μ dependent coefficients.

As far as the cutoff function is concerned, it is first written as $C(k, \mu) = \mu^{\eta-2} C(k^2/\mu^2)$ so as to include the expected anomalous dimensions of the ϕ propagator. To simplify things further, it is then assumed for the remaining function of a single variables that $C(q^2) = q^{2p}$ with p a non-negative integer (Morris, 1994). The subsequent derivative expansion gives for example for the $O(N)$ model in $d = 3$ to lowest order $O(\partial^0)$ an anomalous dimensions $\eta = 0$ for all N , and $\nu = 0.73$ for $N = 2$. At the next order $O(\partial^2)$ in the derivative expansion the method gives $\nu = 0.65$, compared to the best theoretical and experimental value $\nu = 0.67$ (Morris and Turner 1997).

In the gravitational case one can proceed in a similar way. First the gravity analog of Eq. (314) is clearly

$$\frac{\partial \Gamma[g]}{\partial \mu} = -\frac{1}{2} \text{tr} \left[\frac{1}{C} \frac{\partial C}{\partial \mu} \cdot \left(1 + C \cdot \frac{\delta^2 \Gamma}{\delta g \delta g} \right)^{-1} \right] \quad (315)$$

where now $g_{\mu\nu} = \delta W / \delta J_{\mu\nu}$ corresponds to the classical metric. The effective action itself contains the Einstein and cosmological terms

$$\Gamma_\mu[g] = -\frac{1}{16\pi G(\mu)} \int d^d x \sqrt{g} [R(g) - 2\lambda(\mu)] + \dots \quad (316)$$

as well as gauge fixing and possibly higher derivative terms (Reuter 1998). After the addition of a background harmonic gauge fixing term with gauge parameter α , the choice of a suitable (scalar) cutoff function is required, $C^{-1}(k, \mu) = (\mu^2 - k^2)\theta(\mu^2 - k^2)$ (Litim 2004), which is inserted into

$$\int [dh] \exp \left\{ -\frac{1}{2} h \cdot C^{-1} \cdot h - I_\Lambda[g] + J \cdot h \right\} \quad (317)$$

Note that this added momentum-dependent cutoff term violates both the weak field general coordinate invariance [see for instance Eq. (11)], as well as the general rescaling invariance of Eq. (264).

The solution of the resulting renormalization group equation for the two couplings $G(\mu)$ and $\lambda(\mu)$ is then truncated to the Einstein and cosmological term, a procedure which is equivalent to the derivative expansion discussed previously. A nontrivial fixed point in both couplings (G^*, λ^*) is then found in four dimensions, with complex eigenvalues $\nu^{-1} = 1.667 \pm 4.308i$ for a gauge parameter choice $\alpha \rightarrow \infty$ [for general gauge parameter the exponents can vary by as much as seventy percent (Lauscher and Reuter 2002)]. In the special limit of vanishing cosmological constant the equations simplify further and one finds a trivial Gaussian fixed point at $G = 0$, as well as a non-trivial ultraviolet fixed point with $\nu^{-1} = 2d(d-2)/(d+2)$, which in $d = 4$ gives now $\nu^{-1} = 2.667$. So in spite of the apparent crudeness of the lowest order approximation, an ultraviolet fixed point similar to the one found in the $2 + \epsilon$ expansion is recovered.

6. Running of $\alpha(\mu)$ in *QED* and *QCD*

QED and *QCD* provide two invaluable illustrative cases where the running of the gauge coupling with energy is not only theoretically well understood, but also verified experimentally. This section is intended to provide analogies and distinctions between the two theories, in a way later suitable for a comparison with the gravitational case. Most of the results found in this section are well known (see, for example, Frampton, 2000), but the purpose here is to provide some contrast (and in some instances, a relationship) with the gravitational case.

In *QED* the non-relativistic static Coulomb potential is affected by the vacuum polarization contribution due to electrons (and positrons) of mass m . To lowest order in the fine structure constant, the contribution is from a single Feynman diagram involving a fermion loop. One finds

for the vacuum polarization contribution $\omega_R(\mathbf{k}^2)$ at small \mathbf{k}^2 the well known result (Itzykson and Zuber, 1980)

$$\frac{e^2}{\mathbf{k}^2} \rightarrow \frac{e^2}{\mathbf{k}^2[1 + \omega_R(\mathbf{k}^2)]} \sim \frac{e^2}{\mathbf{k}^2} \left[1 + \frac{\alpha}{15\pi} \frac{\mathbf{k}^2}{m^2} + O(\alpha^2) \right] \quad (318)$$

which, for a Coulomb potential with a charge centered at the origin of strength $-Ze$ leads to well-known Uehling δ -function correction

$$V(r) = \left(1 - \frac{\alpha}{15\pi} \frac{\Delta}{m^2} \right) \frac{-Ze^2}{4\pi r} = \frac{-Ze^2}{4\pi r} - \frac{\alpha}{15\pi} \frac{-Ze^2}{m^2} \delta^{(3)}(\mathbf{x}) \quad (319)$$

It is not necessary though to resort to the small- \mathbf{k}^2 approximation, and in general a static charge of strength e at the origin will give rise to a modified potential

$$\frac{e}{4\pi r} \rightarrow \frac{e}{4\pi r} Q(r) \quad (320)$$

with

$$Q(r) = 1 + \frac{\alpha}{3\pi} \ln \frac{1}{m^2 r^2} + \dots \quad m r \ll 1 \quad (321)$$

for small r , and

$$Q(r) = 1 + \frac{\alpha}{4\sqrt{\pi}} \frac{1}{(mr)^{3/2}} e^{-2mr} + \dots \quad m r \gg 1 \quad (322)$$

for large r . Here the normalization is such that the potential at infinity has $Q(\infty) = 1$ ⁷. The reason we have belabored this example is to show that the screening vacuum polarization contribution would have dramatic effects in QED if for some reason the particle running through the fermion loop diagram had a much smaller (or even close to zero) mass. There are two interesting aspects of the (one-loop) result of Eqs. (321) and (322). The first one is that the exponentially small size of the correction at large r is linked with the fact that the electron mass m_e is not too small: the range of the correction term is $\xi = 2\hbar/mc = 0.78 \times 10^{-10} \text{cm}$, but would have been much larger if the electron mass had been a lot smaller.

In *QCD* (and related Yang-Mills theories) radiative corrections are also known to alter significantly the behavior of the static potential at short distances. The changes in the potential are

⁷ The running of the fine structure constant has recently been verified experimentally at *LEP*. The scale dependence of the vacuum polarization effects gives a fine structure constant changing from $\alpha(0) \sim 1/137.036$ at atomic distances to about $\alpha(m_{Z^0}) \sim 1/128.978$ at energies comparable to the Z^0 boson mass, in good agreement with the theoretical renormalization group prediction.

best expressed in terms of the running strong coupling constant $\alpha_S(\mu)$, whose scale dependence is determined by the celebrated beta function of $SU(3)$ QCD with n_f light fermion flavors

$$\mu \frac{\partial \alpha_S}{\partial \mu} = 2\beta(\alpha_S) = -\frac{\beta_0}{2\pi} \alpha_S^2 - \frac{\beta_1}{4\pi^2} \alpha_S^3 - \frac{\beta_2}{64\pi^3} \alpha_S^4 - \dots \quad (323)$$

with $\beta_0 = 11 - \frac{2}{3}n_f$, $\beta_1 = 51 - \frac{19}{3}n_f$, and $\beta_2 = 2857 - \frac{5033}{9}n_f + \frac{325}{27}n_f^2$. The solution of the renormalization group equation Eq. (323) then gives for the running of $\alpha_S(\mu)$

$$\alpha_S(\mu) = \frac{4\pi}{\beta_0 \ln \mu^2 / \Lambda_{\overline{MS}}^2} \left[1 - \frac{2\beta_1}{\beta_0^2} \frac{\ln |\ln \mu^2 / \Lambda_{\overline{MS}}^2|}{\ln \mu^2 / \Lambda_{\overline{MS}}^2} + \dots \right] \quad (324)$$

(see Fig. 8). The non-perturbative scale $\Lambda_{\overline{MS}}$ appears as an integration constant of the renormalization group equations, and is therefore - by construction - scale independent. The physical value of $\Lambda_{\overline{MS}}$ cannot be fixed from perturbation theory alone, and needs to be determined by experiment, giving $\Lambda_{\overline{MS}} \simeq 220 MeV$.

In principle one can solve for $\Lambda_{\overline{MS}}$ in terms of the coupling at any scale, and in particular at the cutoff scale Λ , obtaining

$$\Lambda_{\overline{MS}} = \Lambda \exp \left(- \int^{\alpha_S(\Lambda)} \frac{d\alpha'_S}{2\beta(\alpha'_S)} \right) = \Lambda \left(\frac{\beta_0 \alpha_S(\Lambda)}{4\pi} \right)^{\beta_1/\beta_0^2} e^{-\frac{2\pi}{\beta_0 \alpha_S(\Lambda)}} [1 + O(\alpha_S(\Lambda))] \quad (325)$$

In lattice QCD this is usually taken as the definition of the running strong coupling constant $\alpha_S(\mu)$. It then leads to an effective potential between quarks and anti-quarks of the form

$$V(\mathbf{k}^2) = -\frac{4}{3} \frac{\alpha_S(\mathbf{k}^2)}{\mathbf{k}^2} \quad (326)$$

and the leading logarithmic correction makes the potential appear softer close to the origin, $V(r) \sim 1/(r \ln r)$.

When the QCD result is contrasted with the QED answer of Eqs. (318) and (319) it appears that the infrared small \mathbf{k}^2 singularity in Eqs. (326) is quite serious. An analogous conclusion is reached when examining Eqs. (324): the coupling strength $\alpha_S(\mathbf{k}^2)$ diverges in the infrared due to the singularity at $k^2 = 0$. In phenomenological approaches to low energy QCD (Richardson, 1979) the uncontrolled growth in $\alpha_S(\mathbf{k}^2)$ due to the spurious small- k^2 divergence is regulated by the dynamically generated QCD infrared cutoff $\Lambda_{\overline{MS}}$, which can then be shown to give a confining linear potential at large distances.

Not all physical properties can be computed reliably in weak coupling perturbation theory. In non-Abelian gauge theories a confining potential is found at strong coupling by examining the behavior of the Wilson loop (Wilson, 1973), defined for a large closed loop C as

$$W(C) = \langle \text{tr} \mathcal{P} \exp \left\{ ig \oint_C A_\mu(x) dx^\mu \right\} \rangle, \quad (327)$$

with $A_\mu \equiv t_a A_\mu^a$ and the t_a 's the group generators of $SU(N)$ in the fundamental representation. In the pure gauge theory at strong coupling, the leading contribution to the Wilson loop can be shown to follow an area law for sufficiently large loops

$$W(C) \underset{A \rightarrow \infty}{\sim} \exp(-A(C)/\xi^2) \quad (328)$$

where $A(C)$ is the minimal area spanned by the planar loop C . The quantity ξ is the gauge field correlation length, defined for example from the exponential decay of the Euclidean correlation function of two infinitesimal loops separated by a distance $|x|$,

$$G_\square(x) = \langle \text{tr} \mathcal{P} \exp\left\{ig \oint_{C_\epsilon} A_\mu(x') dx'^\mu\right\}(x) \text{tr} \mathcal{P} \exp\left\{ig \oint_{C_\epsilon} A_\mu(x'') dx''^\mu\right\}(0) \rangle_c, \quad (329)$$

Here the C_ϵ 's are two infinitesimal loops centered around x and 0 respectively, suitably defined on the lattice as elementary square loops, and for which one has at sufficiently large separations

$$G_\square(x) \underset{|x| \rightarrow \infty}{\sim} \exp(-|x|/\xi) \quad (330)$$

The inverse of the correlation length ξ corresponds to the lowest mass excitation in the gauge theory, the scalar glueball, $m_0 = 1/\xi$. Notice that since the glueball mass m_0 is expected to be proportional to the parameter $\Lambda_{\overline{MS}}$ of Eq. (325) for small g , it is non-analytic in the gauge coupling.

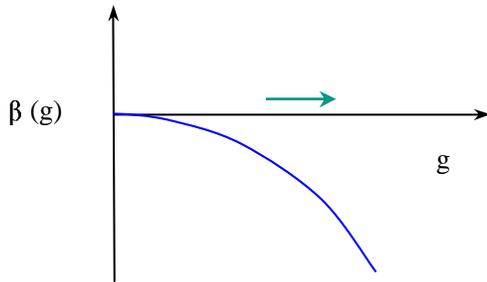


FIG. 8 The QCD β -function in four dimensions, with an ultraviolet stable fixed point at $g = 0$.

III. LATTICE REGULARIZED QUANTUM GRAVITY

The following sections are based on the lattice discretized description of gravity known as Regge calculus, where the Einstein theory is expressed in terms of a simplicial decomposition of space-time manifolds. Its use in quantum gravity is prompted by the desire to make use of techniques

developed in lattice gauge theories (Wilson, 1973)⁸, but with a lattice which reflects the structure of space-time rather than just providing a flat passive background (Regge, 1961). It also allows one to use powerful nonperturbative analytical techniques of statistical mechanics as well as numerical methods. A regularized lattice version of the continuum field theory is also usually perceived as a necessary prerequisite for a rigorous study of the latter.

In Regge gravity the infinite number of degrees of freedom in the continuum is restricted by considering Riemannian spaces described by only a finite number of variables, the geodesic distances between neighboring points. Such spaces are taken to be flat almost everywhere and are called *piecewise linear* (Singer and Thorpe, 1967). The elementary building blocks for d -dimensional space-time are *simplices* of dimension d . A 0-simplex is a point, a 1-simplex is an edge, a 2-simplex is a triangle, a 3-simplex is a tetrahedron. A d -simplex is a d -dimensional object with $d + 1$ vertices and $d(d + 1)/2$ edges connecting them. It has the important property that the values of its edge lengths specify the shape, and therefore the relative angles, uniquely.

A simplicial complex can be viewed as a set of simplices glued together in such a way that either two simplices are disjoint or they touch at a common face. The relative position of points on the lattice is thus completely specified by the *incidence matrix* (it tells which point is next to which) and the *edge lengths*, and this in turn induces a metric structure on the piecewise linear space. Finally the polyhedron constituting the union of all the simplices of dimension d is called a geometrical complex or *skeleton*. The transition from a smooth triangulation of a sphere to the corresponding secant approximation is illustrated in Fig. 9.

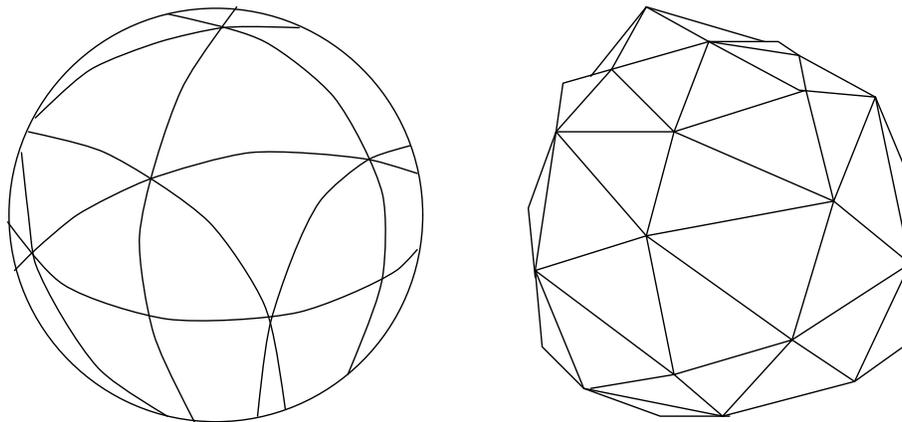


FIG. 9 Polydedral approximation to a sphere.

⁸ As an example of a state-of-the-art calculation of hadron properties in the lattice formulation of $SU(3)$ QCD see (Aoki et al., 2003).

A manifold can then be defined by its relationship to a piecewise linear space: a topological space is called a closed d -dimensional manifold if it is homeomorphic to a connected polyhedron, and furthermore, if its points possess neighborhoods which are homeomorphic to the interior of the d -dimensional sphere.

A. General Formulation

We will consider here a general simplicial lattice in d dimensions, made out of a collection of flat d -simplices glued together at their common faces so as to constitute a triangulation of a smooth continuum manifold, such as the d -torus or the surface of a sphere. If we focus on one such d -simplex, it will itself contain sub-simplices of smaller dimensions; as an example in four dimensions a given 4-simplex will contain 5 tetrahedra, 10 triangles (also referred to as hinges in four dimensions), 10 edges and 5 vertices. In general, an n -simplex will contain $\binom{n+1}{k+1}$ k -simplices in its boundary. It will be natural in the following to label simplices by the letter s , faces by f and hinges by h . A general connected, oriented simplicial manifold consisting of N_s d -simplices will also be characterized by an incidence matrix $I_{s,s'}$, whose matrix element $I_{s,s'}$ is chosen to be equal to one if the two simplices labeled by s and s' share a common face, and zero otherwise.

The geometry of the interior of a d -simplex is assumed to be flat, and is therefore completely specified by the lengths of its $d(d+1)/2$ edges. Let $x^\mu(i)$ be the μ -th coordinate of the i -th site. For each pair of neighboring sites i and j the link length squared is given by the usual expression

$$l_{ij}^2 = \eta_{\mu\nu} [x(i) - x(j)]^\mu [x(i) - x(j)]^\nu \quad (331)$$

with $\eta_{\mu\nu}$ the flat metric. It is therefore natural to associate, within a given simplex s , an edge vector $l_{ij}^\mu(s)$ with the edge connecting site i to site j .

1. Volumes and Angles

When focusing on one such n -simplex it will be convenient to label the vertices by $0, 1, 2, 3, \dots, n$ and denote the square edge lengths by $l_{01}^2 = l_{10}^2, \dots, l_{0n}^2$. The simplex can then be spanned by the set of n vectors e_1, \dots, e_n connecting the vertex 0 to the other vertices. To the remaining edges within the simplex one then assigns vectors $e_{ij} = e_i - e_j$ with $1 \leq i < j \leq n$. One has therefore n independent vectors, but $\frac{1}{2}n(n+1)$ invariants given by all the edge lengths squared within s .

In the interior of a given n -simplex one can also assign a second, orthonormal (Lorentz) frame, which we will denote in the following by $\Sigma(s)$. The expansion coefficients relating this orthonormal

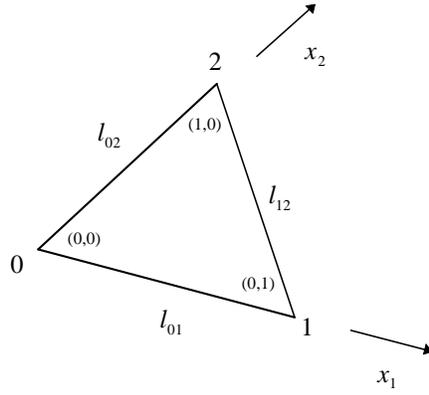


FIG. 10 Coordinates chosen along edges of a simplex, here a triangle.

frame to the one specified by the n directed edges of the simplex associated with the vectors e_i is the lattice analogue of the n -bein or tetrad e_μ^a .

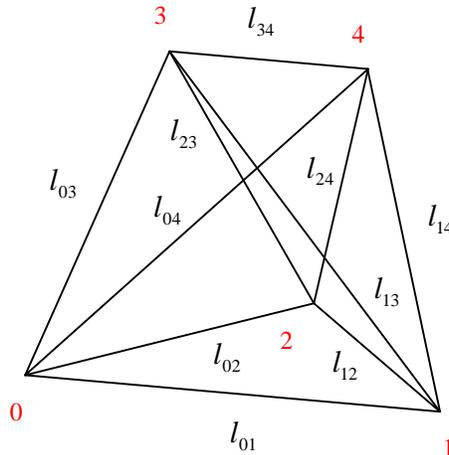


FIG. 11 A four-simplex, the four-dimensional analog of a tetrahedron. It contains five vertices, ten edges, ten triangles and five tetrahedra.

Within each n -simplex one can define a metric

$$g_{ij}(s) = e_i \cdot e_j , \tag{332}$$

with $1 \leq i, j \leq n$, and which in the Euclidean case is positive definite. In components one has

$g_{ij} = \eta_{ab} e_i^a e_j^b$. In terms of the edge lengths $l_{ij} = |e_i - e_j|$, the metric is given by

$$g_{ij}(s) = \frac{1}{2} \left(l_{0i}^2 + l_{0j}^2 - l_{ij}^2 \right) . \quad (333)$$

Comparison with the standard expression for the invariant interval $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ confirms that for the metric in Eq. (333) coordinates have been chosen along the n e_i directions.

The volume of a general n -simplex is given by the n -dimensional generalization of the well-known formula for a tetrahedron, namely

$$V_n(s) = \frac{1}{n!} \sqrt{\det g_{ij}(s)} . \quad (334)$$

An equivalent, but more symmetric, form for the volume of an n -simplex can be given in terms of the bordered determinant of an $(n+2) \times (n+2)$ matrix (Wheeler, 1964)

$$V_n(s) = \frac{(-1)^{\frac{n+1}{2}}}{n! 2^{n/2}} \begin{vmatrix} 0 & 1 & 1 & \dots \\ 1 & 0 & l_{01}^2 & \dots \\ 1 & l_{10}^2 & 0 & \dots \\ 1 & l_{20}^2 & l_{21}^2 & \dots \\ \dots & \dots & \dots & \dots \\ 1 & l_{n,0}^2 & l_{n,1}^2 & \dots \end{vmatrix}^{1/2} . \quad (335)$$

It is possible to associate p -forms with lower dimensional objects within a simplex, which will become useful later (Hartle, 1984). With each face f of an n -simplex (in the shape of a tetrahedron in four dimensions) one can associate a vector perpendicular to the face

$$\omega(f)_\alpha = \epsilon_{\alpha\beta_1\dots\beta_{n-1}} e_{(1)}^{\beta_1} \dots e_{(n-1)}^{\beta_{n-1}} \quad (336)$$

where $e_{(1)} \dots e_{(n-1)}$ are a set of oriented edges belonging to the face f , and $\epsilon_{\alpha_1\dots\alpha_n}$ is the sign of the permutation $(\alpha_1 \dots \alpha_n)$.

The volume of the face f is then given by

$$V_{n-1}(f) = \left(\sum_{\alpha=1}^n \omega_\alpha^2(f) \right)^{1/2} \quad (337)$$

Similarly, one can consider a hinge (a triangle in four dimensions) spanned by edges $e_{(1)}, \dots, e_{(n-2)}$. One defines the (un-normalized) hinge bivector

$$\omega(h)_{\alpha\beta} = \epsilon_{\alpha\beta\gamma_1\dots\gamma_{n-2}} e_{(1)}^{\gamma_1} \dots e_{(n-2)}^{\gamma_{n-2}} \quad (338)$$

with the area of the hinge then given by

$$V_{n-2}(h) = \frac{1}{(n-2)!} \left(\sum_{\alpha<\beta} \omega_{\alpha\beta}^2(h) \right)^{1/2} \quad (339)$$

Next, in order to introduce curvature, one needs to define the *dihedral angle* between faces in an n -simplex. In an n -simplex s two $n-1$ -simplices f and f' will intersect on a common $n-2$ -simplex h , and the dihedral angle at the specified hinge h is defined as

$$\cos \theta(f, f') = \frac{\omega(f)_{n-1} \cdot \omega(f')_{n-1}}{V_{n-1}(f) V_{n-1}(f')} \quad (340)$$

where the scalar product appearing on the r.h.s. can be re-written in terms of squared edge lengths using

$$\omega_n \cdot \omega'_n = \frac{1}{(n!)^2} \det(e_i \cdot e'_j) \quad (341)$$

and $e_i \cdot e'_j$ in turn expressed in terms of squared edge lengths by the use of Eq. (333). (Note that the dihedral angle θ would have to be defined as π minus the arccosine of the expression on the r.h.s. if the orientation for the e 's had been chosen in such a way that the ω 's would all point from the face f inward into the simplex s). As an example, in two dimensions and within a given triangle, two edges will intersect at a vertex, giving θ as the angle between the two edges. In three dimensions within a given simplex two triangles will intersect at a given edge, while in four dimension two tetrahedra will meet at a triangle. For the special case of an equilateral n -simplex, one has simply $\theta = \arccos \frac{1}{n}$. A related and often used formula for the sine of the dihedral angle θ is

$$\sin \theta(f, f') = \frac{n}{n-1} \frac{V_n(s) V_{n-2}(h)}{V_{n-1}(f) V_{n-1}(f')} \quad (342)$$

but is less useful for practical calculations, as the sine of the angle does not unambiguously determine the angle itself, which is needed in order to compute the local curvature.

In a piecewise linear space curvature is detected by going around elementary loops which are dual to a $(d-2)$ -dimensional subspace. From the dihedral angles associated with the faces of the simplices meeting at a given hinge h one can compute the *deficit angle* $\delta(h)$, defined as

$$\delta(h) = 2\pi - \sum_{s \supset h} \theta(s, h) \quad (343)$$

where the sum extends over all simplices s meeting on h . It then follows that the deficit angle δ is a measure of the curvature at h . The two-dimensional case is illustrated in Fig. 12, while the three- and four-dimensional cases are shown in Fig. 13.

2. Rotations, Parallel Transports and Voronoi Loops

Since the interior of each simplex s is assumed to be flat, one can assign to it a Lorentz frame $\Sigma(s)$. Furthermore inside s one can define a d -component vector $\phi(s) = (\phi_0 \dots \phi_{d-1})$. Under a

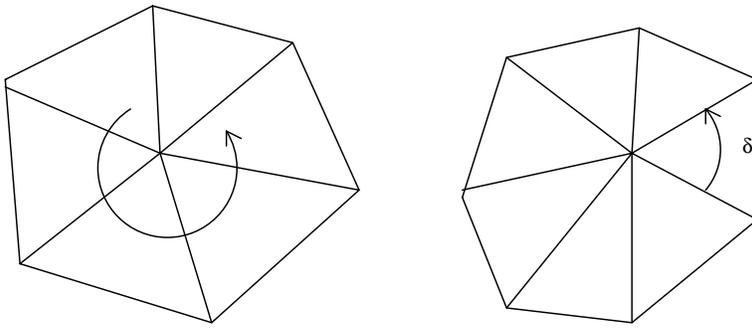


FIG. 12 Illustration of the deficit angle δ in two dimensions, where several flat triangles meet at a vertex.

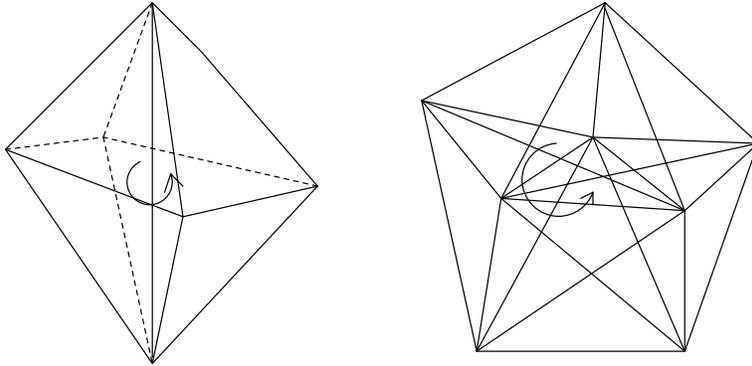


FIG. 13 Deficit angle in three dimensions (where several flat tetrahedra meet at an edge), and in four dimensions (where several flat four-simplices meet at a triangle).

Lorentz transformation of $\Sigma(s)$, described by the $d \times d$ matrix $\Lambda(s)$ satisfying the usual relation for Lorentz transformation matrices

$$\Lambda^T \eta \Lambda = \eta \quad (344)$$

the vector $\phi(s)$ will rotate to

$$\phi'(s) = \Lambda(s) \phi(s) \quad (345)$$

The base edge vectors $e_i^\mu = l_{0i}^\mu(s)$ themselves are of course an example of such a vector.

Next consider two d -simplices, individually labeled by s and s' , sharing a common face $f(s, s')$ of dimensionality $d - 1$. It will be convenient to label the d edges residing in the common face f by indices $i, j = 1 \dots d$. Within the first simplex s one can then assign a Lorentz frame $\Sigma(s)$, and similarly within the second s' one can assign the frame $\Sigma(s')$. The $\frac{1}{2}d(d - 1)$ edge vectors on the common interface $f(s, s')$ (corresponding physically to the same edges, viewed from two different coordinate systems) are expected to be related to each other by a Lorentz rotation \mathbf{R} ,

$$l_{ij}^\mu(s') = R^\mu_\nu(s', s) l_{ij}^\nu(s) \quad (346)$$

Under individual Lorentz rotations in s and s' one has of course a corresponding change in \mathbf{R} , namely $\mathbf{R} \rightarrow \Lambda(s') \mathbf{R}(s', s) \Lambda(s)$. In the Euclidean d -dimensional case \mathbf{R} is an orthogonal matrix, element of the group $SO(d)$.

In the absence of torsion, one can use the matrix $\mathbf{R}(s', s)$ to describes the parallel transport of any vector ϕ from simplex s to a neighboring simplex s' ,

$$\phi^\mu(s') = R^\mu{}_\nu(s', s) \phi^\nu(s) \quad (347)$$

\mathbf{R} therefore describes a lattice version of the connection (Lee, 1983). Indeed in the continuum such a rotation would be described by the matrix

$$R^\mu{}_\nu = \left(e^{\Gamma \cdot dx} \right)^\mu{}_\nu \quad (348)$$

with $\Gamma^\lambda_{\mu\nu}$ the affine connection. The coordinate increment dx is interpreted as joining the center of s to the center of s' , thereby intersecting the face $f(s, s')$. On the other hand, in terms of the Lorentz frames $\Sigma(s)$ and $\Sigma(s')$ defined within the two adjacent simplices, the rotation matrix is given instead by

$$R^a{}_b(s', s) = e^a{}_\mu(s') e^\nu{}_b(s) R^\mu{}_\nu(s', s) \quad (349)$$

(this last matrix reduces to the identity if the two orthonormal bases $\Sigma(s)$ and $\Sigma(s')$ are chosen to be the same, in which case the connection is simply given by $R(s', s)_\mu{}^\nu = e_\mu{}^a e^\nu{}_a$). Note that it is possible to choose coordinates so that $\mathbf{R}(s, s')$ is the unit matrix for one pair of simplices, but it will not then be unity for all other pairs.

This last set of results will be useful later when discussing lattice Fermions. Let us consider here briefly the problem of how to introduce lattice *spin rotations*. Given in d dimensions the above rotation matrix $\mathbf{R}(s', s)$, the spin connection $\mathbf{S}(s, s')$ between two neighboring simplices s and s' is defined as follows. Consider \mathbf{S} to be an element of the 2^ν -dimensional representation of the covering group of $SO(n)$, $Spin(d)$, with $d = 2\nu$ or $d = 2\nu + 1$, and for which S is a matrix of dimension $2^\nu \times 2^\nu$. Then \mathbf{R} can be written in general as

$$\mathbf{R} = \exp \left[\frac{1}{2} \sigma^{\alpha\beta} \theta_{\alpha\beta} \right] \quad (350)$$

where $\theta_{\alpha\beta}$ is an antisymmetric matrix The σ 's are $\frac{1}{2}d(d-1) d \times d$ matrices, generators of the Lorentz group ($SO(d)$ in the Euclidean case, and $SO(d-1, 1)$ in the Lorentzian case), whose explicit form is

$$[\sigma_{\alpha\beta}]^\gamma{}_\delta = \delta^\gamma{}_\alpha \eta_{\beta\delta} - \delta^\gamma{}_\beta \eta_{\alpha\delta} \quad (351)$$

so that, for example,

$$\sigma_{12} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (352)$$

For Fermions the corresponding spin rotation matrix is then obtained from

$$\mathbf{S} = \exp \left[\frac{i}{4} \gamma^{\alpha\beta} \theta_{\alpha\beta} \right] \quad (353)$$

with generators $\gamma^{\alpha\beta} = \frac{1}{2i}[\gamma^\alpha, \gamma^\beta]$, and with the Dirac matrices γ^α satisfying as usual $\gamma^\alpha\gamma^\beta + \gamma^\beta\gamma^\alpha = 2\eta^{\alpha\beta}$. Taking appropriate traces, one can obtain a direct relationship between the original rotation matrix $\mathbf{R}(s, s')$ and the corresponding spin rotation matrix $\mathbf{S}(s, s')$

$$R_{\alpha\beta} = \text{tr} \left(\mathbf{S}^\dagger \gamma_\alpha \mathbf{S} \gamma_\beta \right) / \text{tr} \mathbf{1} \quad (354)$$

which determines the spin rotation matrix up to a sign.

One can consider a sequence of rotations along an arbitrary path $P(s_1, \dots, s_{n+1})$ going through simplices $s_1 \dots s_{n+1}$, whose combined rotation matrix is given by

$$\mathbf{R}(P) = \mathbf{R}(s_{n+1}, s_n) \cdots \mathbf{R}(s_2, s_1) \quad (355)$$

and which describes the parallel transport of an arbitrary vector from the interior of simplex s_1 to the interior of simplex s_{n+1} ,

$$\phi^\mu(s_{n+1}) = R^\mu_\nu(P) \phi^\nu(s_1) \quad (356)$$

If the initial and final simplices s_{n+1} and s_1 coincide, one obtains a closed path $C(s_1, \dots, s_n)$, for which the associated expectation value can be considered as the gravitational analog of the Wilson loop. Its combined rotation is given by

$$\mathbf{R}(C) = \mathbf{R}(s_1, s_n) \cdots \mathbf{R}(s_2, s_1) \quad (357)$$

Under Lorentz transformations within each simplex s_i along the path one has a pairwise cancellation of the $\Lambda(s_i)$ matrices except at the endpoints, giving in the closed loop case

$$\mathbf{R}(C) \rightarrow \Lambda(s_1) \mathbf{R}(C) \Lambda^T(s_1) \quad (358)$$

Clearly the deviation of the matrix $\mathbf{R}(C)$ from unity is a measure of curvature. Also, the trace $\text{tr} \mathbf{R}(C)$ is independent of the choice of Lorentz frames.

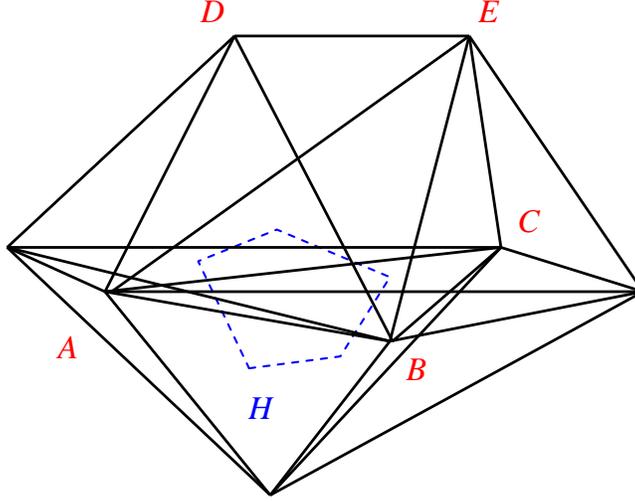


FIG. 14 Elementary polygonal path around a hinge (triangle) in four dimensions. The hinge ABC , contained in the simplex $ABCDE$, is encircled by the polygonal path H connecting the surrounding vertices which reside in the dual lattice. One such vertex is contained within the simplex $ABCDE$.

Of particular interest is the elementary loop associated with the smallest non-trivial, segmented parallel transport path one can build on the lattice. One such polygonal path in four dimensions is shown in Fig. 14. In general consider a $(d - 2)$ -dimensional simplex (hinge) h , which will be shared by a certain number m of d -simplices, sequentially labeled by $s_1 \dots s_m$, and whose common faces $f(s_1, s_2) \dots f(s_{m-1}, s_m)$ will also contain the hinge h . Thus in four dimensions several four-simplices will contain, and therefore encircle, a given triangle (hinge). In three dimensions the path will encircle an edge, while in two dimensions it will encircle a site. Thus for each hinge h there is a unique elementary closed path C_h for which one again can define the ordered product

$$\mathbf{R}(C_h) = \mathbf{R}(s_1, s_m) \cdots \mathbf{R}(s_2, s_1) \quad (359)$$

The hinge h , being geometrically an object of dimension $(d - 2)$, is naturally represented by a tensor of rank $(d - 2)$, referred to a coordinate system in h : an edge vector l_h^μ in $d = 3$, and an area bi-vector $\frac{1}{2}(l_h^\mu l_h'^\nu - l_h'^\nu l_h^\mu)$ in $d = 4$ etc. Following Eq. (338) it will therefore be convenient to define a hinge bi-vector U in any dimension as

$$U_{\mu\nu}(h) = \mathcal{N} \epsilon_{\mu\nu\alpha_1\alpha_{d-2}} l_{(1)}^{\alpha_1} \cdots l_{(d-2)}^{\alpha_{d-2}} \quad , \quad (360)$$

normalized, by the choice of the constant \mathcal{N} , in such a way that $U_{\mu\nu}U^{\mu\nu} = 2$. In four dimensions

$$U_{\mu\nu}(h) = \frac{1}{2A_h} \epsilon_{\mu\nu\alpha\beta} l_1^\alpha l_2^\beta \quad (361)$$

where $l_1(h)$ and $l_2(h)$ two independent edge vectors associated with the hinge h , and A_h the area of the hinge.

An important aspect related to the rotation of an arbitrary vector, when parallel transported around a hinge h , is the fact that, due to the hinge's intrinsic orientation, only components of the vector in the plane perpendicular to the hinge are affected. Since the direction of the hinge h is specified locally by the bivector $U_{\mu\nu}$ of Eq. (361), one can write for the loop rotation matrix \mathbf{R}

$$R^\mu{}_\nu(C) = \left(e^{\delta U} \right)^\mu{}_\nu \quad (362)$$

where C is now the small polygonal loop entangling the hinge h , and δ the deficit angle at h , previously defined in Eq. (343). One particularly noteworthy aspect of this last result is the fact that the area of the loop C does not enter in the expression for the rotation matrix, only the deficit angle and the hinge direction.

At the same time, in the continuum a vector V carried around an infinitesimal loop of area A_C will change by

$$\Delta V^\mu = \frac{1}{2} R^\mu{}_{\nu\lambda\sigma} A^{\lambda\sigma} V^\nu \quad (363)$$

where $A^{\lambda\sigma}$ is an area bivector in the plane of C , with squared magnitude $A_{\lambda\sigma} A^{\lambda\sigma} = 2A_C^2$. Since the change in the vector V is given by $\delta V^\alpha = (\mathbf{R} - \mathbf{1})^\alpha{}_\beta V^\beta$ one is led to the identification

$$\frac{1}{2} R^\alpha{}_{\beta\mu\nu} A^{\mu\nu} = (\mathbf{R} - \mathbf{1})^\alpha{}_\beta . \quad (364)$$

Thus the above change in V can equivalently be re-written in terms of the infinitesimal rotation matrix

$$R^\mu{}_\nu(C) = \left(e^{\frac{1}{2} R \cdot A} \right)^\mu{}_\nu \quad (365)$$

(where the Riemann tensor appearing in the exponent on the r.h.s. should not be confused with the rotation matrix \mathbf{R} on the l.h.s.).

It is then immediate to see that the two expressions for the rotation matrix \mathbf{R} in Eqs. (362) and (365) will be compatible provided one uses for the Riemann tensor at a hinge h the expression

$$R_{\mu\nu\lambda\sigma}(h) = \frac{\delta(h)}{A_C(h)} U_{\mu\nu}(h) U_{\lambda\sigma}(h) \quad (366)$$

expected to be valid in the limit of small curvatures, with $A_C(h)$ the area of the loop entangling the hinge h . Here use has been made of the geometric relationship $U_{\mu\nu} A^{\mu\nu} = 2A_C$. Note that the bivector U has been defined to be perpendicular to the $(d - 2)$ edge vectors spanning the hinge

h , and lies therefore in the same plane as the loop C . Furthermore, the expression of Eq. (366) for the *Riemann tensor at a hinge* has the correct algebraic symmetry properties, such as the antisymmetry in the first and second pair of indices, as well as the swap symmetry between first and second pair, and is linear in the curvature, with the correct dimensions of one over length squared.

The area A_C is most suitably defined by introducing the notion of a *dual lattice*, i.e. a lattice constructed by assigning centers to the simplices, with the polygonal curve C connecting these centers sequentially, and then assigning an area to the interior of this curve. One possible way of assigning such centers is by introducing perpendicular bisectors to the faces of a simplex, and locate the vertices of the dual lattice at their common intersection, a construction originally discussed in (Voronoi, 1908) and in (Meijering, 1953). Another, and perhaps even simpler, possibility is to use a barycentric subdivision (Singer and Thorpe, 1967).

3. Invariant Lattice Action

The first step in writing down an invariant lattice action, analogous to the continuum Einstein-Hilbert action, is to find the lattice analogue of the Ricci scalar. From the expression for the Riemann tensor at a hinge given in Eq. (366) one obtains by contraction

$$R(h) = 2 \frac{\delta(h)}{A_C(h)} \quad (367)$$

The continuum expression $\sqrt{g} R$ is then obtained by multiplication with the volume element $V(h)$ associated with a hinge. The latter is defined by first joining the vertices of the polyhedron C , whose vertices lie in the dual lattice, with the vertices of the hinge h , and then computing its volume.

By *defining* the polygonal area A_C as $A_C(h) = dV(h)/V^{(d-2)}(h)$, where $V^{(d-2)}(h)$ is the volume of the hinge (an area in four dimensions), one finally obtains for the Euclidean lattice action for pure gravity

$$I_R(l^2) = -k \sum_{\text{hinges } h} \delta(h) V^{(d-2)}(h) , \quad (368)$$

with the constant $k = 1/(8\pi G)$. One would have obtained the same result for the single-hinge contribution to the lattice action if one had contracted the infinitesimal form of the rotation matrix $R(h)$ in Eq. (362) with the hinge bivector $\omega_{\alpha\beta}$ of Eq. (338) (or equivalently with the bivector $U_{\alpha\beta}$ of Eq. (361) which differs from $\omega_{\alpha\beta}$ by a constant). The fact that the lattice action only involves

the content of the hinge $V^{(d-2)}(h)$ (the area of a triangle in four dimensions) is quite natural in view of the fact that the rotation matrix at a hinge in Eq. (362) only involves the deficit angle, and not the polygonal area $A_C(h)$.

An alternative form for the lattice action (Fröhlich, 1981) can be obtained instead by contracting the elementary rotation matrix $\mathbf{R}(C)$ of Eq. (362), and not just its infinitesimal form, with the hinge bivector of Eq. (338),

$$I_{\text{com}}(l^2) = -k \sum_{\text{hinges } h} \frac{1}{2} \omega_{\alpha\beta}(h) R^{\alpha\beta}(h) \quad (369)$$

The above construction can be regarded as analogous to Wilson's lattice gauge theory, for which the action also involves traces of products of $SU(N)$ color rotation matrices (Wilson, 1973). For small deficit angles one can of course use $\omega_{\alpha\beta} = (d-2)! V^{(d-2)} U_{\alpha\beta}$ to show the equivalence of the two lattice actions.

But in general, away from a situation of small curvatures, the two lattice action are not equivalent, as can be seen already in two dimensions. Writing the rotation matrix at a hinge as $\mathbf{R}(h) = \begin{pmatrix} \cos \delta & \sin \delta \\ -\sin \delta & \cos \delta \end{pmatrix}$, expressed for example in terms of Pauli matrices, and taking the appropriate trace ($\omega_{\alpha\beta} = \epsilon_{\alpha\beta}$ in two dimensions) one finds

$$\text{tr} \left[\frac{1}{2} (-i\sigma_y)(\cos \delta_p + i\sigma_y \sin \delta_p) \right] = \sin \delta_p \quad (370)$$

and therefore $I_{\text{com}} = -k \sum_p \sin \delta_p$. In general one can show that the compact action I_{com} in d dimensions involves the sine of the deficit angle, instead of just the angle itself as in the Regge case. In the weak field limit the two actions should lead to similar expansions, while away from the weak field limit one would have to verify that the same universal long distance properties are recovered.

The preceding observations can in fact be developed into a consistent first order (Palatini) formulation of Regge gravity, with suitably chosen independent transformation matrices and metrics, related to each other by a set of appropriate lattice equations of motion (Caselle, d'Adda and Magnea, 1986). Ultimately one would expect the first and second order formulations to describe the same quantum theory, with common universal long-distance properties. How to consistently define finite rotations, frames and connections in Regge gravity was first discussed systematically in (Fröhlich, 1981).

One important result that should be mentioned in this context is the rigorous proof of convergence in the sense of measures of the Regge lattice action towards the continuum Einstein-Hilbert action (Cheeger, Müller and Schrader, 1984). Some general aspects of this result have recently

been reviewed from a mathematical point of view in (Lafontaine, 1986). A derivation of the Regge action from its continuum counterpart was later presented in (Lee, Feinberg and Friedberg, 1984).

Other terms can be added to the lattice action. Consider for example a cosmological constant term, which in the continuum theory takes the form $\lambda_0 \int d^d x \sqrt{g}$. The expression for the cosmological constant term on the lattice involves the total volume of the simplicial complex. This may be written as

$$V_{\text{total}} = \sum_{\text{simplices } s} V_s \quad (371)$$

or equivalently as

$$V_{\text{total}} = \sum_{\text{hinges } h} V_h \quad (372)$$

where V_h is the volume associated with each hinge via the construction of a dual lattice, as described above. Thus one may regard the local volume element $\sqrt{g} d^d x$ as being represented by either V_h (centered on h) or V_s (centered on s).

The Regge and cosmological constant term then lead to the combined action

$$I_{\text{latt}}(l^2) = \lambda_0 \sum_{\text{simplices } s} V_s^{(d)} - \sum_{\text{hinges } h} \delta_h V_h^{(d-2)} \quad (373)$$

One would then write for the lattice regularized version of the Euclidean Feynman path integral

$$Z_{\text{latt}}(\lambda_0, k) = \int [dl^2] \exp\left(-I_{\text{latt}}(l^2)\right) \quad , \quad (374)$$

where $[dl^2]$ is an appropriate functional integration measure over the edge lengths, to be discussed later.

The structure of the gravitational action of Eq. (373) leads naturally to some rather general observations, which we will pursue here. The first, cosmological constant, term represents the total four-volume of space-time. As such, it does not contain any derivatives (or finite differences) of the metric and is completely local; it does not contribute to the propagation of gravitational degrees of freedom and is more akin to a mass term (as is already clear from the weak field expansion of $\int \sqrt{g}$ in the continuum). In an ensemble in which the total four-volume is fixed in the thermodynamic limit (number of simplices tending to infinity) one might in fact take the lattice coupling $\lambda_0 = 1$, since different values of λ_0 just correspond to a trivial rescaling of the overall four-volume (of course in a traditional renormalization group approach to field theory, the overall four-volume is always kept fixed while the scale or q^2 dependence of the action and couplings are investigated). Alternatively,

one might even want to choose directly an ensemble for which the probability distribution in the total four-volume V is

$$\mathcal{P}(V) \propto \delta(V - V_0) \quad (375)$$

in analogy with the microcanonical ensemble of statistical mechanics.

The second, curvature contribution to the action contains, as in the continuum, the proper kinetic term. This should already be clear from the derivation of the lattice action given above, and will be made even more explicit in the section dedicated to the lattice weak field expansion. Such a term now provides the necessary coupling between neighboring lattice metrics, but the coupling still remains local. Geometrically, it can be described as a sum of elementary loop contributions, as it contains as its primary ingredient the deficit angle associated with an elementary parallel transport loop around the hinge h . When $k = 0$ one resides in the extreme strong coupling regime. There the fluctuations in the metric are completely unconstrained by the action, insofar as only the total four-volume of the manifold is kept constant.

At this point it might be useful to examine some specific cases with regards to the overall dimensionality of the simplicial complex. In *two dimensions* the Regge action reduces to a sum over lattice sites p of the $2 - d$ deficit angle, giving the discrete analog of the Gauss-Bonnet theorem

$$\sum_{\text{sites } p} \delta_p = 2\pi\chi \quad (376)$$

where $\chi = 2 - 2g$ is the Euler characteristic of the surface, and g the genus (the number of handles). In this case the action is therefore a topological invariant, and the above lattice expression is therefore completely analogous to the well known continuum result

$$\frac{1}{2} \int d^2x \sqrt{g} R = 2\pi\chi \quad (377)$$

This remarkable identity ensures that two-dimensional lattice R -gravity is as trivial as in the continuum, since the variation of the local action density under a small variation of an edge length l_{ij} is still zero. Of course there is a much simpler formula for the Euler characteristic of a simplicial complex, namely

$$\chi = \sum_{i=0}^d (-1)^i N_i \quad (378)$$

where N_i is the number of simplices of dimension i . Also it should be noted that in two dimensions the compact action $I_{com} = \sum_p \sin \delta_p$ does not satisfy the Gauss-Bonnet relation.

In *three dimensions* the Regge lattice action reads

$$I_R = -k \sum_{\text{edges } h} l_h \delta_h \quad (379)$$

where δ_h is the deficit angle around the edge labeled by h . Variation with respect to an edge length l_h gives two terms, of which only the term involving the variation of the edge is non-zero

$$\delta I_R = -k \sum_{\text{edges } h} \delta l_h \cdot \delta_h. \quad (380)$$

In fact it was shown by Regge that for any $d > 2$ the term involving the variation of the deficit angle does not contribute to the equations of motion (just as in the continuum the variation of the Ricci tensor does not contribute to the equations of motion either). Therefore in three dimensions the lattice equations of motion, in the absence of sources and cosmological constant term, reduce to

$$\delta_h = 0 \quad (381)$$

implying that all deficit angles have to vanish, i.e. a flat space.

In four dimensions variation of I_R with respect to the edge lengths gives the simplicial analogue of Einstein's field equations, whose derivation is again, as mentioned, simplified by the fact that the contribution from the variation of the deficit angle is zero

$$\delta I_R = \sum_{\text{triangles } h} \delta(A_h) \cdot \delta_h \quad (382)$$

In the discrete case the field equations reduce therefore to

$$\lambda_0 \sum_s \frac{\partial V_s}{\partial l_{ij}} - k \sum_h \delta_h \frac{\partial A_h}{\partial l_{ij}} = 0 \quad (383)$$

and the derivatives can then be worked out for example from Eq. (335). Alternatively, a rather convenient and compact expression can be given (Hartle, 1984) for the derivative of the squared volume V_n^2 of an arbitrary n -simplex with respect to one of its squared edge lengths

$$\frac{\partial V_n^2}{\partial l_{ij}^2} = \frac{1}{n^2} \omega_{n-1} \cdot \omega'_{n-1} \quad (384)$$

where the $\omega_{(n-1)}^\alpha$'s (here referring to two $(n-1)$ simplices part of the same n -simplex) are given in Eqs. (336) and (341). In the above expression the ω 's are meant to be associated with vertex labels $0, \dots, i-1, i+1, \dots, n$ for ω_{n-1} , and $0, \dots, j-1, j+1, \dots, n$ for ω'_{n-1} respectively.

Then in the absence of a cosmological term one finds the remarkably simple expression for the lattice field equations

$$\frac{1}{2} l_p \sum_{h \supset l_p} \delta_h \cot \theta_{ph} = 0 \quad (385)$$

where the sum is over hinges (triangles) labeled by h meeting on the common edge l_p , and θ_{ph} is the angle in the hinge h opposite to the edge l_p . This is illustrated in Fig. 15.

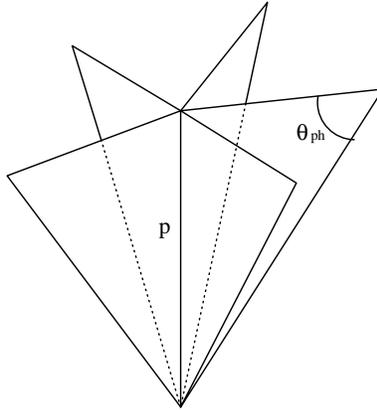


FIG. 15 Angles appearing in the Regge equations.

The discrete equations given above represent the lattice analogs of the Einstein field equations in a vacuum, for which suitable solutions can be searched for by adjusting the edge lengths. Since the equations are in general non-linear, the existence of multiple solutions cannot in general be ruled out (Misner, Thorne and Wheeler, 1973). A number of papers have addressed the general issue of convergence to the continuum in the framework of the classical formulation (Brewin and Gentle, 2001). Several authors have discussed non-trivial applications of the Regge equations to problems in classical general relativity such as the Schwarzschild and Reissner-Nordstrom geometries (Wong, 1971), the Friedmann and Tolman universes (Collins and Williams, 1974), and the problem of radial motion and circular (actually polygonal) orbits (Williams and Ellis, 1980). Spherically symmetric, as well as more generally inhomogeneous, vacuum spacetimes were studied using a discrete $3 + 1$ formulation with a variety of time-slicing prescriptions in (Porter, 1987), and later extended (Dubal, 1989) to a systematic investigation of the axis-symmetric non-rotating vacuum solutions and to the problem of relativistic spherical collapse for polytropic perfect fluids.

In classical gravity the general time evolution problem plays of course a central role. The $3 + 1$ time evolution problem in Regge gravity was discussed originally in (Sorkin, 1975) and later re-examined from a numerical, practical perspective in (Barrett, Galassi, Miller, Sorkin, Tuckey

and Williams, 1994) using a discrete time step formulation, whereas in (Piran and Williams, 1986) a continuous time formalism was proposed. The choice of lapse and shift functions in Regge gravity were discussed further in (Tuckey, 1989; Galassi, 1993) and in (Gentle and Miller, 1998), and applied to the Kasner cosmology in the last reference. An alternative so-called null-strut approach was proposed in (Miller and Wheeler, 1985) which builds up a spacelike-foliated spacetime with a maximal number of null edges, but seems difficult to implement in practice. Finally in (Khatsymovsky, 1991) and (Immirzi, 1996) a continuous time Regge gravity formalism in the tetrad-connection variables was developed, in part targeted towards quantum gravity calculations. A recent comprehensive review of classical applications of Regge gravity can be found for example in (Gentle, 2002), as well as a more complete set of references.

4. Lattice Diffeomorphism Invariance

Consider the two-dimensional flat skeleton shown in Fig. 16. It is clear that one can move around a point on the surface, keeping all the neighbors fixed, without violating the triangle inequalities and leave all curvature invariants unchanged.

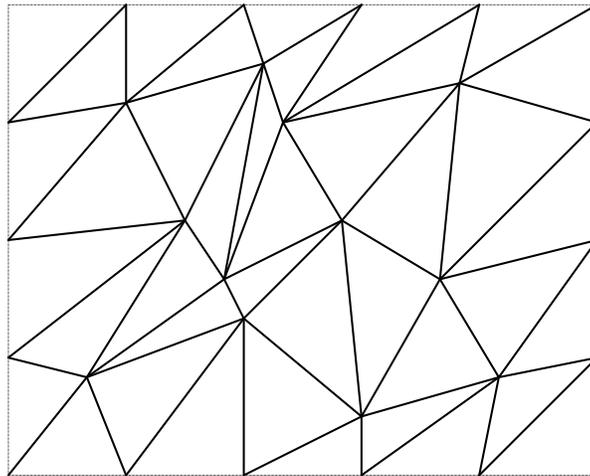


FIG. 16 On a random simplicial lattice there are in general no preferred directions.

In d dimensions this transformation has d parameters and is an exact invariance of the action. When space is slightly curved, the invariance is in general only an approximate one, even though for piecewise linear spaces piecewise diffeomorphisms can still be defined as the set of local motions

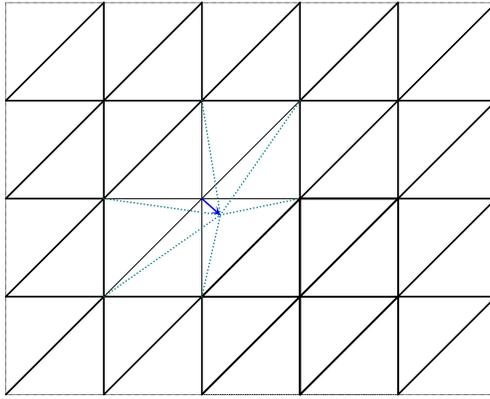


FIG. 17 Example of a lattice diffeomorphism, the local gauge transform of a flat lattice, corresponding to a d -parameter local deformation of the edges.

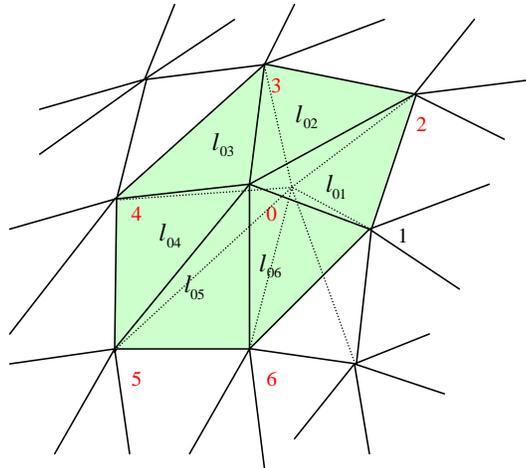


FIG. 18 Another example of a lattice diffeomorphism, the gauge deformation of a lattice around a vertex 0, leaving the local action contribution from that vertex invariant.

of points that leave the local contribution to the action, the measure and the lattice analogues of the continuum curvature invariants unchanged (Hamber and Williams, 1998). Note that in general the gauge deformations of the edges are still constrained by the triangle inequalities. The general situation is illustrated in Figs. 16, 17 and 18. In the limit when the number of edges becomes very large, the full continuum diffeomorphism group should be recovered.

In general the structure of lattice local gauge transformations is rather complicated and will not be given here; it can be found in the above quoted reference. These are defined as transformations acting locally on a given set of edges which leave the local lattice curvature invariant. The simplest

context in which this local invariance can be exhibited explicitly is the lattice weak field expansion, which will be discussed later in Sec. III.B.1. The local gauge invariance corresponding to continuum diffeomorphism is given there in Eq. (527). From the transformation properties of the edge lengths it is clear that their transformation properties are related to those of the local metric, as already suggested for example by the identification of Eqs. (333) and (398).

5. Lattice Bianchi Identities

Consider therefore a closed path C_h encircling a hinge h and passing through each of the simplices that meet at that hinge. In particular one may take C_h to be the boundary of the polyhedral dual (or Voronoi) area surrounding the hinge. We recall that the Voronoi polyhedron dual to a vertex P is the set of all points on the lattice which are closer to P than any other vertex; the corresponding new vertices then represent the sites on the dual lattice. A unique closed parallel transport path can then be assigned to each hinge, by suitably connecting sites in the dual lattice.

With each neighboring pair of simplices $s, s + 1$ one associates a Lorentz transformation $\mathbf{R}_{\beta}^{\alpha}(s, s + 1)$, which describes how a given vector V_{μ} transforms between the local coordinate systems in these two simplices. As discussed previously, the above transformation is directly related to the continuum path-ordered (P) exponential of the integral of the local affine connection $\Gamma_{\mu\nu}^{\lambda}$ via

$$R^{\mu}_{\nu} = \left[\mathcal{P} e^{\int_{\text{path between simplices}} \Gamma^{\lambda} dx_{\lambda}} \right]_{\nu}^{\mu} . \quad (386)$$

The connection here has support only on the common interface between the two simplices.

Just as in the continuum, where the affine connection and therefore the infinitesimal rotation matrix is determined by the metric and its first derivatives, on the lattice one expects that the elementary rotation matrix between simplices $\mathbf{R}_{s,s+1}$ is fixed by the difference between the g_{ij} 's of Eq. (333) within neighboring simplices.

For a vector V transported once around a Voronoi loop, i.e. a loop formed by Voronoi edges surrounding a chosen hinge, the change in the vector V is given by

$$\delta V^{\alpha} = (\mathbf{R} - \mathbf{1})^{\alpha}_{\beta} V^{\beta} , \quad (387)$$

where $\mathbf{R} \equiv \prod_s \mathbf{R}_{s,s+1}$ is now the total rotation matrix associated with the given hinge, given by

$$\left[\prod_s \mathbf{R}_{s,s+1} \right]_{\nu}^{\mu} = \left[e^{\delta(h)U(h)} \right]_{\nu}^{\mu} . \quad (388)$$

It is these lattice parallel transporters around closed elementary loops that satisfy the lattice analogues of the Bianchi identities. These are derived by considering paths which encircle more than one hinge and yet are topologically trivial, in the sense that they can be shrunk to a point without entangling any hinge (Regge, 1961).

Thus, for example, the ordered product of rotation matrices associated with the triangles meeting on a given edge has to give one, since a single path can be constructed which sequentially encircles all the triangles and is topologically trivial

$$\prod_{\substack{\text{hinges } h \\ \text{meeting on edge } p}} \left[e^{\delta(h)U(h)} \right]_{\nu}^{\mu} = 1 \quad (389)$$

Other identities might be derived by considering paths that encircle several hinges meeting on one point. Regge has shown that the above lattice relations correspond precisely to the continuum Bianchi identities. One can therefore explicitly construct exact lattice analogues of the continuum uncontracted, partially contracted, and fully contracted Bianchi identities (Hamber and Kagal, 2004). The lattice Bianchi identities are illustrated in Figs. 19 for the three-dimensional case, and 20 for the four-dimensional case.

The resulting lattice equations are quite similar in structure to the Bianchi identities in $SU(N)$ lattice gauge theories, where one considers identities arising from the multiplication of group elements associated with the square faces of a single cube part of a hypercubic lattice (Wilson, 1973). The motivation there was the possible replacement of the integration over the group elements by an integration over the “plaquette variables” associated with an elementary square (thereby involving the ordered product of four group elements), provided the Bianchi identity constraint is included as well in the lattice path integral.

6. Gravitational Wilson Loop

We have seen that with each neighboring pair of simplices $s, s + 1$ one can associate a Lorentz transformation $\mathbf{R}^{\mu}_{\nu}(s, s + 1)$, which describes how a given vector V^{μ} transforms between the local coordinate systems in these two simplices, and that the above transformation is directly related to the continuum path-ordered (P) exponential of the integral of the local affine connection $\Gamma^{\lambda}_{\mu\nu}(x)$ via

$$R^{\mu}_{\nu} = \left[P e^{\int_{\text{path between simplices}} \Gamma_{\lambda} dx^{\lambda}} \right]_{\nu}^{\mu} . \quad (390)$$

with the connection having support only on the common interface between the two simplices. Also, for a closed elementary path C_h encircling a hinge h and passing through each of the simplices that

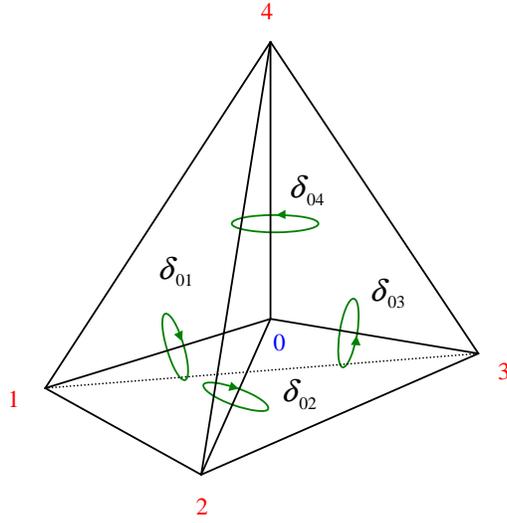


FIG. 19 Illustration of the lattice Bianchi identity in the case of three dimensions, where several hinges (edges) meet on a vertex. The combined rotation for a path that sequentially encircles several hinges and which can be shrunk to a point is given by the identity matrix.

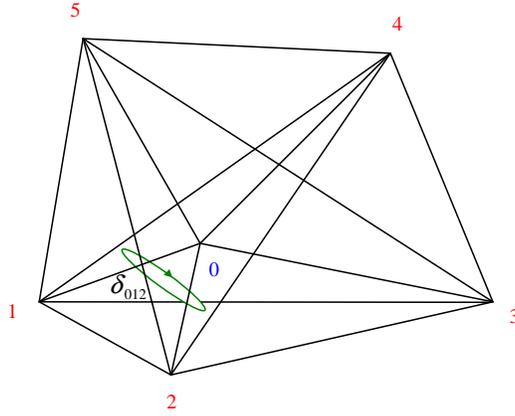


FIG. 20 Another illustration of the lattice Bianchi identity, now in four dimensions. Here several hinges (triangles) meet at the vertex labelled by 0. Around each hinge one has a corresponding rotation and therefore a deficit angle δ . The product of rotation matrices that sequentially encircle several hinges and is topologically trivial gives the identity matrix.

meet at that hinge one has for the total rotation matrix $\mathbf{R} \equiv \prod_s \mathbf{R}_{s,s+1}$ associated with the given hinge

$$\left[\prod_s \mathbf{R}_{s,s+1} \right]_{\nu}^{\mu} = \left[e^{\delta(h)U(h)} \right]_{\nu}^{\mu} . \quad (391)$$

Equivalently, this last expression can be re-written in terms of a surface integral of the Riemann tensor, projected along the surface area element bivector $A^{\alpha\beta}(C_h)$ associated with the loop,

$$\left[\prod_s \mathbf{R}_{s,s+1} \right]_{\nu}^{\mu} \approx \left[e^{\frac{1}{2} \int_S R^{\cdot\alpha\beta} A^{\alpha\beta}(C_h)} \right]_{\nu}^{\mu} . \quad (392)$$

More generally one might want to consider a near-planar, but non-infinitesimal, closed loop C , as shown in Fig. 21. Along this closed loop the overall rotation matrix will still be given by

$$R^{\mu}_{\nu}(C) = \left[\prod_{s \subset C} \mathbf{R}_{s,s+1} \right]_{\nu}^{\mu} \quad (393)$$

In analogy with the infinitesimal loop case, one would like to state that for the overall rotation matrix one has

$$R^{\mu}_{\nu}(C) \approx \left[e^{\delta(C)U(C)} \right]_{\nu}^{\mu} , \quad (394)$$

where $U_{\mu\nu}(C)$ is now an area bivector perpendicular to the loop - which will work only if the loop is close to planar so that $U_{\mu\nu}$ can be taken to be approximately constant along the path C .

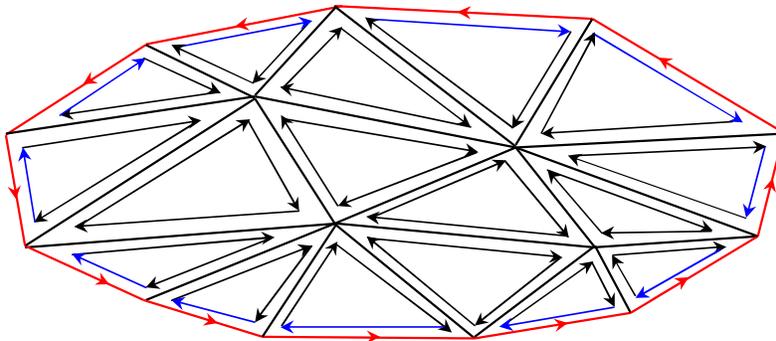


FIG. 21 Gravitational analog of the Wilson loop. A vector is parallel-transported along the larger outer loop. The enclosed minimal surface is tiled with parallel transport polygons, here chosen to be triangles for illustrative purposes. For each link of the dual lattice, the elementary parallel transport matrices $\mathbf{R}(s, s')$ are represented by arrows. In spite of the fact that the (Lorentz) matrices \mathbf{R} can fluctuate strongly in accordance with the local geometry, two contiguous, oppositely oriented arrows always give $\mathbf{R}\mathbf{R}^{-1} = 1$.

If that is true, then one can define, again in analogy with the infinitesimal loop case, an appropriate coordinate scalar by contracting the above rotation matrix $\mathbf{R}(C)$ with the bivector of Eq. (338), namely

$$W(C) = \omega_{\alpha\beta}(C) R^{\alpha\beta}(C) \quad (395)$$

where the loop bivector, $\omega_{\alpha\beta}(C) = (d-2)! V^{(d-2)} U_{\alpha\beta} = 2 A_C U_{\alpha\beta}(C)$ in four dimensions, is now intended as being representative of the overall geometric features of the loop. For example, it can be taken as an average of the hinge bivector $\omega_{\alpha\beta}(h)$ along the loop.

In the quantum theory one is of course interested in the average of the above loop operator $W(C)$, as in Eq. (327). The previous construction is indeed quite analogous to the Wilson loop definition in ordinary lattice gauge theories (Wilson, 1973), where it is defined via the trace of path ordered products of $SU(N)$ color rotation matrices. In gravity though the Wilson loop does not give any information about the static potential (Modanese, 1993; Hamber, 1993). It seems that the Wilson loop in gravity provides instead some insight into the large-scale curvature of the manifold, just as the infinitesimal loop contribution entering the lattice action of Eqs. (368) and (369) provides, through its averages, insight into the very short distance, local curvature. Of course for any continuum manifold one can define locally the parallel transport of a vector around a near-planar loop C . Indeed parallel transporting a vector around a closed loop represents a suitable operational way of detecting curvature locally. If the curvature of the manifold is small, one can treat the larger loop the same way as the small one; then the expression of Eq. (394) for the rotation matrix $\mathbf{R}(C)$ associated with a near-planar loop can be re-written in terms of a surface integral of the large-scale Riemann tensor, projected along the surface area element bivector $A^{\alpha\beta}(C)$ associated with the loop,

$$R^\mu{}_\nu(C) \approx \left[e^{\frac{1}{2} \int_S R \cdot_{\alpha\beta} A^{\alpha\beta}(C)} \right]^\mu{}_\nu. \quad (396)$$

Thus a direct calculation of the Wilson loop provides a way of determining the *effective* curvature at large distance scales, even in the case where short distance fluctuations in the metric may be significant. Conversely, the rotation matrix appearing in the elementary Wilson loop of Eqs. (359) and (362) only provides information about the parallel transport of vectors around *infinitesimal* loops, with size comparable to the ultraviolet cutoff.

One would expect that for a geometry fluctuating strongly at short distances (corresponding therefore to the small k limit) the infinitesimal parallel transport matrices $\mathbf{R}(s, s')$ should be distributed close to randomly, with a measure close to the uniform Haar measure, and with little correlation between neighboring hinges. In such instance one would have for the local quantum averages of the infinitesimal lattice parallel transports $\langle \mathbf{R} \rangle = 0$, but $\langle \mathbf{R}\mathbf{R}^{-1} \rangle \neq 0$, which would require, for a non-vanishing lowest order contribution to the Wilson loop, that the loop at least be tiled by elementary action contributions from Eqs. (368) or (369), thus forming a minimal surface spanning the loop. Then, in close analogy to the Yang-Mills case of Eq. (328) (see for example

Peskin and Schroeder, 1995), the leading contribution to the gravitational Wilson loop would be expected to follow an area law,

$$W(C) \sim \text{const. } k^{A(C)} \sim \exp(-A(C)/\xi^2) \quad (397)$$

where $A(C)$ is the minimal physical area spanned by the near-planar loop C , and $\xi = 1/\sqrt{|\ln k|}$ the gravitational correlation length at small k . Thus for a close-to-circular loop of perimeter P one would use $A(C) \approx P^2/4\pi$.

The rapid decay of the gravitational Wilson loop as a function of the area is seen here simply as a general and direct consequence of the disorder in the fluctuations of the parallel transport matrices $\mathbf{R}(s, s')$ at strong coupling. It should then be clear from the above discussion that the gravitational Wilson loop provides in a sense a measure of the magnitude of the large-scale, averaged curvature, where the latter is most suitably defined by the process of parallel-transporting test vectors around very large loops, and is therefore, from the above expression, computed to be of the order $R \sim 1/\xi^2$, at least in the small k limit. A direct calculation of the Wilson loop should therefore provide, among other things, a direct insight into whether the manifold is de Sitter or anti-de Sitter at large distances. More details on the lattice construction of the gravitational Wilson loop, the various issues that arise in its precise definition, and a sample calculation in the strong coupling limit of lattice gravity, can be found in (Hamber and Williams, 2007).

Finally we note that the definition of the gravitational Wilson loop is based on a surface with a given boundary C , in the simplest case the minimal surface spanning the loop. It is possible though to consider other surfaces built out of elementary parallel transport loops. An example of such a generic closed surface tiled with elementary parallel transport polygons (here chosen for illustrative purposes to be triangles) will be given later in Fig. 29.

Later similar surfaces will arise naturally in the context of the strong coupling (small k) expansion for gravity, as well as in the high dimension (large d) expansion.

7. Lattice Regularized Path Integral

As the edge lengths l_{ij} play the role of the continuum metric $g_{\mu\nu}(x)$, one would expect the discrete measure to involve an integration over the squared edge lengths (Hamber, 1984; Hartle, 1984; Hamber and Williams, 1999). Indeed the induced metric at a simplex is related to the squared edge lengths within that simplex, via the expression for the invariant line element $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$. After choosing coordinates along the edges emanating from a vertex, the relation between metric

perturbations and squared edge length variations for a given simplex based at 0 in d dimensions is

$$\delta g_{ij}(l^2) = \frac{1}{2} (\delta l_{0i}^2 + \delta l_{0j}^2 - \delta l_{ij}^2) . \quad (398)$$

For one d -dimensional simplex labeled by s the integration over the metric is thus equivalent to an integration over the edge lengths, and one has the identity

$$\left(\frac{1}{d!} \sqrt{\det g_{ij}(s)} \right)^\sigma \prod_{i \geq j} dg_{ij}(s) = \left(-\frac{1}{2} \right)^{\frac{d(d-1)}{2}} [V_d(l^2)]^\sigma \prod_{k=1}^{d(d+1)/2} dl_k^2 \quad (399)$$

There are $d(d+1)/2$ edges for each simplex, just as there are $d(d+1)/2$ independent components for the metric tensor in d dimensions (Cheeger, Müller and Schrader, 1981). Here one is ignoring temporarily the triangle inequality constraints, which will further require all sub-determinants of g_{ij} to be positive, including the obvious restriction $l_k^2 > 0$.

Let us discuss here briefly the simplicial inequalities that need to be imposed on the edge lengths (Wheeler 1964). These are conditions on the edge lengths l_{ij} such that the sites i can then be considered the vertices of a d -simplex embedded in flat d -dimensional Euclidean space. In one dimension, $d = 1$, one requires trivially for all edge lengths

$$l_{ij}^2 > 0 . \quad (400)$$

In two dimensions, $d = 2$, the conditions on the edge lengths are again $l_{ij}^2 > 0$ as in one dimensions, as well as

$$A_\Delta^2 = \left(\frac{1}{2!} \right)^2 \det g_{ij}^{(2)}(s) > 0 \quad (401)$$

which is equivalent, by virtue of Heron's formula for the area of a triangle $A_\Delta^2 = s(s - l_{ij})(s - l_{jk})(s - l_{ki})$ where s is the semi-perimeter $s = \frac{1}{2}(l_{ij} + l_{jk} + l_{ki})$, to the requirement that the area of the triangle be positive. In turn Eq. (401) implies that the triangle inequalities must be satisfied for all three edges,

$$\begin{aligned} l_{ij} + l_{jk} &> l_{ik} \\ l_{jk} + l_{ki} &> l_{ji} \\ l_{ki} + l_{ij} &> l_{kj} \end{aligned} \quad (402)$$

In three dimensions, $d = 3$, the conditions on the edge lengths are again such that one recovers a physical tetrahedron. One therefore requires for the individual edge lengths the condition of

Eq. (400), the reality and positivity of all four triangle areas as in Eq. (401), as well as the requirement that the volume of the tetrahedron be real and positive,

$$V_{\text{tetrahedron}}^2 = \left(\frac{1}{3!}\right)^2 \det g_{ij}^{(3)}(s) > 0 \quad (403)$$

The generalization to higher dimensions is such that one requires all triangle inequalities and their higher dimensional analogs to be satisfied,

$$\begin{aligned} l_{ij}^2 &> 0 \\ V_k^2 &= \left(\frac{1}{k!}\right)^2 \det g_{ij}^{(k)}(s) > 0 \end{aligned} \quad (404)$$

with $k = 2 \dots d$ for every possible choice of sub-simplex (and therefore sub-determinant) within the original simplex s .

The extension of the measure to many simplices glued together at their common faces is then immediate. For this purpose one first needs to identify edges $l_k(s)$ and $l_{k'}(s')$ which are shared between simplices s and s' ,

$$\int_0^\infty dl_k^2(s) \int_0^\infty dl_{k'}^2(s') \delta[l_k^2(s) - l_{k'}^2(s')] = \int_0^\infty dl_k^2(s) . \quad (405)$$

After summing over all simplices one derives, up to an irrelevant numerical constant, the unique functional measure for simplicial geometries

$$\int [dl^2] = \int_\epsilon^\infty \prod_s [V_d(s)]^\sigma \prod_{ij} dl_{ij}^2 \Theta[l_{ij}^2] . \quad (406)$$

Here $\Theta[l_{ij}^2]$ is a (step) function of the edge lengths, with the property that it is equal to one whenever the triangle inequalities and their higher dimensional analogs are satisfied, and zero otherwise. The quantity ϵ has been introduced as a cutoff at small edge lengths. If the measure is non-singular for small edges, one can safely take the limit $\epsilon \rightarrow 0$. In four dimensions the lattice analog of the DeWitt measure ($\sigma = 0$) takes on a particularly simple form, namely

$$\int [dl^2] = \int_0^\infty \prod_{ij} dl_{ij}^2 \Theta[l_{ij}^2] . \quad (407)$$

Lattice measures over the space of squared edge lengths have been used extensively in numerical simulations of simplicial quantum gravity (Hamber and Williams, 1984; Hamber, 1984; Berg 1985). The derivation of the above lattice measure closely parallels the analogous procedure in the continuum.

There is no obstacle in defining a discrete analog of the supermetric, as a way of introducing an invariant notion of distance between simplicial manifolds, as proposed in (Cheeger, Müller and

Schrader, 1984). It leads to an alternative way of deriving the lattice measure in Eq. (407), by considering the discretized distance between induced metrics $g_{ij}(s)$

$$\|\delta g(s)\|^2 = \sum_s G^{ijkl}[g(s)] \delta g_{ij}(s) \delta g_{kl}(s) , \quad (408)$$

with the inverse of the lattice DeWitt supermetric now given by the expression

$$G^{ijkl}[g(s)] = \frac{1}{2} \sqrt{g(s)} \left[g^{ik}(s) g^{jl}(s) + g^{il}(s) g^{jk}(s) + \lambda g^{ij}(s) g^{kl}(s) \right] \quad (409)$$

and with again $\lambda \neq -2/d$. This procedure defines a metric on the tangent space of positive real symmetric matrices $g_{ij}(s)$. After computing the determinant of G , the resulting functional measure is

$$\int d\mu[l^2] = \int \prod_s [\det G(g(s))]^{\frac{1}{2}} \prod_{i \geq j} dg_{ij}(s) , \quad (410)$$

with the determinant of the super-metric $G^{ijkl}(g(s))$ given by the local expression

$$\det G[g(s)] \propto (1 + \frac{1}{2}d\lambda) [g(s)]^{(d-4)(d+1)/4} , \quad (411)$$

Using Eq. (399), and up to irrelevant constants, one obtains again the standard lattice measure of Eq. (406). Of course the same procedure can be followed for the Misner-like measure, leading to a similar result for the lattice measure, but with a different power σ .

One might be tempted to try to find alternative lattice measures by looking directly at the discrete form for the supermetric, written as a quadratic form in the squared edge lengths (instead of the metric components), and then evaluating the resulting determinant. The main idea, inspired by work described in a paper (Lund and Regge, 1974) on the 3+1 formulation of simplicial gravity, is as follows. First one considers a lattice analog of the DeWitt supermetric by writing

$$\|\delta l^2\|^2 = \sum_{ij} G_{ij}(l^2) \delta l_i^2 \delta l_j^2 , \quad (412)$$

with $G_{ij}(l^2)$ now defined on the space of squared edge lengths (Hartle, Miller and Williams, 1997). The next step is to find an appropriate form for $G_{ij}(l^2)$ expressed in terms of known geometric objects. One simple way of constructing the explicit form for $G_{ij}(l^2)$, in any dimension, is to first focus on one simplex, and write the squared volume of a given simplex in terms of the induced metric components within the *same* simplex s ,

$$V^2(s) = \left(\frac{1}{d!}\right)^2 \det g_{ij}(l^2(s)) . \quad (413)$$

One computes to linear order

$$\frac{1}{V(l^2)} \sum_i \frac{\partial V^2(l^2)}{\partial l_i^2} \delta l_i^2 = \frac{1}{d!} \sqrt{\det(g_{ij})} g^{ij} \delta g_{ij} , \quad (414)$$

and to quadratic order

$$\frac{1}{V(l^2)} \sum_{ij} \frac{\partial^2 V^2(l^2)}{\partial l_i^2 \partial l_j^2} \delta l_i^2 \delta l_j^2 = \frac{1}{d!} \sqrt{\det(g_{ij})} \left[g^{ij} g^{kl} \delta g_{ij} \delta g_{kl} - g^{ij} g^{kl} \delta g_{jk} \delta g_{li} \right] . \quad (415)$$

The right hand side of this equation contains precisely the expression appearing in the continuum supermetric of Eq. (163), for the specific choice of the parameter $\lambda = -2$. One is led therefore to the identification

$$G_{ij}(l^2) = -d! \sum_s \frac{1}{V(s)} \frac{\partial^2 V^2(s)}{\partial l_i^2 \partial l_j^2} , \quad (416)$$

and therefore for the norm

$$\|\delta l^2\|^2 = \sum_s V(s) \left\{ -\frac{d!}{V^2(s)} \sum_{ij} \frac{\partial^2 V^2(s)}{\partial l_i^2 \partial l_j^2} \delta l_i^2 \delta l_j^2 \right\} . \quad (417)$$

One could be tempted at this point to write down a lattice measure, in parallel with Eq. (165), and write

$$\int [dl^2] = \int \prod_i \sqrt{\det G_{ij}^{(\omega')}(l^2)} dl_i^2 . \quad (418)$$

with

$$G_{ij}^{(\omega')}(l^2) = -d! \sum_s \frac{1}{[V(s)]^{1+\omega'}} \frac{\partial^2 V^2(s)}{\partial l_i^2 \partial l_j^2} , \quad (419)$$

where one has allowed for a parameter ω' , possibly different from zero, interpolating between apparently equally acceptable measures. The reasoning here is that, as in the continuum, different edge length measures, here parametrized by ω' , are obtained, depending on whether the local volume factor $V(s)$ is included in the supermetric or not.

One rather undesirable, and puzzling, feature of the lattice measure of Eq. (418) is that in general it is non-local, in spite of the fact that the original continuum measure of Eq. (167) is completely local (although it is clear that for some special choices of ω' and d , one does recover a local measure; thus in two dimensions and for $\omega' = -1$ one obtains again the simple result $\int [dl^2] = \int_0^\infty \prod_i dl_i^2$). Unfortunately irrespective of the value chosen for ω' , one can show (Hamber and Williams, 1999) that the measure of Eq. (418) disagrees with the continuum measure of Eq. (167) already to lowest order in the weak field expansion, and does not therefore describe an acceptable lattice measure.

The lattice action for pure four-dimensional Euclidean gravity contains a cosmological constant and Regge scalar curvature term

$$I_{latt} = \lambda_0 \sum_h V_h(l^2) - k \sum_h \delta_h(l^2) A_h(l^2) , \quad (420)$$

with $k = 1/(8\pi G)$, as well as possibly higher derivative terms. It only couples edges which belong either to the same simplex or to a set of neighboring simplices, and can therefore be considered as *local*, just like the continuum action, and leads to the regularized lattice functional integral

$$Z_{latt} = \int [d l^2] e^{-\lambda_0 \sum_h V_h + k \sum_h \delta_h A_h} , \quad (421)$$

where, as customary, the lattice ultraviolet cutoff is set equal to one (i.e. all length scales are measured in units of the lattice cutoff).

The lattice partition function Z_{latt} should then be compared to the continuum Euclidean Feynman path integral of Eq. (183),

$$Z_{cont} = \int [d g_{\mu\nu}] e^{-\lambda_0 \int dx \sqrt{g} + \frac{1}{16\pi G} \int dx \sqrt{g} R} , \quad (422)$$

Occasionally it can be convenient to include the λ_0 -term in the measure. For this purpose one defines

$$d\mu(l^2) \equiv [d l^2] e^{-\lambda_0 \sum_h V_h} . \quad (423)$$

It should be clear that this last expression represents a fairly non-trivial quantity, both in view of the relative complexity of the expression for the volume of a simplex, Eq. (335), and because of the generalized triangle inequality constraints already implicit in $[d l^2]$. But, like the continuum functional measure, it is certainly *local*, to the extent that each edge length appears only in the expression for the volume of those simplices which explicitly contain it. Furthermore, λ_0 sets the overall scale and can therefore be set equal to one without any loss of generality.

8. An Elementary Example

In the very simple case of one dimension ($d = 1$) one can work out explicitly a number of details, and see how potential problems with the functional measure arise, and how they are resolved.

In one dimension one discretizes the line by introducing N points, with lengths l_n associated with the edges, and periodic boundary conditions, $l_{N+1} = l_1$. Here l_n is the distance between

points n and $n + 1$. The only surviving invariant term in one dimension is then the overall length of a curve,

$$L(l) = \sum_{n=1}^N l_n , \quad (424)$$

which corresponds to

$$\int dx \sqrt{g(x)} = \int dx e(x) \quad (425)$$

(with $g(x) \equiv g_{00}(x)$) in the continuum. Here $e(x)$ is the “einbein”, and satisfies the obvious constraint $\sqrt{g(x)} = e(x) > 0$. In this context the discrete action is unique, preserving the geometric properties of the continuum definition. From the expression for the invariant line element, $ds^2 = gdx^2$, one associates $g(x)$ with l_n^2 (and therefore $e(x)$ with l_n). One can further take the view that distances can only be assigned between vertices which appear on some lattice in the ensemble, although this is not strictly necessary, as distances can also be defined for locations that do not coincide with any specific vertex.

The gravitational measure then contains an integration over the elementary lattice degrees of freedom, the lattice edge lengths. For the edges one writes the lattice integration measure as

$$\int d\mu[l] = \prod_{n=1}^N \int_0^\infty dl_n^2 l_n^\sigma , \quad (426)$$

where σ is a parameter interpolating between different local measures. The positivity of the edge lengths is all that remains of the triangle inequality constraints in one dimension. The factor l_n^σ plays a role analogous to the $g^{\sigma/2}$ which appears for continuum measures in the Euclidean functional integral.

The functional measure does not have compact support, and the cosmological term (with a coefficient $\lambda_0 > 0$) is therefore necessary to obtain convergence of the functional integral, as can be seen for example from the expression for the average edge length,

$$\langle L(l) \rangle = \langle \sum_{n=1}^N l_n \rangle = \mathcal{Z}_N^{-1} \prod_{n=1}^N \int_0^\infty dl_n^2 l_n^\sigma \exp\left(-\lambda_0 \sum_{n=1}^N l_n\right) \sum_{n=1}^N l_n = \frac{2 + \sigma}{\lambda_0} N \quad (427)$$

with

$$\mathcal{Z}_N(\lambda_0) = \prod_{n=1}^N \int_0^\infty dl_n^2 l_n^\sigma \exp\left(-\lambda_0 \sum_{n=1}^N l_n\right) = \left[\frac{2 \Gamma(2 + \sigma)}{\lambda_0^{2 + \sigma}} \right]^N \quad (428)$$

Similarly one finds for the fluctuation in the total length $\Delta L/L = 1/\sqrt{(2 + \sigma)N}$, which requires $\sigma > -2$. Different choices for λ_0 then correspond to trivial rescalings of the average lattice spacing, $l_0 \equiv \langle l \rangle = (2 + \sigma)/\lambda_0$.

In the continuum, the action of Eq. (425) is invariant under continuous reparametrizations

$$x \rightarrow x'(x) = x - \epsilon(x) \quad (429)$$

$$g(x) \rightarrow g'(x') = \left(\frac{dx}{dx'} \right)^2 g(x) = g(x) + 2 g(x) \left(\frac{d\epsilon}{dx} \right) + O(\epsilon^2) , \quad (430)$$

or equivalently

$$\delta g(x) \equiv g'(x') - g(x) = 2g\partial\epsilon , \quad (431)$$

and we have set $\partial \equiv d/dx$. A gauge can then be chosen by imposing $g'(x') = 1$, which can be achieved by the choice of coordinates $x' = \int dx \sqrt{g(x)}$.

The discrete analog of the transformation rule is

$$\delta l_n = \epsilon_{n+1} - \epsilon_n , \quad (432)$$

where the ϵ_n 's represent continuous gauge transformations defined on the lattice vertices. In order for the edge lengths to remain positive, one needs to require $\epsilon_n - \epsilon_{n+1} < l_n$, which is certainly satisfied for sufficiently small ϵ 's. The above continuous symmetry is an exact invariance of the lattice action of Eq. (424), since

$$\delta L = \sum_{n=1}^N \delta l_n = \sum_{n=1}^N \epsilon_{n+1} - \sum_{n=1}^N \epsilon_n = 0 , \quad (433)$$

and we have used $\epsilon_{N+1} = \epsilon_1$. Moreover, it is the only local symmetry of the action of Eq. (424).

The infinitesimal local invariance property defined in Eq. (432) formally selects a unique measure over the edge lengths, corresponding to $\prod_n dl_n$ ($\sigma = -1$ in Eq. (426)), as long as we ignore the effects of the lower limit of integration. On the other hand for sufficiently large lattice diffeomorphisms, the lower limit of integration comes into play (since we require $l_n > 0$ always) and the measure is no longer invariant. Thus a measure $\int_{-\infty}^{\infty} \prod dl_n$ would not be acceptable on physical grounds; it would violate the constraint $\sqrt{g} > 0$ or $e > 0$.

The same functional measure can be obtained from the following physical consideration. Define the gauge invariant distance d between two configurations of edge lengths $\{l_n\}$ and $\{l'_n\}$ by

$$d^2(l, l') \equiv [L(l) - L'(l')]^2 = \left(\sum_{n=1}^N l_n - \sum_{n'=1}^N l'_{n'} \right)^2 = \sum_{n=1}^N \sum_{n'=1}^N \delta l_n M_{n,n'} \delta l_{n'} , \quad (434)$$

with $M_{n,n'} = 1$. Since M is independent of l_n and $l'_{n'}$, the ensuing measure is again simply proportional to $\prod dl_n$. Note that the above metric over edge length deformations δl is non-local.

In the continuum, the functional measure is usually determined by considering the following (local) norm in function space,

$$\|\delta g\|^2 = \int dx \sqrt{g(x)} \delta g(x) G(x) \delta g(x) , \quad (435)$$

and diffeomorphism invariance would seem to require $G(x) = 1/g^2(x)$. The volume element in function space is then an ultraviolet regulated version of $\prod_x \sqrt{G(x)} dg(x) = \prod_x dg(x)/g(x)$, which is the Misner measure in one dimension. Its naive discrete counterpart would be $\prod dl_n/l_n$, which is not invariant under the transformation of Eq. (432) (it is invariant under $\delta l_n = l_n(\epsilon_{n+1} - \epsilon_n)$, which is not an invariance of the action).

The point of the discussion of the one-dimensional case is to bring to the surface the several non-trivial issues that arise when defining a properly regulated version of the continuum Feynman functional measure $[dg_{\mu\nu}]$, and how they can be systematically resolved.

9. Lattice Higher Derivative Terms

So far only the gravitational Einstein-Hilbert contribution to the lattice action and the cosmological constant term have been considered. There are several motivations for extending the discussion to lattice higher derivative terms, which would include the fact that these terms a) might appear in the original microscopic action, or might have to be included to cure the classical unboundedness problem of the Euclidean Einstein-Hilbert action, b) that they are in any case generated by radiative corrections, and c) that on a more formal level they may shed new light on the relationship between the lattice and continuum expressions for curvature terms as well as quantities such as the Riemann tensor on a hinge, Eq. (366).

For these reasons we will discuss here a generalization of the Regge gravity equivalent of the Einstein action to curvature squared terms. When considering contributions quadratic in the curvature there are overall six possibilities, listed in Eq. (94). Among the two topological invariants, the Euler characteristic χ for a simplicial decomposition may be obtained from a particular case of the general formula for the analogue of the Lipschitz-Killing curvatures of smooth Riemannian manifolds for piecewise flat spaces. The formula of (Cheeger, Müller and Schrader, 1984) reduces in four dimensions to

$$\chi = \sum_{\sigma^0} \left[1 - \sum_{\sigma^2 \supset \sigma^0} (0, 2) - \sum_{\sigma^4 \supset \sigma^0} (0, 4) + \sum_{\sigma^4 \supset \sigma^2 \supset \sigma^0} (0, 2)(2, 4) \right] \quad (436)$$

where σ^i denotes an i -dimensional simplex and (i, j) denotes the (internal) dihedral angle at an i -dimensional face of a j -dimensional simplex. Thus, for example, $(0, 2)$ is the angle at the vertex

of a triangle and $(2, 4)$ is the dihedral angle at a triangle in a 4-simplex (The normalization of the angles is such that the volume of a sphere in any dimension is one; thus planar angles are divided by 2π , 3-dimensional solid angles by 4π and so on).

Of course, as noted before, there is a much simpler formula for the Euler characteristic of a simplicial complex

$$\chi = \sum_{i=0}^d (-1)^i N_i \quad (437)$$

where N_i is the number of simplices of dimension i . However, it may turn out to be useful in quantum gravity calculations to have a formula for χ in terms of the angles, and hence of the edge lengths, of the simplicial decomposition. These expressions are interesting and useful, but do not shed much light on how the other curvature squared terms should be constructed.

In a piecewise linear space curvature is detected by going around elementary loops which are dual to a $(d - 2)$ -dimensional subspace. The area of the loop itself is not well defined, since any loop inside the d -dimensional simplices bordering the hinge will give the same result for the deficit angle. On the other hand the hinge has a content (the length of the edge in $d = 3$ and the area of the triangle in $d = 4$), and there is a natural volume associated with each hinge, defined by dividing the volume of each simplex touching the hinge into a contribution belonging to that hinge, and other contributions belonging to the other hinges on that simplex (Hamber and Williams, 1984). The contribution belonging to that simplex will be called *dihedral volume* V_d . The volume V_h associated with the hinge h is then naturally the sum of the dihedral volumes V_d belonging to each simplex

$$V_h = \sum_{\substack{\text{d-simplices} \\ \text{meeting on h}}} V_d \quad (438)$$

The dihedral volume associated with each hinge in a simplex can be defined using dual volumes, a barycentric subdivision, or some other natural way of dividing the volume of a d -simplex into $d(d + 1)/2$ parts. If the theory has some reasonable continuum limit, then the final results should not depend on the detailed choice of volume type.

As mentioned previously, there is a well-established procedure for constructing a dual lattice for any given lattice. This involves constructing polyhedral cells, known in the literature as Voronoi polyhedra, around each vertex, in such a way that the cell around each particular vertex contains all points which are nearer to that vertex than to any other vertex. Thus the cell is made up from $(d - 1)$ -dimensional subspaces which are the perpendicular bisectors of the edges in the original lattice, $(d - 2)$ -dimensional subspaces which are orthogonal to the 2-dimensional subspaces of the

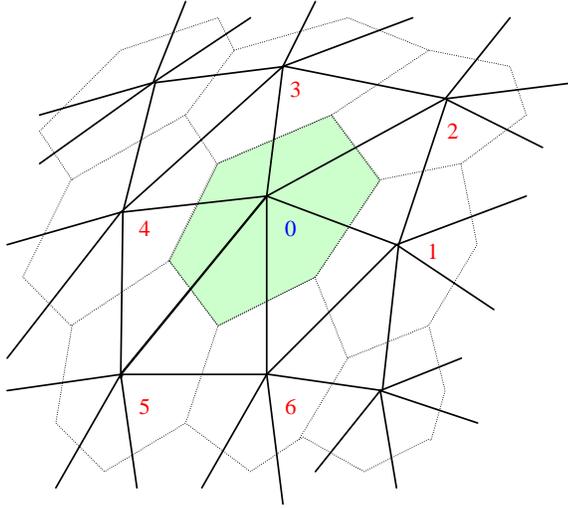


FIG. 22 Illustration of dual volumes in two dimensions. The vertices of the polygons reside in the dual lattice. The shaded region describes the dual area associated with the vertex 0.

original lattice, and so on. General formulas for dual volumes are given in (Hamber and Williams, 1986). In the case of the barycentric subdivision, the dihedral volume is just $2/d(d+1)$ times the volume of the simplex. This leads one to conclude that there is a natural area A_{C_h} associated with each hinge

$$A_{C_h} = \frac{V_h}{A_h^{(d-2)}} \quad (439)$$

obtained by dividing the volume per hinge (which is d -dimensional) by the volume of the hinge (which is $(d-2)$ -dimensional).

The next step is to find terms equivalent to the continuum expression of Eq. (94), and the remainder of this section will be devoted to this problem. It may be objected that since in Regge calculus where the curvature is restricted to the hinges which are subspaces of dimension 2 less than that of the space considered, then the curvature tensor involves δ -functions with support on the hinges, and so higher powers of the curvature tensor are not defined. But this argument clearly does not apply to the Euler characteristic and the Hirzebruch signature of Eq. Eq. (94), which are both integrals of 4-forms. However it is a common procedure in lattice field theory to take powers of fields defined at the same point, and there is no reason why one should not consider similar terms in lattice gravity. Of course one would like the expressions to correspond to the continuum ones as the edge lengths of the simplicial lattice become smaller and smaller.

Since the curvature is restricted to the hinges, it is natural that expressions for curvature

integrals should involve sums over hinges as in Eq. (368). The curvature tensor, which involves second derivatives of the metric, is of dimension L^{-2} . Therefore $\frac{1}{4} \int d^d x \sqrt{g} R^n$ is of dimension L^{d-2n} . Thus if one postulates that an R^2 -type term will involve the square of $A_h \delta_h$, which is of dimension $L^{2(d-2)}$, then one will need to divide by some d -dimensional volume to obtain the correct dimension for the extra term in the action. Now any hinge is surrounded by a number of d -dimensional simplices, so the procedure of dividing by a d -dimensional volume seems ambiguous. The crucial step is to realize that there *is* a unique d -dimensional volume associated with each hinge, as described above.

If one regards the invariant volume element $\sqrt{g} d^d x$ as being represented by V_h of Eq. (438) when one performs the sum over hinges as in Eq. (368), then this means that one may regard the scalar curvature R contribution as being represented at each hinge by $2A_h \delta_h / V_h$

$$\frac{1}{2} \int d^d x \sqrt{g} R \rightarrow \sum_{\text{hinges } h} V_h \frac{A_h \delta_h}{V_h} \equiv \sum_{\text{hinges } h} A_h \delta_h \quad (440)$$

It is then straightforward to see that a candidate curvature squared term is

$$\sum_{\text{hinges } h} V_h \left(\frac{A_h \delta_h}{V_h} \right)^2 \equiv \sum_{\text{hinges } h} V_h \left(\frac{\delta_h}{A_{C_h}} \right)^2 \quad (441)$$

where A_{C_h} was defined in Eq. 439). Since the expression in Eq. (441) vanishes if and only if all deficit angles are zero, it is naturally identified with the continuum Riemann squared term,

$$\frac{1}{4} \int d^d x \sqrt{g} R_{\mu\nu\lambda\sigma} R^{\mu\nu\lambda\sigma} \rightarrow \sum_{\text{hinges } h} V_h \left(\frac{\delta_h}{A_{C_h}} \right)^2 \quad (442)$$

The above construction then leaves open the question of how to construct the remaining curvature squared terms in four dimensions. If one takes the form given previously in Eq. (366) for the Riemann tensor on a hinge and contracts one obtains

$$R(h) = 2 \frac{\delta_h}{A_{C_h}} \quad (443)$$

which agrees with the form used in the Regge action for R . But one also finds readily that with this choice the higher derivative terms are all proportional to each other (Hamber and Williams, 1986),

$$\frac{1}{4} R_{\mu\nu\rho\sigma}(h) R^{\mu\nu\rho\sigma}(h) = \frac{1}{2} R_{\mu\nu}(h) R^{\mu\nu}(h) = \frac{1}{4} R(h)^2 = \left(\frac{\delta_h}{A_{C_h}} \right)^2 \quad (444)$$

Furthermore if one uses the above expression for the Riemann tensor to evaluate the contribution to the Euler characteristic on each hinge one obtains zero, and becomes clear that at least in this case one needs cross terms involving contributions from different hinges.

The next step is therefore to embark on a slightly more sophisticated approach, and construct the full Riemann tensor by considering more than one hinge. Define the Riemann tensor for a simplex s as a weighted sum of hinge contributions

$$\left[R_{\mu\nu\rho\sigma} \right]_s = \sum_{h \subset s} \omega_h \left[\frac{\delta}{A_C} U_{\mu\nu} U_{\rho\sigma} \right]_h \quad (445)$$

where the ω_h are dimensionless weights, to be determined later. After squaring one obtains

$$\left[R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} \right]_s = \sum_{h, h' \subset s} \omega_h \omega_{h'} \left[\frac{\delta}{A_C} U_{\mu\nu} U_{\rho\sigma} \right]_h \left[\frac{\delta}{A_C} U^{\mu\nu} U^{\rho\sigma} \right]_{h'} \quad (446)$$

The question of the weights ω_h introduced in Eq. (445) will now be addressed. Consider the expression for the scalar curvature of a simplex defined as

$$\left[R \right]_s = \sum_{h \subset s} \omega_h \left[2 \frac{\delta}{A_C} \right]_h \quad (447)$$

It is clear from the formulae given above for the lattice curvature invariants (constructed in a simplex by summing over hinge contributions) that there is again a natural volume associated with them : the sum of the volumes of the hinges in the simplex

$$V_s = \sum_{h \subset s} V_h \quad (448)$$

where V_h is the volume around the hinge, as defined in Eq. (438). Summing the scalar curvature over all simplices, one should recover Regge's expression

$$\sum_s V_s \left[R \right]_s = \sum_s \sum_{h \subset s} \omega_h \left[2 \frac{\delta}{A_C} \right]_h = \sum_h \delta_h A_h \quad (449)$$

which implies

$$N_{2,4} V_s \omega_h \frac{\delta_h}{A_{C_h}} \equiv N_{2,4} V_s \omega_h \frac{\delta_h A_h}{V_h} = \delta_h A_h \quad (450)$$

where $N_{2,4}$ is the number of simplices meeting on that hinge. Therefore the correct choice for the weights is

$$\omega_h = \frac{V_h}{N_{2,4} V_s} = \frac{V_h}{N_{2,4} \sum_{h \subset s} V_h} \quad (451)$$

Thus the weighting factors that reproduce Regge's formula for the Einstein action are just the volume fractions occupied by the various hinges in a simplex, which is not surprising (of course the above formulae are not quite unique, since one might have done the above construction of higher derivative terms by considering a point p instead of a four-simplex s).

In particular the following form for the Weyl tensor squared was given in (Hamber and Williams, 1986)

$$\int d^d x \sqrt{g} C_{\mu\nu\lambda\sigma} C^{\mu\nu\lambda\sigma} \sim \frac{2}{3} \sum_s V_s \sum_{h, h' \subset s} \epsilon_{h, h'} \left(\omega_h \left[\frac{\delta}{A_C} \right]_h - \omega_{h'} \left[\frac{\delta}{A_C} \right]_{h'} \right)^2 \quad (452)$$

which introduces a short range coupling between deficit angles. The numerical factor $\epsilon_{h, h'}$ is equal to 1 if the two hinges h, h' have one edge in common and -2 if they do not. Note that this particular interaction term has the property that it requires neighboring deficit angles to have similar values, but it does not require them to be small, which is a key property one would expect from the Weyl curvature squared term.

In conclusion the formulas given above allow one to construct the remaining curvature squared terms in four dimensions, and in particular to write, for example, the lattice analog of the continuum curvature squared action of Eq. (104)

$$I = \int d^4 x \sqrt{g} \left[\lambda_0 - k R - b R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} + \frac{1}{2} (a + 4b) C_{\mu\nu\rho\sigma} C^{\mu\nu\rho\sigma} \right] \quad (453)$$

To compare with the form of Eq. (104) use has been made of $R^2 = 3R_{\mu\nu\rho\sigma}^2 - 6C^2$ and $R_{\mu\nu}^2 = R_{\mu\nu\rho\sigma}^2 - \frac{3}{2}C^2$, up to additive constant contributions. A special case corresponds to $b = -a/4$, which gives a pure $R_{\mu\nu\rho\sigma}^2$ contribution. The latter vanishes if and only if the curvature is locally zero, which is not true of any of the other curvature squared terms.

10. Scalar Matter Fields

In the previous section we have discussed the construction and the invariance properties of a lattice action for pure gravity. Next a scalar field can be introduced as the simplest type of dynamical matter that can be coupled invariantly to gravity. In the continuum the scalar action for a single component field $\phi(x)$ is usually written as

$$I[g, \phi] = \frac{1}{2} \int dx \sqrt{g} [g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + (m^2 + \xi R) \phi^2] + \dots \quad (454)$$

where the dots denote scalar self-interaction terms. Thus for example a scalar field potential $U(\phi)$ could be added containing quartic field terms, whose effects could be of interest in the context of cosmological models where spontaneously broken symmetries play an important role. The dimensionless coupling ξ is arbitrary; two special cases are the minimal ($\xi = 0$) and the

conformal ($\xi = \frac{1}{6}$) coupling case. In the following we shall mostly consider the case $\xi = 0$. Also, it will be straightforward to extend later the treatment to the case of an N_s -component scalar field ϕ_i^a with $a = 1, \dots, N_s$.

One way to proceed is to introduce a lattice scalar ϕ_i defined at the vertices of the simplices. The corresponding lattice action can then be obtained through a procedure by which the original continuum metric is replaced by the induced lattice metric, with the latter written in terms of squared edge lengths as in Eq. (333). For illustrative purposes only the two-dimensional case will be worked out explicitly here (Christ, Friedberg and Lee, 1982; Itzykson, 1983; Bander and Itzykson, 1983; Jevicki and Ninomiya, 1984). The generalization to higher dimensions is straightforward, and in the end the final answer for the lattice scalar action is almost identical to the two dimensional form. Furthermore in two dimensions it leads to a natural discretization of the bosonic string action (Polyakov, 1989).

In two dimensions the simplicial lattice is built out of triangles. For a given triangle it will be convenient to use the notation of Fig. 23, which will display more readily the symmetries of the resulting scalar lattice action. Here coordinates will be picked in each triangle along the (1,2) and (1,3) directions.

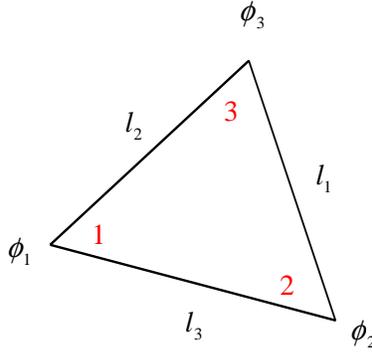


FIG. 23 Labeling of edges and fields for the construction of the scalar field action.

To construct a lattice action for the scalar field, one performs in two dimensions the replacement

$$\begin{aligned}
 g_{\mu\nu}(x) &\longrightarrow g_{ij}(\Delta) \\
 \det g_{\mu\nu}(x) &\longrightarrow \det g_{ij}(\Delta) \\
 g^{\mu\nu}(x) &\longrightarrow g^{ij}(\Delta) \\
 \partial_\mu \phi \partial_\nu \phi &\longrightarrow \Delta_i \phi \Delta_j \phi
 \end{aligned} \tag{455}$$

with the following definitions

$$g_{ij}(\Delta) = \begin{pmatrix} l_3^2 & \frac{1}{2}(-l_1^2 + l_2^2 + l_3^2) \\ \frac{1}{2}(-l_1^2 + l_2^2 + l_3^2) & l_2^2 \end{pmatrix}, \quad (456)$$

$$\det g_{ij}(\Delta) = \frac{1}{4} [2(l_1^2 l_2^2 + l_2^2 l_3^2 + l_3^2 l_1^2) - l_1^4 - l_2^4 - l_3^4] \equiv 4A_\Delta^2, \quad (457)$$

$$g^{ij}(\Delta) = \frac{1}{\det g(\Delta)} \begin{pmatrix} l_2^2 & \frac{1}{2}(l_1^2 - l_2^2 - l_3^2) \\ \frac{1}{2}(l_1^2 - l_2^2 - l_3^2) & l_3^2 \end{pmatrix}. \quad (458)$$

The scalar field derivatives get replaced as usual by finite differences

$$\partial_\mu \phi \longrightarrow (\Delta_\mu \phi)_i = \phi_{i+\mu} - \phi_i. \quad (459)$$

where the index μ labels the possible directions in which one can move away from a vertex within a given triangle. Then

$$\Delta_i \phi \Delta_j \phi = \begin{pmatrix} (\phi_2 - \phi_1)^2 & (\phi_2 - \phi_1)(\phi_3 - \phi_1) \\ (\phi_2 - \phi_1)(\phi_3 - \phi_1) & (\phi_3 - \phi_1)^2 \end{pmatrix}, \quad (460)$$

Then the discrete scalar field action takes the form

$$I = \frac{1}{16} \sum_{\Delta} \frac{1}{A_\Delta} [l_1^2 (\phi_2 - \phi_1)(\phi_3 - \phi_1) + l_2^2 (\phi_3 - \phi_2)(\phi_1 - \phi_2) + l_3^2 (\phi_1 - \phi_3)(\phi_2 - \phi_3)] . \quad (461)$$

where the sum is over all triangles on the lattice. Using the identity

$$(\phi_i - \phi_j)(\phi_i - \phi_k) = \frac{1}{2} [(\phi_i - \phi_j)^2 + (\phi_i - \phi_k)^2 - (\phi_j - \phi_k)^2], \quad (462)$$

one obtains after some re-arrangements the slightly more appealing expression for the action of a massless scalar field (Bander and Itzykson, 1984)

$$I(l^2, \phi) = \frac{1}{2} \sum_{\langle ij \rangle} A_{ij} \left(\frac{\phi_i - \phi_j}{l_{ij}} \right)^2, \quad (463)$$

A_{ij} is the dual (Voronoi) area associated with the edge ij , and the symbol $\langle ij \rangle$ denotes a sum over nearest neighbor lattice vertices. It is immediate to generalize the action of Eq. (463) to higher dimensions, with the two-dimensional Voronoi volumes replaced by their higher dimensional analogs, leading to

$$I(l^2, \phi) = \frac{1}{2} \sum_{\langle ij \rangle} V_{ij}^{(d)} \left(\frac{\phi_i - \phi_j}{l_{ij}} \right)^2, \quad (464)$$

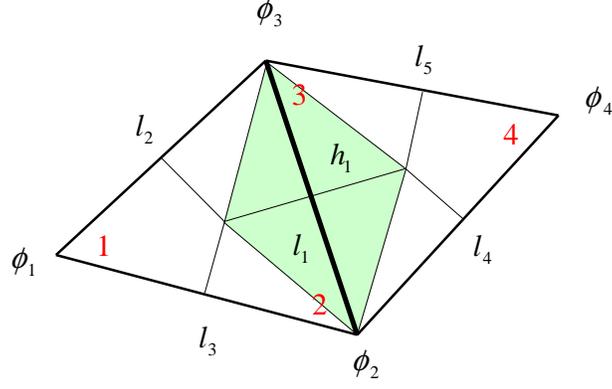


FIG. 24 Dual area associated with the edge l_1 (shaded area), and the corresponding dual link h_1 .

Here $V_{ij}^{(d)}$ is the dual (Voronoi) volume associated with the edge ij , and the sum is over all links on the lattice.

In two dimensions, in terms of the edge length l_{ij} and the dual edge length h_{ij} , connecting neighboring vertices in the dual lattice, one has $A_{ij} = \frac{1}{2}h_{ij}l_{ij}$ (see Fig. 24). Other choices for the lattice subdivision will lead to a similar formula for the lattice action, with the Voronoi dual volumes replaced by their appropriate counterparts for the new lattice. Explicitly, for an edge of length l_1 the dihedral dual volume contribution is given by

$$A_{l_1} = \frac{l_1^2(l_2^2 + l_3^2 - l_1^2)}{16A_{123}} + \frac{l_1^2(l_4^2 + l_5^2 - l_1^2)}{16A_{234}} = \frac{1}{2}l_1h_1 \quad , \quad (465)$$

with h_1 is the length of the edge dual to l_1 .

On the other hand the barycentric dihedral area for the same edge would be simply

$$A_{l_1} = (A_{123} + A_{234})/3 \quad . \quad (466)$$

It is well known that one of the disadvantages of the Voronoi construction is the lack of positivity of the dual volumes, as pointed out in (Hamber and Williams, 1984). Thus some of the weights appearing in Eq. (463) can be negative for such an action. For the barycentric subdivision this problem does not arise, as the areas A_{ij} are always positive due to the enforcement of the triangle inequalities. Thus from a practical point of view the barycentric volume subdivision is the simplest to deal with.

The scalar action of Eq. (464) has a very natural form: it involves the squared difference of fields at neighboring points divided by their invariant distances $(\phi_i - \phi_j)/l_{ij}$, weighted by the appropriate

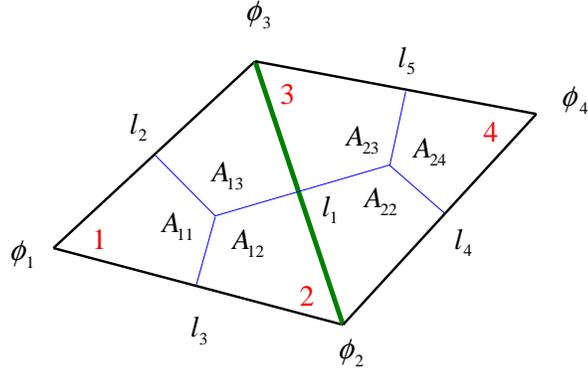


FIG. 25 More dual areas appearing in the scalar field action.

space-time volume element $V_{ij}^{(d)}$ associated with the lattice link ij . This suggests that one could just as well define the scalar fields on the vertices of the dual lattice, and write

$$I(l^2, \phi) = \frac{1}{2} \sum_{\langle rs \rangle} V_{rs}^{(d)} \left(\frac{\phi_r - \phi_s}{l_{rs}} \right)^2, \quad (467)$$

with l_{rs} the length of the edge connecting the dual lattice vertices r and s , and consequently $V_{rs}^{(d)}$ the spacetime volume fraction associated with the dual lattice edge rs . One would expect both forms to be equivalent in the continuum limit.

Continuing on with the two-dimensional case, mass and curvature terms such as the ones appearing in Eq. (454) can be added to the action, so that the total scalar lattice action contribution becomes

$$I = \frac{1}{2} \sum_{\langle ij \rangle} A_{ij} \left(\frac{\phi_i - \phi_j}{l_{ij}} \right)^2 + \frac{1}{2} \sum_i A_i (m^2 + \xi R_i) \phi_i^2. \quad (468)$$

The term containing the discrete analog of the scalar curvature involves the quantity

$$A_i R_i \equiv \sum_{h \supset i} \delta_h \sim \sqrt{g} R. \quad (469)$$

In the above expression for the scalar action, A_{ij} is the area associated with the edge l_{ij} , while A_i is associated with the site i . Again there is more than one way to define the volume element A_i , (Hamber and Williams, 1986), but under reasonable assumptions, such as positivity, one expects to get equivalent results in the lattice continuum limit, if it exists.

In higher dimensions one would use

$$I = \frac{1}{2} \sum_{\langle ij \rangle} V_{ij}^{(d)} \left(\frac{\phi_i - \phi_j}{l_{ij}} \right)^2 + \frac{1}{2} \sum_i V_i^{(d)} (m^2 + \xi R_i) \phi_i^2 \quad (470)$$

where the term containing the discrete analog of the scalar curvature involves

$$V_i^{(d)} R_i \equiv \sum_{h \supset i} \delta_h V_h^{(d-2)} \sim \sqrt{g} R \quad (471)$$

In the expression for the scalar action, $V_{ij}^{(d)}$ is the (dual) volume associated with the edge l_{ij} , while $V_i^{(d)}$ is the (dual) volume associated with the site i .

The lattice scalar action contains a mass parameter m , which has to be tuned to zero in lattice units to achieve the lattice continuum limit for scalar correlations. The agreement between different lattice actions in the smooth limit can be shown explicitly in the lattice weak field expansion. But in general, as is already the case for the purely gravitational action, the correspondence between lattice and continuum operators is true classically only up to higher derivative corrections. But such higher derivative corrections in the scalar field action are expected to be irrelevant when looking at the large distance limit, and they will not be considered here any further.

As an extreme case one could even consider a situation in which the matter action by itself is the only action contribution, without any additional term for the gravitational field, but still with a non-trivial gravitational measure; integration over the scalar field would then give rise to an effective non-local gravitational action.

Finally let us take notice here of the fact that if an N_s -component scalar field is coupled to gravity, the power σ appearing in the gravitational functional measure has to be modified to include an additional factor of $\prod_x (\sqrt{g})^{N_s/2}$. The additional measure factor insures that the integral

$$\int \prod_x \left[d\phi (\sqrt{g})^{\frac{N_s}{2}} \right] \exp \left(-\frac{1}{2} m^2 \int \sqrt{g} \phi^2 \right) = \left(\frac{2\pi}{m^2} \right)^{\frac{N_s V}{2}} \quad (472)$$

evaluates to a constant. Thus for large mass m the scalar field completely decouples, leaving only the dynamics of the pure gravitational field.

The quadratic scalar field action of Eq. (464) can be written in terms of a matrix $\Delta_{ij}(l^2)$

$$I(l^2, \phi) = -\frac{1}{2} \sum_{\langle ij \rangle} \phi_i \Delta_{ij}(l^2) \phi_j \quad (473)$$

The matrix $\Delta_{ij}(l^2)$ can then be regarded as a lattice version of the continuum scalar Laplacian,

$$\Delta(g) = \frac{1}{\sqrt{g}} \partial_\mu \sqrt{g} g^{\mu\nu} \partial_\nu \quad (474)$$

for a given background metric. This then allows one to define the massless lattice scalar propagator as the inverse of the above matrix, $G_{ij}(l^2) = \Delta_{ij}^{-1}(l^2)$. The continuum scalar propagator for a finite

scalar mass m and in a given background geometry, evaluated for large separations $d(x, y) \gg m^{-1}$,

$$G(x, y|g) = \langle x | \frac{1}{-\Delta(g) + m^2} | y \rangle$$

$$\underset{d(x,y) \rightarrow \infty}{\sim} d^{-(d-1)/2}(x, y) \exp\{-m d(x, y)\} \quad (475)$$

involves the geodesic distance $d(x, y)$ between points x and y ,

$$d(x, y|g) = \int_{\tau(x)}^{\tau(y)} d\tau \sqrt{g_{\mu\nu}(\tau) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}} \quad (476)$$

Analogously, one can define the discrete massive lattice scalar propagator

$$G_{ij}(l^2) = \left[\frac{1}{-\Delta(l^2) + m^2} \right]_{ij}$$

$$\underset{d(i,j) \rightarrow \infty}{\sim} d^{-(d-1)/2}(i, j) \exp\{-m d(i, j)\} \quad (477)$$

where $d(i, j)$ is the lattice geodesic distance between vertex i and vertex j . The inverse can be computed, for example, via the recursive expansion (valid for $m^2 > 0$ to avoid the zero eigenvalue of the Laplacian)

$$\frac{1}{-\Delta(l^2) + m^2} = \frac{1}{m^2} \sum_{n=0}^{\infty} \left(\frac{1}{m^2} \Delta(l^2) \right)^n . \quad (478)$$

The large distance behavior of the Euclidean (flat space) massive free field propagator in d dimensions is known in the statistical mechanics literature as the Ornstein-Zernike result.

As a consequence, the lattice propagator $G_{ij}(l^2)$ can be used to estimate the lattice geodesic distance $d(i, j|l^2)$ between any two lattice points i and j in a fixed background lattice geometry (provided again that their mutual separation is such that $d(i, j) \gg m^{-1}$).

$$d(i, j) \underset{d(i,j) \rightarrow \infty}{\sim} -\frac{1}{m} \ln G_{ij}(l^2) \quad (479)$$

11. Invariance Properties of the Scalar Action

In the very simple case of one dimension ($d = 1$) one can work out the details to any degree of accuracy, and see how potential problems arise and how they are resolved. Introduce a scalar field ϕ_n defined on the sites, with action

$$I(\phi) = \frac{1}{2} \sum_{n=1}^N V_1(l_n) \left(\frac{\phi_{n+1} - \phi_n}{l_n} \right)^2 + \frac{1}{2} \omega \sum_{n=1}^N V_0(l_n) \phi_n^2 , \quad (480)$$

with $\phi(N+1) = \phi(1)$. It is natural in one dimension to take for the ‘‘volume per edge’’ $V_1(l_n) = l_n$, and for the ‘‘volume per site’’ $V_0(l_n) = (l_n + l_{n-1})/2$. Here ω plays the role of a mass for the scalar

field, $\omega = m^2$. In addition one needs a term

$$\lambda L(l) = \lambda_0 \sum_{n=1}^N l_n \quad (481)$$

which is necessary in order to make the dl_n integration convergent at large l . Varying the action with respect to ϕ_n gives

$$\frac{2}{l_{n-1} + l_n} \left[\frac{\phi_{n+1} - \phi_n}{l_n} - \frac{\phi_n - \phi_{n-1}}{l_{n-1}} \right] = \omega \phi_n . \quad (482)$$

This is the discrete analog of the equation $g^{-1/2} \partial g^{-1/2} \partial \phi = \omega \phi$. The spectrum of the Laplacian of Eq. (482) corresponds to $\Omega \equiv -\omega > 0$. Variation with respect to l_n gives instead

$$\frac{1}{2l_n^2} (\phi_{n+1} - \phi_n)^2 = \lambda_0 + \frac{1}{4} \omega (\phi_n^2 + \phi_{n+1}^2) . \quad (483)$$

For $\omega = 0$ it suggests the well-known interpretation of the fields ϕ_n as coordinates in embedding space. In the following we shall only consider the case $\omega = 0$, corresponding to a massless scalar field.

It is instructive to look at the invariance properties of the scalar field action under the continuous lattice diffeomorphisms defined in Eq. (432). Physically, these local gauge transformations, which act on the vertices, correspond to re-assignments of edge lengths which leave the distance between two fixed points unchanged. In the simplest case, only two neighboring edge lengths are changed, leaving the total distance between the end points unchanged. On physical grounds one would like to maintain such an invariance also in the case of coupling to matter, just as is done in the continuum.

The scalar nature of the field requires that in the continuum under a change of coordinates $x \rightarrow x'$,

$$\phi'(x') = \phi(x) , \quad (484)$$

where x and x' refer to the same physical point in the two coordinate systems. On the lattice, as discussed previously, diffeomorphisms move the points around, and at the same vertex labeled by n we expect

$$\phi_n \rightarrow \phi'_n \approx \phi_n + \left(\frac{\phi_{n+1} - \phi_n}{l_n} \right) \epsilon_n , \quad (485)$$

One can determine the exact form of the change needed in ϕ_n by requiring that the local variation of the scalar field action be zero. Solving the resulting quadratic equation for $\Delta\phi_n$ one obtains a

rather unwieldy expression, given to lowest order by

$$\begin{aligned} \Delta\phi_n = & \frac{\epsilon_n}{2} \left[\frac{\phi_n - \phi_{n-1}}{l_{n-1}} + \frac{\phi_{n+1} - \phi_n}{l_n} \right] \\ & + \frac{\epsilon_n^2}{8} \left[-\frac{\phi_n - \phi_{n-1}}{l_{n-1}^2} + \frac{\phi_{n+1} - \phi_n}{l_n^2} + \frac{\phi_{n+1} - 2\phi_n + \phi_{n-1}}{l_{n-1}l_n} \right] + O(\epsilon_n^3) , \end{aligned} \quad (486)$$

and which is indeed of the expected form (as well as symmetric in the vertices $n-1$ and $n+1$). For fields which are reasonably smooth, this correction is suppressed if $|\phi_{n+1} - \phi_n|/l_n \ll 1$. On the other hand it should be clear that the measure $d\phi_n$ is no longer manifestly invariant, due to the rather involved transformation property of the scalar field.

The full functional integral for N sites then reads

$$\mathcal{Z}_N = \prod_{n=1}^N \int_0^\infty dl_n^2 l_n^\sigma \int_{-\infty}^\infty d\phi_n \exp \left\{ -\lambda_0 \sum_{n=1}^N l_n - \frac{1}{2} \sum_{n=1}^N \frac{1}{l_n} (\phi_{n+1} - \phi_n)^2 \right\} . \quad (487)$$

In the absence of the scalar field one just has the $\mathcal{Z}_N(\lambda_0)$ of Eq. (428). The trivial translational mode in ϕ can be eliminated for example by setting $\sum_{n=1}^N \phi_n = 0$.

It is possible to further constrain the measure over the edge lengths by examining some local averages. Under a rescaling of the edge lengths $l_n \rightarrow \alpha l_n$ one can derive the following identity for \mathcal{Z}_N

$$\mathcal{Z}_N(\lambda_0, z) = \lambda_0^{-(5/2+\sigma)N} z^{-N/2} \mathcal{Z}_N(1, 1) , \quad (488)$$

where we have replaced the coefficient $1/2$ of the scalar kinetic term by $z/2$. It follows then that

$$\langle l \rangle \equiv \frac{1}{N} \langle \sum_{n=1}^N l_n \rangle = \left(\frac{5}{2} + \sigma \right) \lambda_0^{-1} \quad (489)$$

and

$$\frac{1}{N} \langle \sum_{n=1}^N \frac{1}{l_n} (\phi_{n+1} - \phi_n)^2 \rangle = 1 . \quad (490)$$

Without loss of generality we can fix the average edge length to be equal to one, $\langle l \rangle = 1$, which then requires $\lambda_0 = \frac{5}{2} + \sigma$. In order for the model to be meaningful, the measure parameter is constrained by $\sigma > -5/2$, i.e. the measure over the edges cannot be too singular.

12. Lattice Fermions, Tetrads and Spin Rotations

On a simplicial manifold spinor fields ψ_s and $\bar{\psi}_s$ are most naturally placed at the center of each d -simplex s . In the following we will restrict our discussion for simplicity to the four-dimensional

case, and largely follow the original discussion given by (Fröhlich, 1981) and (Drummond 1986). As in the continuum (see for example Veltman, 1974), the construction of a suitable lattice action requires the introduction of the Lorentz group and its associated tetrad fields $e_\mu^a(s)$ within each simplex labeled by s .

Within each simplex one can choose a representation of the Dirac gamma matrices, denoted here by $\gamma^\mu(s)$, such that in the local coordinate basis

$$\{\gamma^\mu(s), \gamma^\nu(s)\} = 2g^{\mu\nu}(s) \quad (491)$$

These in turn are related to the ordinary Dirac gamma matrices γ^a , which obey

$$\{\gamma^a, \gamma^b\} = 2\eta^{ab} \quad , \quad (492)$$

by

$$\gamma^\mu(s) = e_a^\mu(s) \gamma^a \quad (493)$$

so that within each simplex the tetrads $e_\mu^a(s)$ satisfy the usual relation

$$e_a^\mu(s) e_b^\nu(s) \eta^{ab} = g^{\mu\nu}(s) \quad (494)$$

In general the tetrads are not fixed uniquely within a simplex, being invariant under the local Lorentz transformations discussed earlier in Sec. III.A.2.

In the continuum the action for a massless spinor field is given by

$$I = \int dx \sqrt{g} \bar{\psi}(x) \gamma^\mu D_\mu \psi(x) \quad (495)$$

where $D_\mu = \partial_\mu + \frac{1}{2}\omega_{\mu ab}\sigma^{ab}$ is the spinorial covariant derivative containing the spin connection $\omega_{\mu ab}$. It will be convenient to first consider only two neighboring simplices s_1 and s_2 , covered by a common coordinate system x^μ . When the two tetrads in s_1 and s_2 are made to coincide, one can then use a common set of gamma matrices γ^μ within both simplices. Then in the absence of torsion the covariant derivative D_μ in Eq. (495) reduces to just an ordinary derivative. The fermion field $\psi(x)$ within the two simplices can then be suitably interpolated, by writing for example

$$\psi(x) = \theta(n \cdot x) \psi(s_1) + (1 - \theta(n \cdot x)) \psi(s_2) \quad (496)$$

where n_μ is the common normal to the face $f(s_1, s_2)$ shared by the simplices s_1 and s_2 , and chosen to point into s_1 . Inserting the expression for $\psi(x)$ from Eq. (496) into Eq. (495) and applying the

divergence theorem (or equivalently using the fact that the derivative of a step function only has support at the origin) one obtains

$$I = \frac{1}{2} V^{(d-1)}(f) (\bar{\psi}_1 + \bar{\psi}_2) \gamma^\mu n_\mu (\psi_1 - \psi_2) \quad (497)$$

where $V^{(d-1)}(f)$ represents the volume of the $(d-1)$ -dimensional common interface f , a tetrahedron in four dimensions. But the contributions from the diagonal terms containing $\bar{\psi}_1\psi_1$ and $\bar{\psi}_2\psi_2$ vanish when summed over the faces of an n -simplex, by virtue of the useful identity

$$\sum_{p=1}^{n+1} V(f^{(p)}) n_\mu^{(p)} = 0 \quad (498)$$

where $V(f^{(p)})$ are the volumes of the p faces of a given simplex, and $n_\mu^{(p)}$ the inward pointing unit normals to those faces.

So far the above partial expression for the lattice spinor action was obtained by assuming that the tetrads $e_a^\mu(s_1)$ and $e_a^\mu(s_2)$ in the two simplices coincide. If they do not, then they will be related by a matrix $\mathbf{R}(s_2, s_1)$ such that

$$e_a^\mu(s_2) = R^\mu{}_\nu(s_2, s_1) e_a^\nu(s_1) \quad (499)$$

and whose spinorial representation \mathbf{S} was given previously for example in Eq. (354). Such a matrix $\mathbf{S}(s_2, s_1)$ is now needed to additionally parallel transport the spinor ψ from a simplex s_1 to the neighboring simplex s_2 .

The invariant lattice action for a massless spinor takes therefore the form

$$I = \frac{1}{2} \sum_{\text{faces } f(ss')} V(f(s, s')) \bar{\psi}_s \mathbf{S}(\mathbf{R}(s, s')) \gamma^\mu(s') n_\mu(s, s') \psi_{s'} \quad (500)$$

where the sum extends over all interfaces $f(s, s')$ connecting one simplex s to a neighboring simplex s' . As shown in (Drummond 1986) it can be further extended to include a dynamical torsion field.

The above spinorial action can be considered analogous to the lattice Fermion action proposed originally in (Wilson 1973) for non-Abelian gauge theories. It is possible that it still suffers from the Fermion doubling problem, although the situation is less clear for a dynamical lattice (Nielsen and Ninomiya, 1981; Christ and Lee, 1982).

13. Alternate Discrete Formulations

The simplicial lattice formulation offers a natural way of representing gravitational degrees in a discrete framework by employing inherently geometric concepts such as areas, volumes and angles.

It is possible though to formulate quantum gravity on a flat hypercubic lattice, in analogy to Wilson's discrete formulation for gauge theories, by putting the connection centerstage. In this new set of theories the natural variables are then lattice versions of the spin connection and the vierbein. Also, because the spin connection variables appear from the very beginning, it is much easier to incorporate fermions later. Some lattice models have been based on the pure Einstein theory (Smolin, 1979; Das, Kaku and Townsend, 1979; Mannion and Taylor, 1981; Caracciolo and Pelissetto, 1987), while others attempt to incorporate higher derivative terms (Tomboulis 1984; Kondo 1984).

Difficult arise when attempting to put quantum gravity on a flat hypercubic lattice a la Wilson, since it is not entirely clear what the gravity analogue of the Yang-Mills connection is. In continuum formulations invariant under the Poincaré or de Sitter group the action is invariant under a local extension of the Lorentz transformations, but not under local translations (Kibble, 1961). Local translations are replaced by diffeomorphisms which have a different nature. One set of lattice discretizations starts from the action of (MacDowell and Mansouri, 1977; West, 1978) whose local invariance group is the de Sitter group $Sp(4)$. In the lattice formulation of (Smolin, 1979; Das, Kaku and Townsend, 1979) the lattice variables are gauge potentials $e_{a\mu}(n)$ and $\omega_{\mu ab}(n)$ defined on lattice sites n , generating local $Sp(4)$ matrix transformations with the aid of the de Sitter generators P_a and M_{ab} . The resulting lattice action reduces classically to the Einstein action with cosmological term in first order form in the limit of the lattice spacing $a \rightarrow 0$; to demonstrate the quantum equivalence one needs an additional zero torsion constraint. In the end the issue of lattice diffeomorphism invariance remains somewhat open, with the hope that such an invariance will be restored in the full quantum theory.

As an example, we will discuss here the approach of (Mannion and Taylor, 1981) which relies on a four-dimensional lattice discretization of the Einstein-Cartan theory with gauge group $SL(2, C)$, and does not initially require the presence of a cosmological constant, as would be the case if one had started out with the de Sitter group $Sp(4)$. On a lattice of spacing a with vertices labelled by n and directions by μ one relates the relative orientations of nearest-neighbor local $SL(2, C)$ frames by

$$U_\mu(n) = [U_{-\mu}(n + \mu)]^{-1} = \exp[i B_\mu(n)] \quad (501)$$

with $B_\mu = \frac{1}{2}aB_\mu^{ab}(n)J_{ba}$, J_{ba} being the set of six generators of $SL(2, C)$, the covering group of the Lorentz group $SO(3, 1)$, usually taken to be

$$\sigma_{ab} = \frac{1}{2i}[\gamma_a, \gamma_b] \quad (502)$$

with γ_a 's the Dirac gamma matrices. The local lattice curvature is then obtained in the usual way by computing the product of four parallel transport matrices around an elementary lattice square,

$$U_\mu(n) U_\nu(n + \mu) U_{-\mu}(n + \mu + \nu) U_{-\nu}(n + \nu) \quad (503)$$

giving in the limit of small a by the Baker-Hausdorff formula the value $\exp[iaR_{\mu\nu}(n)]$, where $R_{\mu\nu}$ is the Riemann tensor defined in terms of the spin connection B_μ

$$R_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu + i[B_\mu, B_\nu] \quad (504)$$

If one were to write for the action the usual Wilson lattice gauge form

$$\sum_{n,\mu,\nu} \text{tr}[U_\mu(n) U_\nu(n + \mu) U_{-\mu}(n + \mu + \nu) U_{-\nu}(n + \nu)] \quad (505)$$

then one would obtain a curvature squared action proportional to $\sim \int R_{\mu\nu}^{ab} R_{ab}^{\mu\nu}$ instead of the Einstein-Hilbert one. One needs therefore to introduce lattice vierbeins $e_\mu^b(n)$ on the sites by defining the matrices

$$E_\mu(n) = a e_\mu^a \gamma_a \quad (506)$$

Then a suitable lattice action is given by

$$I = \frac{i}{16\kappa^2} \sum_{n,\mu,\nu,\lambda,\sigma} \text{tr}[\gamma_5 U_\mu(n) U_\nu(n + \mu) U_{-\mu}(n + \mu + \nu) U_{-\nu}(n + \nu) E_\sigma(n) E_\lambda(n)] \quad (507)$$

The latter is invariant under local $SL(2, C)$ transformations $\Lambda(n)$ defined on the lattice vertices

$$U_\mu \rightarrow \Lambda(n) U_\mu(n) \Lambda^{-1}(n + \mu) \quad (508)$$

for which the curvature transforms as

$$\begin{aligned} & U_\mu(n) U_\nu(n + \mu) U_{-\mu}(n + \mu + \nu) U_{-\nu}(n + \nu) \\ & \rightarrow \Lambda(n) U_\mu(n) U_\nu(n + \mu) U_{-\mu}(n + \mu + \nu) U_{-\nu}(n + \nu) \Lambda^{-1}(n) \end{aligned} \quad (509)$$

and the vierbein matrices as

$$E_\mu(n) \rightarrow \Lambda(n) E_\mu(n) \Lambda^{-1}(n) \quad (510)$$

Since $\Lambda(n)$ commutes with γ_5 , the expression in Eq. (507) is invariant. The metric is then obtained as usual by

$$g_{\mu\nu}(n) = \frac{1}{4} \text{tr}[E_\mu(n) E_\nu(n)] \quad (511)$$

From the expression for the lattice curvature $R_{\mu\nu}^{ab}$ given above it follows immediately that the lattice action in the continuum limit becomes

$$I = \frac{a^4}{4\kappa^2} \sum_n \epsilon^{\mu\nu\lambda\sigma} \epsilon_{abcd} R_{\mu\nu}^{ab}(n) e_\lambda^c(n) e_\sigma^d(n) + O(a^6) \quad (512)$$

which is the Einstein action in Cartan form

$$I = \frac{1}{4\kappa^2} \int d^4x \epsilon^{\mu\nu\lambda\sigma} \epsilon_{abcd} R_{\mu\nu}^{ab} e_\lambda^c e_\sigma^d \quad (513)$$

with the parameter κ identified with the Planck length. One can add more terms to the action; in this theory a cosmological term can be represented by

$$\lambda_0 \sum_n \epsilon^{\mu\nu\lambda\sigma} \text{tr}[\gamma_5 E_\mu(n) E_\nu(n) E_\sigma(n) E_\lambda(n)] \quad (514)$$

Both Eqs. (507) and Eq. (514) are locally $SL(2, C)$ invariant. The functional integral is then given by

$$Z = \int \prod_{n,\mu} dB_\mu(n) \prod_{n,\sigma} dE_\sigma(n) \exp\{-I(B, E)\} \quad (515)$$

and from it one can then compute suitable quantum averages. Here $dB_\mu(n)$ is the Haar measure for $SL(2, C)$; it is less clear how to choose the integration measure over the E_σ 's, and how it should suitably be constrained, which obscures the issue of diffeomorphism invariance in this theory.

In these theories it is possible to formulate curvature squared terms as well. In general for a hypercubic lattice the formulation of R^2 -type terms in four dimensions involves constraints between the connections and the tetrads, which are a bit difficult to handle. Also there is no simple way of writing down topological invariants, which are either related to the Einstein action (in two dimensions), or are candidates for extra terms to be included in the action. A flat hypercubic lattice action has been written with higher derivative terms which appears to be reflection positive but has a very cumbersome form. These difficulties need not be present on a simplicial lattice (except that it is not known how to write an exact expression for the Hirzebruch signature in lattice terms).

There is another way of discretizing gravity, still using largely geometric concepts as is done in the Regge theory. In the dynamical triangulation approach one fixes the edge lengths to unity, and varies the incidence matrix. As a result the volume of each simplex is fixed at

$$V_d = \frac{1}{d!} \sqrt{\frac{d+1}{2^d}} \quad (516)$$

and all dihedral angles are given by the constant value

$$\cos \theta_d = \frac{1}{d} \tag{517}$$

so that for example in four dimensions one has $\theta_d = \arccos(1/4) \approx 75.5^\circ$. Local curvatures are then determined by how many simplices $n_s(h)$ meet on a given hinge,

$$\delta(h) = 2\pi - n_s(h)\theta_d \tag{518}$$

The action contribution from a single hinge is therefore from Eq. (368) $\delta(h)A(h) = \frac{1}{4}\sqrt{3}[2\pi - n_s(h)\theta_d]$ with n_s a positive integer. The local curvatures are inherently discrete, and there is no equivalent lattice notion of continuous diffeomorphisms, or for that matter of continuous local deformations corresponding for example to shear waves. The hope is that for lattices made of some large number of simplices one recovers some sort of discrete version of diffeomorphism invariance. The Euclidean dynamical triangulation approach has been reviewed recently in (Ambjørn, Carfora and Marzuoli, 1999), and we refer the reader to further references therein. A recent discussion of attempts at simulating the Lorentzian case, which leads to complex weights in the functional integral which are difficult to handle, can be found in (Loll et al, 2006).

Another lattice approach closely related to the Regge theory described in this review is based on the so-called spin foam models, which have their origin in an observation found in (Ponzano and Regge, 1968) relating the geometry of simplicial lattices to the asymptotics of Racah angular momentum addition coefficients. The original Regge-Ponzano concepts were later developed into a spin model for gravity (Hasslacher and Perry, 1981) based on quantum spin variables attached to lattice links. In these models representations of $SU(2)$ label edges. One natural underlying framework for such theories is the canonical $3+1$ approach to quantum gravity, wherein quantum spin variables are naturally related to $SU(2)$ spin connections. Extensions to four dimensions have been attempted, and we refer the reader to the recent review of spin foam models in (Perez, 2003).

B. Analytical Expansion Methods

The following sections will discuss a number of instances where the lattice theory of quantum gravity can be investigated analytically, subject to some simplifying assumptions.

The first problem is the lattice weak field expansion about a flat background. It will be shown that in this case the relevant modes are the lattice analogs of transverse-traceless deformations.

The second problem is the strong coupling (large G) expansion, where the weight factor in

the path integral is expanded in powers of $1/G$. The domain of validity of this expansion can be regarded as somewhat complementary to the weak field limit.

The third case to be discussed is what happens in lattice gravity in the limit of large dimensions d , which formally is similar in some ways to the large- N expansion discussed previously in this review. In this limit one can derive exact estimates for the phase transition point and for the scaling dimensions.

1. Lattice Weak Field Expansion and Transverse-Traceless Modes

One of the simplest possible problems that can be treated in quantum Regge calculus is the analysis of small fluctuations about a fixed flat Euclidean simplicial background (Roček and Williams, 1981). In this case one finds that the lattice graviton propagator in a De Donder-like gauge is precisely analogous to the continuum expression.

To compute an expansion of the lattice Regge action

$$I_R \propto \sum_{\text{hinges}} \delta(l) A(l) \quad (519)$$

to quadratic order in the lattice weak fields one needs first and second variations with respect to the edge lengths. In four dimensions the first variation of the lattice Regge action is given by

$$\delta I_R \propto \sum_{\text{hinges}} \delta \cdot \left(\sum_{\text{edges}} \frac{\partial A}{\partial l} \delta l \right) \quad (520)$$

since Regge has shown that the term involving the variation of the deficit angle δ vanishes (here the variation symbol should obviously not be confused with the deficit angle). Furthermore in flat space all the deficit angles vanish, so that the second variation is given simply by

$$\delta^2 I_R \propto \sum_{\text{hinges}} \left(\sum_{\text{edges}} \frac{\partial \delta}{\partial l} \delta l \right) \cdot \left(\sum_{\text{edges}} \frac{\partial A}{\partial l} \delta l \right) \quad (521)$$

Next a specific lattice structure needs to be chosen as a background geometry. A natural choice is to use a flat hypercubic lattice, made rigid by introducing face diagonals, body diagonals and hyperbody diagonals, which results into a subdivision of each hypercube into $d!$ (here $4!=24$) simplices. This subdivision is shown in Fig. 26.

By a simple translation, the whole lattice can then be constructed from this one elemental hypercube. Consequently there will be $2^d - 1 = 15$ lattice fields per point, corresponding to all the edge lengths emanating in the positive lattice directions from any one vertex. Note that the number of degrees per lattice point is slightly larger than what one would have in the continuum, where

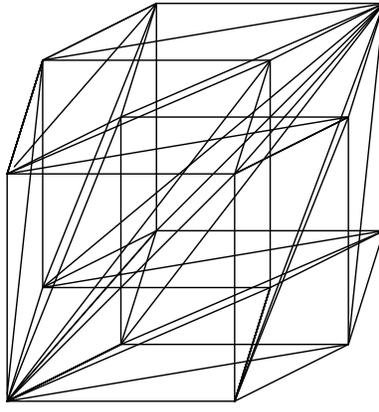


FIG. 26 A four-dimensional hypercube divided up into four-simplices.

the metric $g_{\mu\nu}(x)$ has $d(d+1)/2 = 10$ degrees of freedom per spacetime point x in four dimensions (perturbatively, the physical degrees of freedom in the continuum are much less: $\frac{1}{2}d(d+1) - 1 - d - (d-1) = \frac{1}{2}d(d-3)$, for a traceless symmetric tensor, and after imposing gauge conditions). Thus in four dimensions each lattice hypercube will contain 4 body principals, 6 face diagonals, 4 body diagonals and one hyperbody diagonal. Within a given hypercube it is quite convenient to label the coordinates of the vertices using a binary notation, so that the four body principals with coordinates $(1, 0, 0, 0)$ $(0, 0, 0, 1)$ will be labeled by integers 1,2,4,8, and similarly for the other vertices (thus for example the vertex $(0, 1, 1, 0)$, corresponding to a face diagonal along the second and third Cartesian direction, will be labeled by the integer 6).

For a given lattice of fixed connectivity, the edge lengths are then allowed to fluctuate around an equilibrium value l_i^0

$$l_i = l_i^0 (1 + \epsilon_i) \quad (522)$$

In the case of the hypercubic lattice subdivided into simplices, the unperturbed edge lengths l_i^0 take on the values $1, \sqrt{2}, \sqrt{3}, 2$, depending on edge type. The second variation of the action then reduces to a quadratic form in the 15-component small fluctuation vector ϵ_n

$$\delta^2 I_R \propto \sum_{mn} \epsilon_m^T M_{mn} \epsilon_n \quad (523)$$

Here M is the small fluctuation matrix, whose inverse determines the free lattice graviton propagator, and the indices m and n label the sites on the lattice. But just as in the continuum, M has zero eigenvalues and cannot therefore be inverted until one supplies an appropriate gauge condition. Specifically, one finds that the matrix M in four dimensions has four zero modes corresponding to periodic translations of the lattice, and a fifth zero mode corresponding to periodic fluctuations in

the hyperbody diagonal. After block-diagonalization it is found that 4 modes completely decouple and are constrained to vanish, and thus the remaining degrees of freedom are 10, as in the continuum, where the metric has 10 independent components. The wrong sign for the conformal mode, which is present in the continuum, is also reproduced by the lattice propagator.

Due to the locality of the original lattice action, the matrix M can be considered local as well, since it only couples edge fluctuations on neighboring lattice sites. In Fourier space one can write for each of the fifteen displacements $\epsilon_n^{i+j+k+l}$, defined at the vertex of the hypercube with labels (i, j, k, l) ,

$$\epsilon_n^{i+j+k+l} = (\omega_1)^i (\omega_2)^j (\omega_4)^k (\omega_8)^l \epsilon_n^0 \quad (524)$$

with $\omega_1 = e^{ik_1}$, $\omega_2 = e^{ik_2}$, $\omega_4 = e^{ik_3}$ and $\omega_8 = e^{ik_4}$ (it will be convenient in the following to use binary notation for ω and ϵ , but the regular notation for k_i). Here and in the following we have set the lattice spacing a equal to one.

In this basis the matrix M reduces to a block-diagonal form, with entries given by the 15×15 dimensional matrices

$$M_\omega = \begin{pmatrix} A_{10} & B & 0 \\ B^\dagger & 18I_4 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (525)$$

where A_{10} is a 10×10 dimensional matrix, B a 10×4 dimensional matrix and I_4 is the 4×4 dimensional identity matrix. Explicitly the above cited authors find

$$\begin{aligned} (M_\omega)_{1,1} &= 6 \\ (M_\omega)_{1,2} &= \omega_1(\omega_4 + \omega_8) + \bar{\omega}_2(\bar{\omega}_4 + \bar{\omega}_8) \\ (M_\omega)_{1,3} &= 2 + 2\bar{\omega}_2 \\ (M_\omega)_{1,6} &= 2\omega_1 + 2\bar{\omega}_2\bar{\omega}_4 \\ (M_\omega)_{1,7} &= \bar{\omega}_2 + \bar{\omega}_4 \\ (M_\omega)_{1,14} &= 0 \\ (M_\omega)_{3,3} &= 4 \\ (M_\omega)_{3,5} &= \omega_2 + \bar{\omega}_4 \\ (M_\omega)_{3,12} &= 0 \\ (M_\omega)_{3,7} &= 1 + \bar{\omega}_4 \\ (M_\omega)_{3,13} &= 0 \end{aligned} \quad (526)$$

where the remaining non-vanishing matrix elements can be obtained either by permuting appropriate indices, or by complex conjugation.

Besides one obvious zero eigenvalue, corresponding to a periodic fluctuation in ϵ_{15} , the matrix M_ω exhibits four additional zero modes corresponding to the four-parameter group of translations in flat space. An explicit form for these eigenmodes is

$$\begin{aligned}
\epsilon_i &= (1 - \omega_i) x_i \\
\epsilon_{i+j} &= \frac{1}{2}(1 - \omega_i \omega_j) (x_i + x_j) \\
\epsilon_{i+j+k} &= \frac{1}{3}(1 - \omega_i \omega_j \omega_k) (x_i + x_j + x_k) \\
\epsilon_{i+j+k+l} &= \frac{1}{4}(1 - \omega_i \omega_j \omega_k \omega_l) (x_i + x_j + x_k + x_l)
\end{aligned} \tag{527}$$

with $i, j, k, l = 1, 2, 4, 8$ and $i \neq j \neq k \neq l$.

The next step consists in transforming the lattice action M_ω into a form more suitable for comparison with the continuum action. To this end a set of transformations are performed sequentially, the first of which involves the matrix

$$S = \begin{pmatrix} I_{10} & 0 & 0 \\ -\frac{1}{18}B^\dagger & I_4 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{528}$$

which rotates M_ω into

$$M'_\omega = S^\dagger M_\omega S = \begin{pmatrix} A_{10} - \frac{1}{18}BB^\dagger & 0 & 0 \\ 0 & 18I_4 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{529}$$

thus completely decoupling the (body diagonal) fluctuations $\epsilon_7, \epsilon_{11}, \epsilon_{13}, \epsilon_{14}$. These in turn can now be integrated out, as they appear in the action with no ω (i.e. derivative) term. As a result the number of dynamical degrees of freedom has been reduced from 15 to 10, the same number as in the continuum.

The remaining dynamics is thus encoded in the 10×10 dimensional matrix $L_\omega = A_{10} - \frac{1}{18}BB^\dagger$. By a second rotation, here affected by the matrix T , it can finally be brought into the form

$$\tilde{L}_\omega = T^\dagger L_\omega T = [8 - (\Sigma + \bar{\Sigma})] \begin{pmatrix} \frac{1}{2}\beta & 0 \\ 0 & I_6 \end{pmatrix} - C^\dagger C \tag{530}$$

with the matrix β given by

$$\beta = \frac{1}{2} \begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix} \quad (531)$$

The other matrix C appearing in the second term is given by

$$C = \begin{pmatrix} f_1 & 0 & 0 & 0 & \tilde{f}_2 & \tilde{f}_4 & 0 & \tilde{f}_8 & 0 & 0 \\ 0 & f_2 & 0 & 0 & \tilde{f}_1 & 0 & \tilde{f}_4 & 0 & \tilde{f}_8 & 0 \\ 0 & 0 & f_4 & 0 & 0 & \tilde{f}_1 & \tilde{f}_2 & 0 & 0 & \tilde{f}_8 \\ 0 & 0 & 0 & f_8 & 0 & 0 & 0 & \tilde{f}_1 & \tilde{f}_2 & \tilde{f}_4 \end{pmatrix} \quad (532)$$

with $f_i \equiv \omega_i - 1$ and $\tilde{f}_i \equiv 1 - \bar{\omega}_i$. Furthermore $\Sigma = \sum_i \omega_i$, and for small momenta one finds

$$8 - (\Sigma + \bar{\Sigma}) = 8 - \sum_{i=1}^4 (e^{ik_i} + e^{-ik_i}) \sim k^2 + O(k^4) \quad (533)$$

which shows that the surviving terms in the lattice action are indeed quadratic in k . The rotation matrix T involved in the last transformation is given by

$$T = \begin{pmatrix} \Omega_4 \beta & 0 \\ 0 & I_6 \end{pmatrix} \begin{pmatrix} I_4 & 0 \\ \Omega_6 \gamma & I_6 \end{pmatrix} \quad (534)$$

with $\Omega_4 = \text{diag}(\omega_1, \omega_2, \omega_4, \omega_8)$ and $\Omega_6 = \text{diag}(\omega_1\omega_2, \omega_1\omega_4, \omega_2\omega_4, \omega_1\omega_8, \omega_2\omega_8, \omega_4\omega_8)$, and

$$\gamma = -\frac{1}{2} \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \quad (535)$$

At this point one is finally ready for a comparison with the continuum result, namely with the Lagrangian for pure gravity in the weak field limit as given in Eq. (7)

$$\begin{aligned} \mathcal{L}_{sym} = & -\frac{1}{2} \partial_\lambda h_{\lambda\mu} \partial_\mu h_{\nu\nu} + \frac{1}{2} \partial_\lambda h_{\lambda\mu} \partial_\nu h_{\nu\mu} \\ & - \frac{1}{4} \partial_\lambda h_{\mu\nu} \partial_\lambda h_{\mu\nu} + \frac{1}{4} \partial_\lambda h_{\mu\mu} \partial_\lambda h_{\nu\nu} \end{aligned} \quad (536)$$

The latter can be conveniently split into two parts, as was done already in Eq. (52), as follows

$$\mathcal{L}_{sym} = -\frac{1}{2} \partial_\lambda h_{\alpha\beta} V_{\alpha\beta\mu\nu} \partial_\lambda h_{\mu\nu} + \frac{1}{2} C^2 \quad (537)$$

with

$$V_{\alpha\beta\mu\nu} = \frac{1}{2}\eta_{\alpha\mu}\eta_{\beta\nu} - \frac{1}{4}\eta_{\alpha\beta}\eta_{\mu\nu} \quad (538)$$

or as a matrix,

$$V = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad (539)$$

with metric components 11, 22, 33, 44, 12, 13, 14, 23, 24, 34 more conveniently labeled sequentially by integers 1...10, and the gauge fixing term C_μ given by the term in Eq. (53)

$$C_\mu = \partial_\nu h_{\mu\nu} - \frac{1}{2}\partial_\mu h_{\nu\nu} \quad (540)$$

The above expression is still not quite the same as the lattice weak field action, but a simple transformation to trace reversed variables $\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}h_{\lambda\lambda}$ leads to

$$\mathcal{L}_{sym} = \frac{1}{2}k_\lambda \bar{h}_i V_{ij} k_\lambda \bar{h}_j - \frac{1}{2}\bar{h}_i (C^\dagger C)_{ij} \bar{h}_j \quad (541)$$

with the matrix V given by

$$V_{ij} = \begin{pmatrix} \frac{1}{2}\beta & 0 \\ 0 & I_6 \end{pmatrix} \quad (542)$$

with $k = i\partial$. Now β is the same as the matrix in Eq. (531), and C is nothing but the small k limit of the matrix by the same name in Eq. (532), for which one needs to set $\omega_i - 1 \simeq i k_i$. The resulting continuum expression is then recognized to be identical to the lattice weak field results of Eq. (530).

This concludes the outline of the proof of equivalence of the lattice weak field expansion of the Regge action to the corresponding continuum expression. To summarize, there are several ingredients to this proof, the first of which is a relatively straightforward weak field expansion of both actions, and the second of which is the correct identification of the lattice degrees of

freedom $\epsilon_i(n)$ with their continuum counterparts $h_{\mu\nu}(x)$, which involves a sequence of non-trivial ω -dependent transformations, expressed by the matrices S and T . One more important aspect of the process is the disappearance of redundant lattice variables (five in the case of the hypercubic lattice), whose dynamics turns out to be trivial, in the sense that the associated degrees of freedom are non-propagating.

It is easy to see that the sequence of transformations expressed by the matrices S of Eq. (528) and T of Eq. (534), and therefore ultimately relating the lattice fluctuations $\epsilon_i(n)$ to their continuum counterparts $h_{\mu\nu}(x)$, just reproduces the expected relationship between lattice and continuum fields (Hamber and Williams, 1993). On the one hand one has $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, where $\eta_{\mu\nu}$ is the flat metric. At the same time one has from Eq. (333) for each simplex within a given hypercube

$$g_{ij} = \frac{1}{2}(l_{0i}^2 + l_{0j}^2 - l_{ij}^2) \quad (543)$$

By inserting $l_i = l_i^0 (1 + \epsilon_i)$, with $l_i^0 = 1, \sqrt{2}, \sqrt{3}, 2$ for the body principal ($i = 1, 2, 4, 8$), face diagonal ($i = 3, 5, 6, 9, 10, 12$), body diagonal ($i = 7, 11, 13, 14$) and hyperbody diagonal ($i = 15$), respectively, one gets for example $(1 + \epsilon_1)^2 = 1 + h_{11}$, $(1 + \epsilon_3)^2 = 1 + \frac{1}{2}(h_{11} + h_{22}) + h_{12}$ etc., which in turn can then be solved for the ϵ 's in terms of the $h_{\mu\nu}$'s. One would then obtain

$$\begin{aligned} \epsilon_1 &= -1 + [1 + h_{11}]^{1/2} \\ \epsilon_3 &= -1 + [1 + \frac{1}{2}(h_{11} + h_{22}) + h_{12}]^{1/2} \\ \epsilon_7 &= -1 + [1 + \frac{1}{3}(h_{11} + h_{22} + h_{33}) \\ &\quad + \frac{2}{3}(h_{12} + h_{23} + h_{13})]^{1/2} \\ \epsilon_{15} &= -1 + [1 + \frac{1}{4}(h_{11} + h_{22} + h_{33} + h_{44}) \\ &\quad + \frac{3}{4}(h_{12} + h_{13} + h_{14} + h_{23} + h_{24} + h_{34})]^{1/2} \end{aligned} \quad (544)$$

and so on for the other edges, by suitably permuting indices. These relations can then be expanded out for weak h , giving for example

$$\begin{aligned} \epsilon_1 &= \frac{1}{2} h_{11} + O(h^2) \\ \epsilon_3 &= \frac{1}{2} h_{12} + \frac{1}{4} (h_{11} + h_{22}) + O(h^2) \\ \epsilon_7 &= \frac{1}{6} (h_{12} + h_{13} + h_{23}) + \frac{1}{6} (h_{23} + h_{13} + h_{12}) \\ &\quad + \frac{1}{6} (h_{11} + h_{22} + h_{33}) + O(h^2) \end{aligned} \quad (545)$$

and so on. The above correspondence between the ϵ 's and the $h_{\mu\nu}$ are the underlying reason for the existence of the rotation matrices S and T of Eqs. (528) and (534), with one further important amendment: on the hypercubic lattice four edges within a given simplex are assigned to one vertex, while the remaining six edges are assigned to neighboring vertices, and require therefore a translation back to the base vertex of the hypercube, using the result of Eq. (524). This explains the additional factors of ω appearing in the rotation matrices S and T . More importantly, one would expect such a combined rotation to be independent of what particular term in the lattice action one is considering, implying that it can be used to relate other lattice gravity contributions, such as the cosmological term and higher derivative terms, to their continuum counterparts (Hamber and Williams, 1993).

The choice of gauge in Eq. (540) is of course not the only possible one. Another possible choice is the so-called vacuum gauge for which in the continuum $h_{ik,k} = 0$, $h_{00} = h_{0i} = 0$. Expressed in terms of the lattice small fluctuation variables such a condition reads in momentum space

$$\begin{aligned}
e_8 &= 0 \\
e_9 &= \frac{1}{2} \omega_8 e_1 \\
e_{10} &= \frac{1}{2} \omega_8 e_2 \\
e_{12} &= \frac{1}{2} \omega_8 e_4 \\
e_{11} &= \frac{1}{3} (1 + \omega_8) e_3 - \frac{1}{6} (1 - \omega_8) (\omega_2 e_1 + \omega_1 e_2) \\
e_{13} &= \frac{1}{3} (1 + \omega_8) e_5 - \frac{1}{6} (1 - \omega_8) (\omega_4 e_1 + \omega_1 e_4) \\
e_{14} &= \frac{1}{3} (1 + \omega_8) e_6 - \frac{1}{6} (1 - \omega_8) (\omega_2 e_4 + \omega_4 e_2)
\end{aligned} \tag{546}$$

One can then evaluate the lattice action in such a gauge and again compare to the continuum expression. First one expands again the e_i 's in terms of the h_{ij} 's, as given in Eq. (544), and then expand out the ω 's in powers of k . If one then sets $k_4 = 0$ one finds that the resulting contribution can be re-written as the sum of two parts (Hamber and Williams, 2005), the first part being the transverse-traceless contribution

$$\frac{1}{4} \mathbf{k}^2 \text{Tr} [{}^3h(P {}^3hP - \frac{1}{2} P \text{Tr} (P {}^3h))] = \frac{1}{4} \mathbf{k}^2 \bar{h}_{ij}^{TT}(\mathbf{k}) h_{ij}^{TT}(\mathbf{k}) \tag{547}$$

$$\bar{h}_{ij}^{TT} h_{ij}^{TT} \equiv \text{Tr} [{}^3h(P {}^3hP - \frac{1}{2} P \text{Tr} (P {}^3h))] \tag{548}$$

with $P_{ij} = \delta_{ij} - k_i k_j / \mathbf{k}^2$ acting on the three-metric ${}^3h_{ij}$, and the second part arising due to the trace component of the metric

$$- \frac{1}{4} \mathbf{k}^2 \text{Tr} [P \text{Tr} (P {}^3h) P \text{Tr} (P {}^3h)] = \mathbf{k}^2 \bar{h}_{ij}^T(\mathbf{k}) h_{ij}^T(\mathbf{k}) \tag{549}$$

with $h^T = \frac{1}{2}P \text{Tr} (P^3 h)$. In the vacuum gauge $h_{ik,k} = 0$, $h_{ii} = 0$, $h_{0i} = 0$ one can further solve for the metric components h_{12} , h_{13} , h_{23} and h_{33} in terms of the two remaining degrees of freedom, h_{11} and h_{22} ,

$$\begin{aligned}
h_{12} &= -\frac{1}{2k_1 k_2} (h_{11} k_1^2 + h_{22} k_2^2 + h_{11} k_3^2 + h_{22} k_3^2) \\
h_{13} &= -\frac{1}{2k_1 k_3} (h_{11} k_1^2 - h_{22} k_2^2 - h_{11} k_3^2 - h_{22} k_3^2) \\
h_{23} &= -\frac{1}{2k_2 k_3} (-h_{11} k_1^2 + h_{22} k_2^2 - h_{11} k_3^2 - h_{22} k_3^2) \\
h_{33} &= -h_{11} - h_{22}
\end{aligned} \tag{550}$$

and show that the second (trace) part vanishes.

The above manipulations underscore the fact that the lattice action, in the weak field limit and for small momenta, only propagates transverse-traceless modes, as for linearized gravity in the continuum. They can be used to derive an expression for the lattice analog of the result given in (Kuchar, 1972) and (Hartle, 1982) for the vacuum wave functional of linearized gravity, which gives therefore a suitable starting point for a lattice candidate for the same functional.

A cosmological constant term can be analyzed in the lattice weak field expansion along similar lines. According to Eqs. (371) or (372) it is given on the lattice by the total space-time volume, so that the action contribution is given by

$$I_V = \lambda_0 \sum_{\text{edges } h} V_h, \tag{551}$$

where V_h is defined to be the volume associated with an edge h . The latter is obtained by subdividing the volume of each four-simplex into contributions associated with each hinge (here via a barycentric subdivision), and then adding up the contributions from each four-simplex touched by the given hinge. Expanding out in the small edge fluctuations one has

$$I_V \sim \sum_n (\epsilon_1^{(n)} + \epsilon_2^{(n)} + \epsilon_4^{(n)} + \epsilon_8^{(n)}) + \frac{1}{2} \sum_{mn,ij} \epsilon_i^{(m)T} M_{i,j}^{(m,n)} \epsilon_j^{(n)} \tag{552}$$

One needs to be careful since the expansion of ϵ_i in terms of $h_{\mu\nu}$ contains terms quadratic in $h_{\mu\nu}$, so that there are additional diagonal contributions to the small fluctuation matrix L_ω ,

$$\epsilon_1 + \epsilon_2 + \epsilon_4 + \epsilon_8 = \frac{1}{2} (h_{11} + h_{22} + h_{33} + h_{44}) - \frac{1}{8} (h_{11}^2 + h_{22}^2 + h_{33}^2 + h_{44}^2) + \dots \tag{553}$$

These additional contributions are required for the volume term to reduce to the continuum form of Eq. (41) for small momenta and to quadratic order in the weak field expansion.

Next the same set of rotations needs to be performed as for the Einstein term, in order to go from the lattice variables ϵ_i to the continuum variables $\bar{h}_{\mu\nu}$. After the combined S_ω - and T_ω -matrix rotations of Eqs. (528) and (534) one obtains for the small fluctuation matrix L_ω arising from the gauge-fixed lattice Einstein-Regge term [see Eq. (530)]

$$L_\omega = -\frac{1}{2} k^2 V, \quad (554)$$

with the matrix V given by Eq. (539). Since the lattice cosmological term can also be expressed in terms of the matrix V ,

$$\sqrt{g} = 1 + \frac{1}{2} h_{\mu\mu} - \frac{1}{2} h_{\alpha\beta} V^{\alpha\beta\mu\nu} h_{\mu\nu} + O(h^3), \quad (555)$$

one finds, as in the continuum, for the combined Einstein and cosmological constant terms

$$\lambda_0 \left(1 + \frac{1}{2} h_{\mu\mu} \right) + \frac{1}{2} \cdot \frac{k}{2} h_{\alpha\beta} V^{\alpha\beta\mu\nu} \left(\partial^2 - \frac{2\lambda_0}{k} \right) h_{\mu\nu} + O(h^3) , \quad (556)$$

corresponding in this gauge to the exchange of a particle of "mass" $\mu^2 = 2\lambda_0/k$, in agreement with the continuum weak field result of Eq. (64). As for the Regge-Einstein term, there are higher order lattice corrections to the cosmological constant term of $O(k)$ (which are completely absent in the continuum, since no derivatives are present there). These should be irrelevant in the lattice continuum limit.

2. Strong Coupling Expansion

In this section the strong coupling (large G or small $k = 1/(8\pi G)$) expansion of the lattice gravitational functional integral will be discussed. The resulting series is expected to be useful up to some $k = k_c$, where k_c is the lattice critical point, at which the partition function develops a singularity.

There will be two main aspects to the following discussion. The first aspect will be the development of a systematic expansion for the partition function and the correlation functions in powers of k , and a number of rather general considerations that follow from it. The second main aspect will be a detailed analysis and interpretation of the individual terms which appear order by order in the strong coupling expansion. This second part will lead to a later discussion of what happens for large d .

One starts from the lattice regularized path integral with action Eq. (373) and measure Eq. (406). In the following we will focus at first on the four-dimensional case. Then the four-dimensional Euclidean lattice action contains the usual cosmological constant and Regge scalar

curvature terms of Eq. (420)

$$I_{latt} = \lambda \sum_h V_h(l^2) - k \sum_h \delta_h(l^2) A_h(l^2) , \quad (557)$$

with $k = 1/(8\pi G)$, and possibly additional higher derivative terms as well. The action only couples edges which belong either to the same simplex or to a set of neighboring simplices, and can therefore be considered as *local*, just like the continuum action. It leads to a lattice partition function defined in Eq. (421)

$$Z_{latt} = \int [dl^2] e^{-\lambda_0 \sum_h V_h + k \sum_h \delta_h A_h} , \quad (558)$$

where, as customary, the lattice ultraviolet cutoff is set equal to one (i.e. all length scales are measured in units of the lattice cutoff). For definiteness the measure will be of the form

$$\int [dl^2] = \int_0^\infty \prod_s (V_d(s))^\sigma \prod_{ij} dl_{ij}^2 \Theta[l_{ij}^2] . \quad (559)$$

The lattice partition function Z_{latt} should be compared to the continuum Euclidean Feynman path integral of Eq. (183),

$$Z_{cont} = \int [dg_{\mu\nu}] e^{-\lambda \int dx \sqrt{g} + \frac{1}{16\pi G} \int dx \sqrt{g} R} . \quad (560)$$

When doing an expansion in the kinetic term proportional to k , it will be convenient to include the λ -term in the measure. We will set therefore in this Section as in Eq. (423)

$$d\mu(l^2) \equiv [dl^2] e^{-\lambda_0 \sum_h V_h} . \quad (561)$$

It should be clear that this last expression represents a fairly non-trivial quantity, both in view of the relative complexity of the expression for the volume of a simplex, Eq. (335), and because of the generalized triangle inequality constraints already implicit in $[dl^2]$. But, like the continuum functional measure, it is certainly *local*, to the extent that each edge length appears only in the expression for the volume of those simplices which explicitly contain it. Also, we note that in general the integral $\int d\mu$ can only be evaluated numerically; nevertheless this can be done, at least in principle, to arbitrary precision. Furthermore, λ_0 sets the overall scale and can therefore be set equal to one without any loss of generality.

Thus the effective strong coupling measure of Eq. (561) has the properties that (a) it is local in the lattice metric of Eq. (333), to the same extent that the continuum measure is ultra-local, (b) it restricts all edge lengths to be positive, and (c) it imposes a soft cutoff on large simplices due to the λ_0 -term and the generalized triangle inequalities. Apart from these constraints, it does

not significantly restrict the fluctuations in the lattice metric field at short distances. It will be the effect of the curvature term to restrict such fluctuation, by coupling the metric field between simplices, in the same way as the derivatives appearing in the continuum Einstein term couple the metric between infinitesimally close space-time points.

As a next step, Z_{latt} is expanded in powers of k ,

$$Z_{latt}(k) = \int d\mu(l^2) e^{k \sum_h \delta_h A_h} = \sum_{n=0}^{\infty} \frac{1}{n!} k^n \int d\mu(l^2) \left(\sum_h \delta_h A_h \right)^n . \quad (562)$$

It is easy to show that $Z(k) = \sum_{n=0}^{\infty} a_n k^n$ is analytic at $k = 0$, so this expansion should be well defined up to the nearest singularity in the complex k plane. A quantitative estimate for the expected location of such a singularity in the large- d limit will be given later in Sec. III.B.3. Beyond this singularity $Z(k)$ can sometimes be extended, for example, via Padé or differential approximants⁹. The above expansion is of course analogous to the high temperature expansion in statistical mechanics systems, where the on-site terms are treated exactly and the kinetic or hopping term is treated as a perturbation. Singularities in the free energy or its derivatives can usually be pinned down with the knowledge of a large enough number of terms in the relevant expansion (Domb and Green, 1973).

Next consider a fixed, arbitrary hinge on the lattice, and call the corresponding curvature term in the action δA . Such a contribution will be denoted in the following, as is customary in lattice gauge theories, a *plaquette* contribution. For the average curvature on that hinge one has

$$\langle \delta A \rangle = \frac{\sum_{n=0}^{\infty} \frac{1}{n!} k^n \int d\mu(l^2) \delta A \left(\sum_h \delta_h A_h \right)^n}{\sum_{n=0}^{\infty} \frac{1}{n!} k^n \int d\mu(l^2) \left(\sum_h \delta_h A_h \right)^n} . \quad (563)$$

After expanding out in k the resulting expression, one obtains for the cumulants

$$\langle \delta A \rangle = \sum_{n=0}^{\infty} c_n k^n , \quad (564)$$

with

$$c_0 = \frac{\int d\mu(l^2) \delta A}{\int d\mu(l^2)} , \quad (565)$$

⁹ A first order transition cannot affect the singularity structure of $Z(k)$ as viewed from the strong coupling phase, as the free energy is C_∞ at a first order transition.

whereas to first order in k one has

$$c_1 = \frac{\int d\mu(l^2) \delta A \left(\sum_h \delta_h A_h \right)}{\int d\mu(l^2)} - \frac{\int d\mu(l^2) \delta A \cdot \int d\mu(l^2) \sum_h \delta_h A_h}{\left(\int d\mu(l^2) \right)^2} . \quad (566)$$

This last expression clearly represents a measure of the fluctuation in δA , namely $[\langle (\sum_h \delta_h A_h)^2 \rangle - \langle \sum_h \delta_h A_h \rangle^2] / N_h$, using the homogeneity properties of the lattice $\delta A \rightarrow \sum_h \delta_h A_h / N_h$. Here N_h is the number of hinges in the lattice. Equivalently, it can be written in an even more compact way as $N_h [\langle (\delta A)^2 \rangle - \langle \delta A \rangle^2]$.

To second order in k one has $c_2 = N_h^2 [\langle (\delta A)^3 \rangle - 3\langle \delta A \rangle \langle (\delta A)^2 \rangle + 2\langle \delta A \rangle^3] / 2$. At the next order one has $c_3 = N_h^3 [\langle (\delta A)^4 \rangle - 4\langle \delta A \rangle \langle (\delta A)^3 \rangle - 3\langle (\delta A)^2 \rangle^2 + 12\langle (\delta A)^2 \rangle \langle \delta A \rangle^2 - 6\langle \delta A \rangle^4] / 6$, and so on. Note that the expressions in square parentheses become rapidly quite small, $O(1/N_h^n)$ with increasing order n , as a result of large cancellations that must arise eventually between individual terms inside the square parentheses. In principle, a careful and systematic numerical evaluation of the above integrals (which is quite feasible in practice) would allow the determination of the expansion coefficients in k for the average curvature $\langle \delta A \rangle$ to rather high order.

As an example, consider a non-analyticity in the average scalar curvature

$$\mathcal{R}(k) = \frac{\langle \int dx \sqrt{g(x)} R(x) \rangle}{\langle \int dx \sqrt{g(x)} \rangle} \approx \frac{\langle \sum_h \delta_h A_h \rangle}{\langle \sum_h V_h \rangle} , \quad (567)$$

assumed for concreteness to be of the form of an algebraic singularity at k_c , namely

$$\mathcal{R}(k) \underset{k \rightarrow k_c}{\sim} A_{\mathcal{R}} (k_c - k)^\delta \quad (568)$$

with δ some exponent. It will lead to a behavior, for the general term in the series in k , of the type

$$(-1)^n A_{\mathcal{R}} \frac{(\delta - n + 1)(\delta - n + 2) \dots \delta}{n! k_c^{n-\delta}} k^n . \quad (569)$$

Given enough terms in the series, the singularity structure can then be investigated using a variety of increasingly sophisticated series analysis methods.

It can be advantageous to isolate in the above expressions the *local* fluctuation term, from those terms that involve *correlations* between different hinges. To see this, one needs to go back, for example, to the first order expression in Eq. (566) and isolate in the sum \sum_h the contribution which contains the selected hinge with value δA , namely

$$\sum_h \delta_h A_h = \delta A + \sum_h ' \delta_h A_h , \quad (570)$$

where the primed sum indicates that the term containing δA is *not* included. The result is

$$c_1 = \frac{\int d\mu(l^2) (\delta A)^2}{\int d\mu(l^2)} - \frac{\left(\int d\mu(l^2) \delta A\right)^2}{\left(\int d\mu(l^2)\right)^2} + \frac{\int d\mu(l^2) \delta A \sum_h' \delta_h A_h}{\int d\mu(l^2)} - \frac{\left(\int d\mu(l^2) \delta A\right) \left(\int d\mu(l^2) \sum_h' \delta_h A_h\right)}{\left(\int d\mu(l^2)\right)^2} . \quad (571)$$

One then observes the following: the first two terms describe the *local* fluctuation of δA on a given hinge; the third and fourth terms describe *correlations* between δA terms on *different* hinges. But because the action is *local*, the only non-vanishing contribution to the last two terms comes from edges and hinges which are in the immediate vicinity of the hinge in question. For hinges located further apart (indicated below by “*not nn*”) one has that their fluctuations remain uncorrelated, leading to a vanishing variance

$$\frac{\int d\mu(l^2) \delta A \sum_{h \text{ not } nn} \delta_h A_h}{\int d\mu(l^2)} - \frac{\left(\int d\mu(l^2) \delta A\right) \left(\int d\mu(l^2) \sum_{h \text{ not } nn} \delta_h A_h\right)}{\left(\int d\mu(l^2)\right)^2} = 0 , \quad (572)$$

since for uncorrelated random variables X_n 's, $\langle X_n X_m \rangle - \langle X_n \rangle \langle X_m \rangle = 0$. Therefore the only non-vanishing contributions in the last two terms in Eq. (571) come from hinges which are *close* to each other.

The above discussion makes it clear that a key quantity is the *correlation* between different plaquettes,

$$\langle (\delta A)_h (\delta A)_{h'} \rangle = \frac{\int d\mu(l^2) (\delta A)_h (\delta A)_{h'} e^{k \sum_h \delta_h A_h}}{\int d\mu(l^2) e^{k \sum_h \delta_h A_h}} , \quad (573)$$

or, better, its *connected* part (denoted here by $\langle \dots \rangle_C$)

$$\langle (\delta A)_h (\delta A)_{h'} \rangle_C \equiv \langle (\delta A)_h (\delta A)_{h'} \rangle - \langle (\delta A)_h \rangle \langle (\delta A)_{h'} \rangle , \quad (574)$$

which subtracts out the trivial part of the correlation. Here again the exponentials in the numerator and denominator can be expanded out in powers of k , as in Eq. (563). The lowest order term in k will involve the correlation

$$\int d\mu(l^2) (\delta A)_h (\delta A)_{h'} . \quad (575)$$

But unless the two hinges are close to each other, they will fluctuate in an uncorrelated manner, with $\langle (\delta A)_h (\delta A)_{h'} \rangle - \langle (\delta A)_h \rangle \langle (\delta A)_{h'} \rangle = 0$. In order to achieve a non-trivial correlation, the path between the two hinges h and h' needs to be tiled by at least as many terms from the product $(\sum_h \delta_h A_h)^n$ in

$$\int d\mu(l^2) (\delta A)_h (\delta A)_{h'} \left(\sum_h \delta_h A_h \right)^n \quad (576)$$

as are needed to cover the distance l between the two hinges. One then has

$$\langle (\delta A)_h (\delta A)_{h'} \rangle_C \sim k^l \sim e^{-l/\xi} , \quad (577)$$

with the correlation length $\xi = 1/|\log k| \rightarrow 0$ to lowest order as $k \rightarrow 0$ (here we have used the usual definition of the correlation length ξ , namely that a generic correlation function is expected to decay as $\exp(-\text{distance}/\xi)$ for large separations)¹⁰. This last result is quite general, and holds for example irrespective of the boundary conditions (unless of course $\xi \sim L$, where L is the linear size of the system, in which case a path can be found which wraps around the lattice).

But further thought reveals that the above result is in fact not completely correct, due to the fact that in order to achieve a non-vanishing correlation one needs, at least to lowest order, to connect the two hinges by a narrow tube (Hamber and Williams, 2006). The previous result should then read correctly as

$$\langle (\delta A)_h (\delta A)_{h'} \rangle_C \sim (k^{n_d})^l , \quad (578)$$

where $n_d l$ represents the minimal number of dual lattice polygons needed to form a closed surface connecting the hinges h and h' , with l the actual distance (in lattice units) between the two hinges. Fig. 27 provides an illustration of the situation.

With some additional effort many additional terms can be computed in the strong coupling expansion. In practice the method is generally not really competitive with direct numerical evaluation of the path integral via Monte Carlo methods. But it does provide a new way of looking at the functional integral, and provide the basis for new approaches, such as the large d limit to be discussed in the second half of the next section.

¹⁰ This statement, taken literally, oversimplifies the situation a bit, as depending on the spin (or tensor structure) of the operator appearing in the correlation function, the large distance decay of the corresponding correlator is determined by the lightest excitation in that specific channel. But in the gravitational context one is mostly concerned with correlators involving spin two (transverse-traceless) objects, evaluated at fixed geodesic distance.

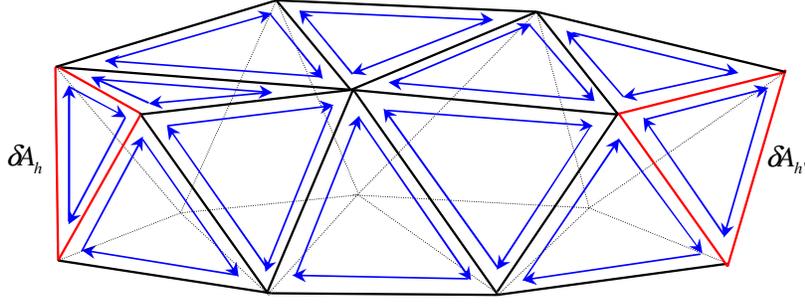


FIG. 27 Correlations between action contributions on hinge h and hinge h' arise to lowest order in the strong coupling expansions from diagrams describing a narrow tube connecting the two hinges. Here vertices represent points in the dual lattice, with the tube-like closed surface tiled with parallel transport polygons. For each link of the dual lattice, the $SO(4)$ parallel transport matrices \mathbf{R} of Sec. III.A.2 are represented by an arrow.

3. Discrete Gravity in the Large- d Limit

In the large- d limit the geometric expressions for volume, areas and angles simplify considerably, and as will be shown below one can obtain a number of interesting results for lattice gravity. These can then be compared to earlier investigations of continuum Einstein gravity in the same limit (Strominger, 1981).

Here we will consider a general simplicial lattice in d dimensions, made out of a collection of flat d -simplices glued together at their common faces so as to constitute a triangulation of a smooth continuum manifold, such as the d -torus or the surface of a sphere. Each simplex is endowed with $d + 1$ vertices, and its geometry is completely specified by assigning the lengths of its $d(d + 1)/2$ edges. We will label the vertices by $1, 2, 3, \dots, d + 1$ and denote the square edge lengths by $l_{12}^2 = l_{21}^2, \dots, l_{1,d+1}^2$.

As discussed in Sec. III.A.1, the volume of a d -simplex can be computed from the determinant of a $(d + 2) \times (d + 2)$ matrix,

$$V_d = \frac{(-1)^{\frac{d+1}{2}}}{d! 2^{d/2}} \begin{vmatrix} 0 & 1 & 1 & \dots \\ 1 & 0 & l_{12}^2 & \dots \\ 1 & l_{21}^2 & 0 & \dots \\ 1 & l_{31}^2 & l_{32}^2 & \dots \\ \dots & \dots & \dots & \dots \\ 1 & l_{d+1,1}^2 & l_{d+1,2}^2 & \dots \end{vmatrix}^{1/2}. \quad (579)$$

If one calls the above matrix M_d then previous expression can be re-written as

$$V_d = \frac{(-1)^{\frac{d+1}{2}}}{d! 2^{d/2}} \sqrt{\det M_d} , \quad (580)$$

In general the formulae for volumes and angles are quite complicated and therefore of limited use in large dimensions. The next step consists in expanding them out in terms of small edge length variations, by setting

$$l_{ij}^2 = l_{ij}^{(0)2} + \delta l_{ij}^2 . \quad (581)$$

From now on we will set $\delta l_{ij}^2 = \epsilon_{ij}$. Unless stated otherwise, we will be considering the expansion about the equilateral case, and set $l_{ij}^{(0)} = 1$; later on this restriction will be relaxed. In the equilateral case one has for the volume of a simplex

$$V_d = \frac{1}{d!} \sqrt{\frac{d+1}{2^d}} , \quad (582)$$

From the well-known expansion for determinants

$$\begin{aligned} \det(1 + M) &= e^{\text{tr} \ln(1+M)} \\ &= 1 + \text{tr} M + \frac{1}{2!} [(\text{tr} M)^2 - \text{tr} M^2] + \dots . \end{aligned} \quad (583)$$

one finds after a little algebra

$$V_d \underset{d \rightarrow \infty}{\sim} \frac{\sqrt{d}}{d! 2^{d/2}} \left\{ 1 - \frac{1}{2} \epsilon_{12}^2 + \dots + \frac{1}{d} (\epsilon_{12} + \dots + \epsilon_{12} \epsilon_{13} + \dots) + O(d^{-2}) \right\} . \quad (584)$$

Note that the terms linear in ϵ , which would have required a shift in the ground state value of ϵ for non-vanishing cosmological constant λ_0 , vanish to leading order in $1/d$. The complete volume term $\lambda_0 \sum V_d$ appearing in the action can then be easily written down using the above expressions.

In d dimensions the dihedral angle in a d -dimensional simplex of volume V_d , between faces of volume V_{d-1} and V'_{d-1} , is obtained from Eq. (342)

$$\sin \theta_d = \frac{d}{d-1} \frac{V_d V_{d-2}}{V_{d-1} V'_{d-1}} . \quad (585)$$

In the equilateral case one has for the dihedral angle

$$\theta_d = \arcsin \frac{\sqrt{d^2 - 1}}{d} \underset{d \rightarrow \infty}{\sim} \frac{\pi}{2} - \frac{1}{d} - \frac{1}{6d^3} + \dots , \quad (586)$$

which will require *four* simplices to meet on a hinge, to give a deficit angle of $2\pi - 4 \times \frac{\pi}{2} \approx 0$ in large dimensions. One notes that in large dimensions the simplices look locally (i.e. at a vertex) more

like hypercubes. Several d -dimensional simplices will meet on a $(d-2)$ -dimensional hinge, sharing a common face of dimension $d-1$ between adjacent simplices. Each simplex has $(d-2)(d-1)/2$ edges “on” the hinge, some more edges are then situated on the two “interfaces” between neighboring simplices meeting at the hinge, and finally one edge lies “opposite” to the hinge in question.

In the large d limit one then obtains, to leading order for the dihedral angle at the hinge with vertices labelled by $1 \dots d-1$

$$\begin{aligned} \theta_d \sim_{d \rightarrow \infty} & \arcsin \frac{\sqrt{d^2-1}}{d} + \epsilon_{d,d+1} + \epsilon_{1,d} \epsilon_{1,d+1} + \dots \\ & + \frac{1}{d} \left(-\epsilon_{1,d} + \dots - \frac{1}{2} \epsilon_{1,d}^2 + \dots - \frac{1}{2} \epsilon_{d,d+1}^2 - \epsilon_{12} \epsilon_{1,d+1} - \epsilon_{1,d} \epsilon_{3,d+1} - \epsilon_{1,d} \epsilon_{d,d+1} + \dots \right) \\ & + O(d^{-2}) . \end{aligned} \quad (587)$$

From the expressions in Eq. (584) for the volume and (587) for the dihedral angle one can then evaluate the d -dimensional Euclidean lattice action, involving cosmological constant and scalar curvature terms as in Eq. (373)

$$I(l^2) = \lambda_0 \sum V_d - k \sum \delta_d V_{d-2} , \quad (588)$$

where δ_d is the d -dimensional deficit angle, $\delta_d = 2\pi - \sum_{\text{simplices}} \theta_d$. The lattice functional integral is then

$$Z(\lambda_0, k) = \int [dl^2] \exp(-I(l^2)) . \quad (589)$$

To evaluate the curvature term $-k \sum \delta_d V_{d-2}$ appearing in the gravitational lattice action one needs the hinge volume V_{d-2} , which is easily obtained from Eq. (584), by reducing $d \rightarrow d-2$.

We now specialize to the case where four simplices meet at a hinge. When expanded out in terms of the ϵ 's one obtains for the deficit angle

$$\begin{aligned} \delta_d = & 2\pi - 4 \cdot \frac{\pi}{2} + \sum_{\text{simplices}} \frac{1}{d} - \epsilon_{d,d+1} + \dots - \epsilon_{1,d} \epsilon_{1,d+1} + \dots \\ & - \frac{1}{d} \left(-\epsilon_{1,d} - \frac{1}{2} \epsilon_{1,d}^2 - \frac{1}{2} \epsilon_{d,d+1}^2 - \epsilon_{12} \epsilon_{1,d+1} - \epsilon_{1,d} \epsilon_{3,d+1} - \epsilon_{1,d} \epsilon_{d,d+1} + \dots \right) + O\left(\frac{1}{d^2}\right) . \end{aligned} \quad (590)$$

The action contribution involving the deficit angle is then, for a single hinge,

$$-k \delta_d V_{d-2} = (-k) \frac{2 d^{3/2} (d-1)}{d! 2^{d/2}} \left(-\epsilon_{d,d+1} + \dots - \epsilon_{1,d} \epsilon_{1,d+1} + \dots \right) . \quad (591)$$

It involves two types of terms: one linear in the (single) edge opposite to the hinge, as well as a term involving a product of two distinct edges, connecting any hinge vertex to the two vertices

opposite to the given hinge. Since there are four simplices meeting on one hinge, one will have 4 terms of the first type, and $4(d-1)$ terms of the second type.

To obtain the total action, a sum over all simplices, resp. hinges, has still to be performed. Dropping the irrelevant constant term and summing over edges one obtains for the total action $\lambda_0 \sum V_d - k \sum \delta_d V_{d-2}$ in the large d limit

$$\lambda_0 \left(-\frac{1}{2} \sum \epsilon_{ij}^2 \right) - 2k d^2 \left(-\sum \epsilon_{jk} - \sum \epsilon_{ij} \epsilon_{ik} \right) , \quad (592)$$

up to an overall multiplicative factor $\sqrt{d}/d! 2^{d/2}$, which will play no essential role in the following.

The next step involves the choice of a specific lattice. Here we will evaluate the action for the cross polytope β_{d+1} . The cross polytope β_n is the regular polytope in n dimensions corresponding to the convex hull of the points formed by permuting the coordinates $(\pm 1, 0, 0, \dots, 0)$, and has therefore $2n$ vertices. It is named so because its vertices are located equidistant from the origin, along the Cartesian axes in n -space. The cross polytope in n dimensions is bounded by $2^n (n-1)$ -simplices, has $2n$ vertices and $2n(n-1)$ edges.

In three dimensions, it represents the convex hull of the octahedron, while in four dimensions the cross polytope is the 16-cell (Coxeter 1948; Coxeter 1974). In the general case it is dual to a hypercube in n dimensions, with the ‘dual’ of a regular polytope being another regular polytope having one vertex in the center of each cell of the polytope one started with. Fig. 28 shows as an example the polytope β_8 .

When we consider the surface of the cross polytope in $d+1$ dimensions, we have an object of dimension $n-1 = d$, which corresponds to a triangulated manifold with no boundary, homeomorphic to the sphere. From Eq. (590) the deficit angle is then given to leading order by

$$\delta_d = 0 + \frac{4}{d} - (\epsilon_{d,d+1} + 3 \text{ terms} + \epsilon_{1,d} \epsilon_{1,d+1} + \dots) + \dots \quad (593)$$

and therefore close to flat in the large d limit (due to our choice of an equilateral starting configuration). Indeed if the choice of triangulation is such that the deficit angle is not close to zero, then the discrete model leads to an average curvature whose magnitude is comparable to the lattice spacing or ultraviolet cutoff, which from a physical point of view does not seem very attractive: one obtains a space-time with curvature radius comparable to the Planck length.

When evaluated on such a manifold the lattice action becomes

$$\frac{\sqrt{d} 2^{d/2}}{d!} 2 \left(\lambda_0 - k d^3 \right) \left[1 - \frac{1}{8} \sum \epsilon_{ij}^2 + \frac{1}{d} \left(\frac{1}{4} \sum \epsilon_{ij} + \frac{1}{8} \sum \epsilon_{ij} \epsilon_{ik} \right) + O(1/d^2) \right] . \quad (594)$$

Dropping the $1/d$ correction the action is proportional to

$$- \frac{1}{2} \left(\lambda_0 - k d^3 \right) \sum \epsilon_{ij}^2 . \quad (595)$$

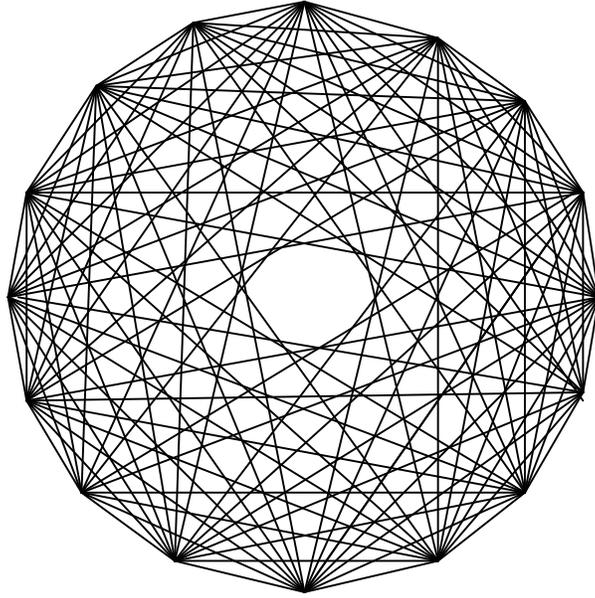


FIG. 28 Cross polytope β_n with $n = 8$ and $2n = 16$ vertices, whose surface can be used to define a simplicial manifold of dimension $d = n - 1 = 7$. For general d , the cross polytope β_{d+1} will have $2(d + 1)$ vertices, connected to each other by $2d(d + 1)$ edges.

Since there are $2d(d + 1)$ edges in the cross polytope, one finds therefore that, at the critical point $kd^3 = \lambda_0$, the quadratic form in ϵ , defined by the above action, develops $2d(d + 1) \sim 2d^2$ zero eigenvalues.

This result is quite close to the $d^2/2$ zero eigenvalues expected in the continuum for large d , with the factor of four discrepancy presumably attributed to an underlying intrinsic ambiguity that arises when trying to identify lattice points with points in the continuum.

It is worth noting here that the competing curvature (k) and cosmological constant (λ_0) terms will have comparable magnitude when

$$k_c = \frac{\lambda_0 l_0^2}{d^3} . \quad (596)$$

Here we have further allowed for the possibility that the average lattice spacing $l_0 = \langle l^2 \rangle^{1/2}$ is not equal to one (in other words, we have restored the appropriate overall scale for the average edge length, which is in fact largely determined by the value of λ_0).

The average lattice spacing l_0 can easily be estimated from the following argument. The volume of a general equilateral simplex is given by Eq. (582), multiplied by an additional factor of l_0^d . In the limit of small k the average volume of a simplex is largely determined by the cosmological term, and can therefore be computed from

$$\langle V \rangle = - \frac{\partial}{\partial \lambda_0} \log \int [dl^2] e^{-\lambda_0 V(l^2)} , \quad (597)$$

with $V(l^2) = (\sqrt{d+1}/d! 2^{d/2}) l^d \equiv c_d l^d$. After doing the integral over l^2 with measure dl^2 and solving this last expression for l_0^2 one obtains

$$l_0^2 = \frac{1}{\lambda_0^{2/d}} \left[\frac{2}{d} \frac{d! 2^{d/2}}{\sqrt{d+1}} \right]^{2/d} \quad (598)$$

(which, for example, gives $l_0 = 2.153$ for $\lambda_0 = 1$ in four dimensions, in reasonable agreement with the actual value $l_0 \approx 2.43$ found near the transition point).

This then gives for $\lambda_0 = 1$ the estimate $k_c = \sqrt{3}/(16 \cdot 5^{1/4}) = 0.0724$ in $d = 4$, to be compared with $k_c = 0.0636(11)$ obtained in (Hamber, 2000) by direct numerical simulation in four dimensions. Even in $d = 3$ one finds again for $\lambda_0 = 1$, from Eqs. (596) and (598), $k_c = 2^{5/3}/27 = 0.118$, to be compared with $k_c = 0.112(5)$ obtained in (Hamber and Williams, 1993) by direct numerical simulation.

Using Eq. (598) inserted into Eq. (596) one obtains in the large d limit for the dimensionless combination $k/\lambda_0^{(d-2)/d}$

$$\frac{k_c}{\lambda_0^{1-2/d}} = \frac{2^{1+2/d}}{d^3} \left[\frac{\Gamma(d)}{\sqrt{d+1}} \right]^{2/d} . \quad (599)$$

To summarize, an expansion in powers of $1/d$ can be developed, which relies on a combined use of the weak field expansion. It can be regarded therefore as a double expansion in $1/d$ and ϵ , valid wherever the fields are smooth enough and the geometry is close to flat, which presumably is the case in the vicinity of the lattice critical point at k_c .

A somewhat complementary $1/d$ expansion can be set up, which does not require weak fields, but relies instead on the strong coupling (small $k = 1/8\pi G$, or large G) limit. As such it will be a double expansion in $1/d$ and k . Its validity will be in a regime where the fields are not smooth, and in fact will involve lattice field configurations which are very far from smooth at short distances.

The general framework for the strong coupling expansion for pure quantum gravity was outlined in the previous section, and is quite analogous to what one does in gauge theories (Balian, Drouffe and Itzykson, 1975). One expands Z_{latt} in powers of k as in Eq. (562)

$$Z_{latt}(k) = \int d\mu(l^2) e^{k \sum_h \delta_h A_h} = \sum_{n=0}^{\infty} \frac{1}{n!} k^n \int d\mu(l^2) \left(\sum_h \delta_h A_h \right)^n . \quad (600)$$

Then one can show that dominant diagrams contributing to Z_{latt} correspond to closed surfaces tiled with elementary transport loops. In the case of the hinge-hinge connected correlation function the leading contribution at strong coupling come from closed surfaces anchored on the two hinges, as in Eq. (578).

It will be advantageous to focus on general properties of the parallel transport matrices \mathbf{R} , discussed previously in Sec. III.A.2. For smooth enough geometries, with small curvatures, these rotation matrices can be chosen to be close to the identity. Small fluctuations in the geometry will then imply small deviations in the \mathbf{R} 's from the identity matrix. But for strong coupling ($k \rightarrow 0$) the measure $\int d\mu(l^2)$ does not significantly restrict fluctuations in the lattice metric field. As a result we will assume that these fields can be regarded, in this regime, as basically unconstrained random variables, only subject to the relatively mild constraints implicit in the measure $d\mu$. The geometry is generally far from smooth since there is no coupling term to enforce long range order (the coefficient of the lattice Einstein term is zero), and one has as a consequence large local fluctuations in the geometry. The matrices \mathbf{R} will therefore fluctuate with the local geometry, and average out to zero, or a value close to zero. In the sense that, for example, the $SO(4)$ rotation

$$\mathbf{R}_\theta = \begin{pmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (601)$$

averages out to zero when integrated over θ . In general an element of $SO(n)$ is described by $n(n-1)/2$ independent parameters, which in the case at hand can be conveniently chosen as the six $SO(4)$ Euler angles. The uniform (Haar) measure over the group is then

$$d\mu_H(\mathbf{R}) = \frac{1}{32\pi^9} \int_0^{2\pi} d\theta_1 \int_0^\pi d\theta_2 \int_0^\pi d\theta_3 \int_0^\pi d\theta_4 \sin \theta_4 \int_0^\pi d\theta_5 \sin \theta_5 \int_0^\pi d\theta_6 \sin^2 \theta_6 \quad (602)$$

This is just a special case of the general n result, which reads

$$d\mu_H(\mathbf{R}) = \left(\prod_{i=1}^n \Gamma(i/2) / 2^n \pi^{n(n+1)/2} \right) \prod_{i=1}^{n-1} \prod_{j=1}^i \sin^{j-1} \theta_j^i d\theta_j^i \quad (603)$$

with $0 \leq \theta_k^1 < 2\pi$, $0 \leq \theta_k^j < \pi$.

These averaging properties of rotations are quite similar of course to what happens in $SU(N)$ Yang-Mills theories, or even more simply in (compact) QED, where the analogs of the $SO(d)$ rotation matrices \mathbf{R} are phase factors $U_\mu(x) = e^{iaA_\mu(x)}$. There one has $\int \frac{dA_\mu}{2\pi} U_\mu(x) = 0$ and $\int \frac{dA_\mu}{2\pi} U_\mu(x) U_\mu^\dagger(x) = 1$. In addition, for two contiguous closed paths C_1 and C_2 sharing a common side one has

$$e^{i \oint_{C_1} \mathbf{A} \cdot d\mathbf{l}} e^{i \oint_{C_2} \mathbf{A} \cdot d\mathbf{l}} = e^{i \oint_C \mathbf{A} \cdot d\mathbf{l}} = e^{i \int_S \mathbf{B} \cdot \mathbf{n} dA} \quad , \quad (604)$$

with C the slightly larger path encircling the two loops. For a closed surface tiled with many contiguous infinitesimal closed loops the last expression evaluates to 1, due to the divergence

theorem. In the lattice gravity case the discrete analog of this last result is considerably more involved, and ultimately represents the (exact) lattice analog of the contracted Bianchi identities. An example of a closed surface tiled with parallel transport polygons (here chosen for simplicity to be triangles) is shown in Fig. 29.

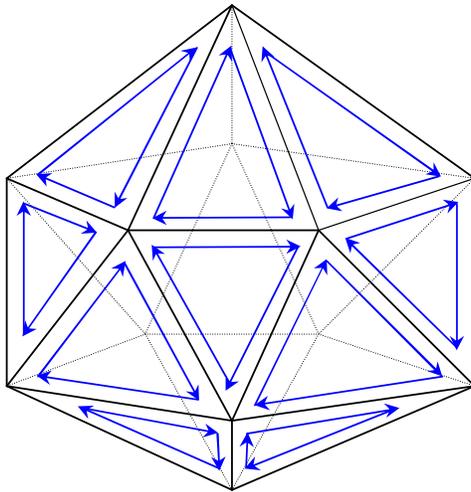


FIG. 29 Elementary closed surface tiled with parallel transport polygons, here chosen to be triangles for illustrative purposes. For each link of the dual lattice, the elementary parallel transport matrices $\mathbf{R}(s, s')$ are represented by an arrow. In spite of the fact that the (Lorentz) matrices \mathbf{R} fluctuate with the local geometry, two contiguous, oppositely oriented arrows always give $\mathbf{R}\mathbf{R}^{-1} = 1$.

As one approaches the critical point, $k \rightarrow k_c$, one is interested in random surfaces which are of very large extent. Let n_p be the number of polygons in the surface, and set $n_p = T^2$ since after all one is describing a surface. The critical point then naturally corresponds to the appearance of surfaces of infinite extent,

$$n_p = T^2 \sim \frac{1}{k_c - k} \rightarrow \infty . \quad (605)$$

A legitimate parallel is to the simpler case of scalar field theories, where random walks of length T describing particle paths become of infinite extent at the critical point, situated where the inverse of the (renormalized) mass $\xi = m^{-1}$, expressed in units of the ultraviolet cutoff, diverges.

In the present case of polygonal random surfaces, one can provide the following concise argument in support of the identification in Eq. (605). First approximate the discrete sums over n , as they appear for example in the strong coupling expansion for the average curvature, Eq. (563) or its

correlation, Eq. (573), by continuous integrals over areas

$$\sum_{n=0}^{\infty} c_n \left(\frac{k}{k_c}\right)^n \rightarrow \int_0^{\infty} dA A^{\gamma-1} \left(\frac{k}{k_c}\right)^A = \Gamma(\gamma) \left(\log \frac{k_c}{k}\right)^{-\gamma}, \quad (606)$$

where $A \equiv T^2$ is the area of a given surface. The $A^{\gamma-1}$ term can be regarded as counting the multiplicity of the surface (its entropy, in statistical mechanics terms). The exponent γ depends on the specific quantity one is looking at. For the average curvature one has from Eq. (568) $\gamma = -\delta$, while for its derivative, the curvature fluctuation (the curvature correlation function at zero momentum), one expects $\gamma = 1 - \delta$. The saddle point is located at

$$A = \frac{(\gamma - 1)}{\log \frac{k_c}{k}} \underset{k \rightarrow k_c}{\sim} \frac{(\gamma - 1) k_c}{k_c - k}. \quad (607)$$

From this discussion one then concludes that close to the critical point very large areas dominate, as claimed in Eq. (605).

Furthermore, one would expect that the universal geometric scaling properties of such a (closed) surface would not depend on its short distance details, such as whether it is constructed out of say triangles or more complex polygons. In general excluded volume effects at finite d will provide constraints on the detailed geometry of the surface, but as $d \rightarrow \infty$ these constraints can presumably be neglected and one is dealing then with a more or less unconstrained random surface. This should be regarded as a direct consequence of the fact that as $d \rightarrow \infty$ there are infinitely many dimensions for the random surface to twist and fold into, giving a negligible contribution from unallowed (by interactions) directions. In the following we will assume that this is indeed the case, and that no special pathologies arise, such as the collapse of the random surface into narrow tube-like, lower dimensional geometric configurations. Then in the large d limit the problem simplifies considerably.

Related examples for what is meant in this context are the simpler cases of random walks in infinite dimensions, random polymers and random surfaces in gauge theories (Drouffe, Parisi and Sourlas, 1979), which have been analysed in detail in the large- d limit. There too the problem simplifies considerably in such a limit since excluded volume effects (self-intersections) can be neglected there as well. A summary of these results, with a short derivation, is given in the appendices of (Hamber and Williams, 2006).

Following (Gross, 1984) one can define the partition function for such an ensemble of unconstrained random surfaces, and one finds that the mean square size of the surface increases logarithmically with the intrinsic area of the surface. This last result is usually interpreted as the statement that an unconstrained random surface has infinite fractal (or Hausdorff) dimension. Although made of very many triangles (or polygons), the random surface remains quite compact

in overall size, as viewed from the original embedding space. In a sense, an unconstrained random surface is a much more compact object than an unconstrained random walk, for which $\langle \mathbf{X}^2 \rangle \sim T$. Identifying the size of the random surface with the gravitational correlation length ξ then gives

$$\xi \sim \sqrt{\log T} \underset{k \rightarrow k_c}{\sim} |\log(k_c - k)|^{1/2} . \quad (608)$$

From the definition of the exponent ν , namely $\xi \sim (k_c - k)^{-\nu}$, the above result then implies $\nu = 0$ (i.e. a weak logarithmic singularity) at $d = \infty$.

It is of interest to contrast the result $\nu \sim 0$ for gravity in large dimensions with what one finds for scalar (Wilson and Fisher, 1972; Wilson 1973) and gauge (Drouffe, Parisi and Sourlas, 1979) fields, in the same limit $d = \infty$. So far, known results can be summarized as follows

$$\begin{aligned} \text{scalar field} & \quad \nu = \frac{1}{2} \\ \text{lattice gauge field} & \quad \nu = \frac{1}{4} \\ \text{lattice gravity} & \quad \nu = 0 . \end{aligned} \quad (609)$$

It should be regarded as encouraging that the new value obtained here, namely $\nu = 0$ for gravitation, appears to some extent to be consistent with the general trend observed for lower spin, at least at infinite dimension. What happens in finite dimensions? The situation becomes much more complicated since the self-intersection properties of the surface have to be taken into account. But a simple geometric argument then suggests in finite but large dimensions $\nu = 1/(d - 1)$ (Hamber and Williams, 2004).

IV. NUMERICAL STUDIES IN FOUR DIMENSIONS

The exact evaluation of the lattice functional integral for quantum gravity by numerical methods allows one to investigate a regime which is generally inaccessible by perturbation theory, where the coupling G is strong and quantum fluctuations are expected to be large.

The hope in the end is to make contact with the analytic results obtained, for example, in the $2 + \epsilon$ expansion, and determine which scenarios are physically realized in the lattice regularized model, and then perhaps even in the real world.

Specifically, one can enumerate several major questions that one would like to get at least partially answered.

- The first one is: which scenarios suggested by perturbation theory are realized in the lattice theory? Perhaps a stable ground state for the quantum theory cannot be found, which would imply that the regulated theory is still inherently pathological.

- Furthermore, if a stable ground state exists for some range of bare parameters, does it require the inclusion of higher derivative couplings in an essential way, or is the minimal theory, with an Einstein and a cosmological term, sufficient?
- Does the presence of dynamical matter, say in the form of a massless scalar field, play an important role, or is the non-perturbative dynamics of gravity largely determined by the pure gravity sector (as in Yang-Mills theories)?
- Is there any indication that the non-trivial ultraviolet fixed point scenario is realized in the lattice theory in four dimensions? This would imply, as in the non-linear sigma model, the existence of at least two physically distinct phases and non-trivial exponents. Which quantity can be used as an order parameter to physically describe, in a *qualitative*, way the two phases?
- A clear physical characterization of the two phases would allow one, at least in principle, to decide which phase, if any, could be realized in nature. Ultimately this might or might not be possible based on purely qualitative aspects. As will be discussed below, the lattice continuum limit is taken in the vicinity of the fixed point, so close to it is the physically most relevant regime.
- At the next level one would hope to be able to establish a *quantitative* connection with those continuum perturbative results which are not affected by uncontrollable errors, such as for example the $2 + \epsilon$ expansion of Sec. II.C.4. Since the lattice cutoff and the method of dimensional regularization cut the theory off in the ultraviolet in rather different ways, one needs to compare universal quantities which are *cutoff-independent*. One example is the critical exponent ν , as well as any other non-trivial scaling dimension that might arise. Within the $2 + \epsilon$ expansion only *one* such exponent appears, to *all* orders in the loop expansion, as $\nu^{-1} = -\beta'(G_c)$. Therefore one central issue in the lattice regularized theory is the value of the universal exponent ν .

Knowledge of ν would allow one to be more specific about the running of the gravitational coupling. One purpose of the discussion in Sec. II.C.2 was to convince the reader that the exponent ν determines the renormalization group running of $G(\mu^2)$ in the vicinity of the fixed point, as in Eq. (207) for the non-linear σ -model, and more appropriately in Eq. (300) for quantized gravity. From a practical point of view, on the lattice it is difficult to determine the running of $G(\mu^2)$

directly from correlation functions, since the effects from the running of G are generally small. Instead one would like to make use of the analog of Eqs. (214), (244) and (245) for the non-linear σ -model, and, again, more appropriately of Eqs. (304) and possibly (310) for gravity to determine ν , and from there the running of G . But the correlation length $\xi = m^{-1}$ is also difficult to compute, since it enters the curvature correlations at fixed geodesic distance, which are hard to compute for (genuinely geometric) reasons to be discussed later. Furthermore, these generally decay exponentially in the distance at strong G , and can therefore be difficult to compute due to the signal to noise problem of numerical simulations.

Fortunately the exponent ν can be determined instead, and with good accuracy, from singularities of the derivatives of the path integral Z , whose singular part is expected, on the basis of very general arguments, to behave in the vicinity of the fixed point as $F \equiv -\frac{1}{V} \ln Z \sim \xi^{-d}$ where ξ is the gravitational correlation length. From Eq. (304) relating $\xi(G)$ to $G - G_c$ and ν one can then determine ν , as well as the critical coupling G_c .

A. Observables, Phase Structure and Critical Exponents

The starting point is once again the lattice regularized path integral with action as in Eq. (373) and measure as in Eq. (406). Then the lattice action for pure four-dimensional Euclidean gravity contains a cosmological constant and Regge scalar curvature term as in Eq. (420)

$$I_{latt} = \lambda_0 \sum_h V_h(l^2) - k \sum_h \delta_h(l^2) A_h(l^2) , \quad (610)$$

with $k = 1/(8\pi G)$, and leads to the regularized lattice functional integral

$$Z_{latt} = \int [dl^2] e^{-\lambda_0 \sum_h V_h + k \sum_h \delta_h A_h} , \quad (611)$$

where, as customary, the lattice ultraviolet cutoff is set equal to one (i.e. all length scales are measured in units of the lattice cutoff). The lattice measure is given in Eq. (406) and will be therefore of the form

$$\int [dl^2] = \int_0^\infty \prod_s (V_d(s))^\sigma \prod_{ij} dl_{ij}^2 \Theta[l_{ij}^2] . \quad (612)$$

with σ a real parameter given below.

Ultimately the above lattice partition function Z_{latt} is intended as a regularized form of the continuum Euclidean Feynman path integral of Eq. (183),

$$Z_{cont} = \int [dg_{\mu\nu}] e^{-\lambda_0 \int dx \sqrt{g} + \frac{1}{16\pi G} \int dx \sqrt{g} R} . \quad (613)$$

with functional measure over the $g_{\mu\nu}(x)$'s of the form

$$\int [dg_{\mu\nu}] \equiv \prod_x [g(x)]^{\sigma/2} \prod_{\mu \geq \nu} dg_{\mu\nu}(x) \ , \quad (614)$$

where σ is a real parameter constrained by the requirement $\sigma \geq -(d+1)$. For $\sigma = \frac{1}{2}(d-4)(d+1)$ one obtains the De Witt measure of Eq. (167), while for $\sigma = -(d+1)$ one recovers the original Misner measure of Eq. (171). In the following we will mostly be interested in the four-dimensional case, for which $d = 4$ and therefore $\sigma = 0$ for the DeWitt measure.

It is possible to add higher derivative terms to the lattice action and investigate how the results are affected. The original motivation was that they would improve the convergence properties of functional integral for the lattice theory, but extensive numerical studies suggest that they don't seem to be necessary after all. In any case, with such terms included the lattice action for pure gravity acquires the two additional terms whose lattice expressions can be found in Eqs. (441) and (452),

$$\begin{aligned} I_{latt} = & \sum_h \left[\lambda_0 V_h - k \delta_h A_h - b A_h^2 \delta_h^2 / V_h \right] \\ & + \frac{1}{3} (a + 4b) \sum_s V_s \sum_{h, h' \subset s} \epsilon_{h, h'} \left(\omega_h \left[\frac{\delta}{AC} \right]_h - \omega_{h'} \left[\frac{\delta}{AC} \right]_{h'} \right)^2 \end{aligned} \quad (615)$$

The above action is intended as a lattice form for the continuum action

$$I = \int dx \sqrt{g} \left[\lambda_0 - \frac{1}{2} k R - \frac{1}{4} b R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} + \frac{1}{2} (a + 4b) C_{\mu\nu\rho\sigma} C^{\mu\nu\rho\sigma} \right] \quad (616)$$

and is therefore of the form in Eqs. (104) and (453). Because of its relative complexity, in the following the Weyl term will not be considered any further, and b will be chosen so that $b = -\frac{1}{4}a$. Thus the only curvature term to be discussed here will be a Riemann squared contribution, with a (small) positive coefficient $+\frac{1}{4}a \rightarrow 0$.

1. Invariant Local Gravitational Averages

Among the simplest quantum mechanical averages is the one associated with the local curvature

$$\mathcal{R}(k) \sim \frac{\langle \int dx \sqrt{g} R(x) \rangle}{\langle \int dx \sqrt{g} \rangle} \ , \quad (617)$$

The curvature associated with the quantity above is the one that would be detected when parallel-transporting vectors around infinitesimal loops, with size comparable to the average lattice spacing l_0 . Closely related to it is the fluctuation in the local curvature

$$\chi\mathcal{R}(k) \sim \frac{\langle (\int dx \sqrt{g} R)^2 \rangle - \langle \int dx \sqrt{g} R \rangle^2}{\langle \int dx \sqrt{g} \rangle} \ . \quad (618)$$

The latter is related to the connected curvature correlation at zero momentum

$$\chi_{\mathcal{R}} \sim \frac{\int dx \int dy \langle \sqrt{g(x)} R(x) \sqrt{g(y)} R(y) \rangle_c}{\langle \int dx \sqrt{g(x)} \rangle} . \quad (619)$$

Both $\mathcal{R}(k)$ and $\chi_{\mathcal{R}}(k)$ are directly related to derivatives of Z with respect to k ,

$$\mathcal{R}(k) \sim \frac{1}{V} \frac{\partial}{\partial k} \ln Z , \quad (620)$$

and

$$\chi_{\mathcal{R}}(k) \sim \frac{1}{V} \frac{\partial^2}{\partial k^2} \ln Z . \quad (621)$$

Thus a divergence or non-analyticity in Z , as caused for example by a phase transition, is expected to show up in these local averages as well. Note that the above expectation values are manifestly invariant, since they are related to derivatives of Z .

On the lattice one prefers to define quantities in such a way that variations in the average lattice spacing $\sqrt{\langle l^2 \rangle}$ are compensated by an appropriate factor determined from dimensional considerations. In the case of the average curvature one defines therefore the lattice quantity \mathcal{R} as

$$\mathcal{R}(k) \equiv \langle l^2 \rangle \frac{\langle 2 \sum_h \delta_h A_h \rangle}{\langle \sum_h V_h \rangle} , \quad (622)$$

and similarly for the curvature fluctuation,

$$\chi_{\mathcal{R}}(k) \equiv \frac{\langle (\sum_h \delta_h A_h)^2 \rangle - \langle \sum_h \delta_h A_h \rangle^2}{\langle \sum_h V_h \rangle} , \quad (623)$$

Fluctuations in the local curvature probe graviton correlations, and are expected to be sensitive to the presence of a massless spin two particle. Note that both of the above expressions are dimensionless, and are therefore unaffected by an overall rescaling of the edge lengths. As in the continuum, they are proportional to first and second derivatives of Z_{latt} with respect to k .

One can contrast the behavior of the preceding averages, related to the curvature, with the corresponding quantities involving the local volumes V_h (the quantity $\sqrt{g} dx$ in the continuum). Consider the average volume per site

$$\langle V \rangle \equiv \frac{1}{N_0} \langle \sum_h V_h \rangle , \quad (624)$$

and its fluctuation, defined as

$$\chi_V(k) \equiv \frac{\langle (\sum_h V_h)^2 \rangle - \langle \sum_h V_h \rangle^2}{\langle \sum_h V_h \rangle} , \quad (625)$$

where V_h is the volume associated with the hinge h . The last two quantities are again simply related to derivatives of Z_{latt} with respect to the bare cosmological constant λ_0 , as for example in

$$\langle V \rangle \sim \frac{\partial}{\partial \lambda_0} \ln Z_{latt} \quad , \quad (626)$$

and

$$\chi_V(k) \sim \frac{\partial^2}{\partial \lambda_0^2} \ln Z_{latt} \quad . \quad (627)$$

Some useful relations and sum rules can be derived, which follow directly from the scaling properties of the discrete functional integral. Thus a simple scaling argument, based on neglecting the effects of curvature terms entirely (which, as will be seen below, vanish in the vicinity of the critical point), gives an estimate of the average volume per edge [for example from Eqs. (597) and (598)]

$$\langle V_l \rangle \sim \frac{2(1 + \sigma d)}{\lambda_0 d} \underset{d=4, \sigma=0}{\sim} \frac{1}{2\lambda_0} \quad . \quad (628)$$

where σ is the functional measure parameter in Eqs. (176) and (406). In four dimensions direct numerical simulations with $\sigma = 0$ (corresponding to the lattice DeWitt measure) agree quite well with the above formula.

Some exact lattice identities can be obtained from the scaling properties of the action and measure. The bare couplings k and λ_0 in the gravitational action are dimensionful in four dimensions, but one can define the dimensionless ratio k^2/λ_0 , and rescale the edge lengths so as to eliminate the overall length scale $\sqrt{k/\lambda_0}$. As a consequence the path integral for pure gravity,

$$Z_{latt}(\lambda_0, k, a, b) = \int [dl^2] e^{-I(l^2)} \quad , \quad (629)$$

obeys the scaling law

$$Z_{latt}(\lambda_0, k, a, b) = (\lambda_0)^{-N_1/2} Z_{latt}\left(1, \frac{k}{\sqrt{\lambda_0}}, a, b\right) \quad (630)$$

where N_1 represents the number of edges in the lattice, and the dl^2 measure ($\sigma = 0$) has been selected. This implies in turn a sum rule for local averages, which for the dl^2 measure reads

$$2\lambda_0 \langle \sum_h V_h \rangle - k \langle \sum_h \delta_h A_h \rangle - N_1 = 0 \quad , \quad (631)$$

and is easily derived from Eq. (630) and the definitions in Eqs. (620) and (626). N_0 represents the number of sites in the lattice, and the averages are defined per site (for the hypercubic lattice divided up into simplices as in Fig. 26, $N_1 = 15$). This last formula can be very useful in checking

the accuracy of numerical evaluations of the path integral. A similar sum rule holds for the fluctuations

$$4 \lambda_0^2 \left[\langle (\sum_h V_h)^2 \rangle - \langle \sum_h V_h \rangle^2 \right] - k^2 \left[\langle (\sum_h \delta_h A_h)^2 \rangle - \langle \sum_h \delta_h A_h \rangle^2 \right] - 2N_1 = 0 . \quad (632)$$

In light of the above discussion one can therefore consider without loss of generality the case of unit bare cosmological coupling $\lambda_0 = 1$ (in units of the cutoff). Then all lengths are expressed in units of the fundamental microscopic length scale $\lambda_0^{-1/4}$.

2. Invariant Correlations at Fixed Geodesic Distance

Compared to ordinary field theories, new issues arise in quantum gravity due to the fact that the physical distance between any two points x and y

$$d(x, y | g) = \min_{\xi} \int_{\tau(x)}^{\tau(y)} d\tau \sqrt{g_{\mu\nu}(\xi) \frac{d\xi^\mu}{d\tau} \frac{d\xi^\nu}{d\tau}} , \quad (633)$$

is a fluctuating function of the background metric $g_{\mu\nu}(x)$. In addition, the Lorentz group used to classify spin states is meaningful only as a local concept.

In the continuum the shortest distance between two events is determined by solving the equation of motion (equation of free fall, or geodesic equation)

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma_{\lambda\sigma}^\mu \frac{dx^\lambda}{d\tau} \frac{dx^\sigma}{d\tau} = 0 \quad (634)$$

On the lattice the geodesic distance between two lattice vertices x and y requires the determination of the shortest lattice path connecting several lattice vertices, and having the two given vertices as endpoints. This can be done at least in principle by enumerating all paths connecting the two points, and then selecting the shortest one. Equivalently it can be computed from the scalar field propagator, as in Eq. (475).

Consequently physical correlations have to be defined at fixed geodesic distance d , as in the following correlation between scalar curvatures

$$\langle \int dx \int dy \sqrt{g} R(x) \sqrt{g} R(y) \delta(|x - y| - d) \rangle \quad (635)$$

Generally these do not go to zero at large separation, so one needs to define the connected part, by subtracting out the value at $d = \infty$. These will be indicated in the following by the connected $\langle \rangle_c$

average, and we will write the resulting connected curvature correlation function at fixed geodesic distance compactly as

$$G_R(d) \sim \langle \sqrt{g} R(x) \sqrt{g} R(y) \delta(|x - y| - d) \rangle_c . \quad (636)$$

One can define several more invariant correlation functions at fixed geodesic distance for other operators involving curvatures (Hamber, 1994). The gravitational correlation function just defined is similar to the one in non-Abelian gauge theories, Eq. (329).

In the lattice regulated theory one can define similar correlations, using for example the correspondence of Eqs. (368) or (440) for the scalar curvature

$$\sqrt{g} R(x) \rightarrow 2 \sum_{h \supset x} \delta_h A_h \quad (637)$$

If the deficit angles are averaged over a number of contiguous hinges h sharing a common vertex x , one is naturally lead to the connected correlation function

$$G_R(d) \equiv \langle \sum_{h \supset x} \delta_h A_h \sum_{h' \supset y} \delta_{h'} A_{h'} \delta(|x - y| - d) \rangle_c , \quad (638)$$

which probes correlations in the scalar curvatures. In practice the above lattice correlations have to be computed by a suitable binning procedure: one averages out all correlations in a geodesic distance interval $[d, d + \Delta d]$ with Δd comparable to one lattice spacing l_0 . See Fig. 30. Similarly one can construct the connected correlation functions for local volumes at fixed geodesic distance

$$G_V(d) \equiv \langle \sum_{h \supset x} V_h \sum_{h' \supset y} V_{h'} \delta(|x - y| - d) \rangle_c , \quad (639)$$

In general one expects for the curvature correlation either a power law decay, for distances sufficiently larger than the lattice spacing l_0 ,

$$\langle \sqrt{g} R(x) \sqrt{g} R(y) \delta(|x - y| - d) \rangle_c \underset{d \gg l_0}{\sim} \frac{1}{d^{2n}} , \quad (640)$$

with n some exponent characterizing the power law decay, or at very large distances an exponential decay, characterized by a correlation length ξ ,

$$\langle \sqrt{g} R(x) \sqrt{g} R(y) \delta(|x - y| - d) \rangle_c \underset{d \gg \xi}{\sim} e^{-d/\xi} . \quad (641)$$

In fact the invariant correlation length ξ is generally defined (in analogy with what one does for other theories) through the long-distance decay of the connected, invariant correlations at fixed geodesic distance d . In the pure power law decay case of Eq. (640) the correlation length ξ is of

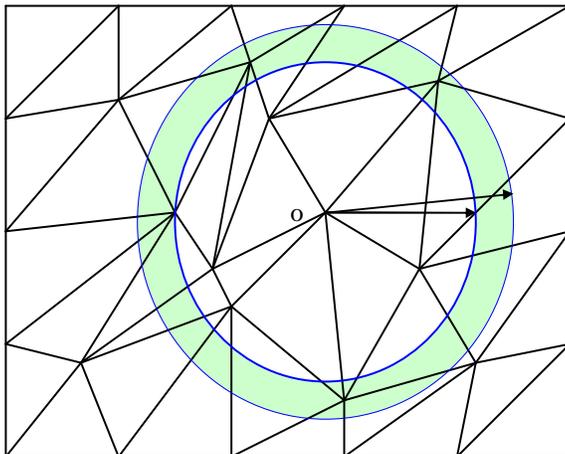


FIG. 30 Geodesic distance and correlations. Correlation functions are computed for lattice vertices in range $[d, d + \Delta d]$.

course infinite. One can show from scaling considerations (see below) that the power n in Eq. (640) is related to the critical exponent ν by $n = 4 - 1/\nu$.

In the presence of a finite correlation length ξ one needs therefore to carefully distinguish between the "short distance" regime

$$l_0 \ll d \ll \xi \quad (642)$$

where Eq. (640) is valid, and the "long distance" regime

$$\xi \ll d \ll L \quad (643)$$

where Eq. (641) is appropriate. Here $l_0 = \sqrt{\langle l^2 \rangle}$ is the average lattice spacing, and $L = V^{1/4}$ the linear size of the system.

Recently the issue of defining diffeomorphism invariant correlations in quantum gravity has been re-examined from a new perspective (Giddings, Marolf and Hartle, 2006).

3. Wilson Lines and Static Potential

In a gauge theory such as QED the static potential can be computed from the manifestly gauge invariant Wilson loop. To this end one considers the process where a particle-antiparticle pair are created at time zero, separated by a fixed distance R , and re-annihilated at a later time T . In QED

the amplitude for such a process associated with the closed loop Γ is given by the Wilson loop

$$W(\Gamma) = \langle \exp\left\{ie \oint_{\Gamma} A_{\mu}(x) dx^{\mu}\right\} \rangle , \quad (644)$$

which is a manifestly gauge invariant quantity. Performing the required Gaussian average using the (Euclidean) free photon propagator one obtains

$$\langle \exp\left\{ie \oint_{\Gamma} A_{\mu} dx^{\mu}\right\} \rangle = \exp\left\{-\frac{1}{2}e^2 \oint_{\Gamma} \oint_{\Gamma} dx^{\mu} dy^{\nu} \Delta_{\mu\nu}(x-y)\right\} . \quad (645)$$

For a rectangular loop of sides R and T one has after a short calculation

$$\langle \exp\left\{ie \oint_{\Gamma} A_{\mu} dx^{\mu}\right\} \rangle \simeq \exp\left\{-\frac{e^2}{4\pi\epsilon}(T+R) + \frac{e^2}{4\pi} \frac{T}{R} + \frac{e^2}{2\pi^2} \log\left(\frac{T}{\epsilon}\right) + \dots\right\} \quad (646)$$

$$\underset{T \gg R}{\sim} \exp\left[-V(R) T\right] , \quad (647)$$

where $\epsilon \rightarrow 0$ is an ultraviolet cutoff. In the last line use has been made of the fact that for large imaginary times the exponent in the amplitude involves the energy for the process multiplied by the time T . Thus for $V(R)$ itself one obtains

$$V(R) = -\lim_{T \rightarrow \infty} \frac{1}{T} \log \langle \exp\left\{ie \oint_{\Gamma} A_{\mu} dx^{\mu}\right\} \rangle \sim \text{cst.} - \frac{e^2}{4\pi R} , \quad (648)$$

which is the correct Coulomb potential for two oppositely charged particles.

To obtain the static potential it is not necessary to consider closed loops. Alternatively, in a periodic box of length T one can introduce two long oppositely oriented parallel lines in the time direction, separated by a distance R and closed by the periodicity of the lattice, and associated with oppositely charged particles,

$$\langle \exp\left\{ie \int_{\Gamma} A_{\mu} dx^{\mu}\right\} \exp\left\{ie \int_{\Gamma'} A_{\nu} dy^{\nu}\right\} \rangle \quad (649)$$

In the large time limit one can then show that the result for the potential $V(R)$ is the same.

In the gravitational case there is no notion of ‘‘oppositely charged particles’’, so one cannot use the closed Wilson loop to extract the potential (Modanese 1995). One is therefore forced to consider a process in which one introduces two separate world-lines for the two particles. It is well known that the equation for free fall can be obtained by extremizing the space-time distance travelled. Thus the quantity

$$\mu \int_{\tau(a)}^{\tau(b)} d\tau \sqrt{g_{\mu\nu}(x) \frac{dx^{\mu}}{d\tau} \frac{dx^{\nu}}{d\tau}} , \quad (650)$$

can be taken as the Euclidean action contribution associated with the heavy spinless particle of mass μ .

Next consider two particles of mass M_1, M_2 , propagating along parallel lines in the ‘time’ direction and separated by a fixed distance R . Then the coordinates for the two particles can be chosen to be $x^\mu = (\tau, r, 0, 0)$ with r either 0 or R . The amplitude for this process is a product of two factors, one for each heavy particle. Each is of the form

$$L(0; M_1) = \exp\left\{-M_1 \int d\tau \sqrt{g_{\mu\nu}(x) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}}\right\} . \quad (651)$$

where the first argument indicates the spatial location of the Wilson line. For the two particles separated by a distance R the amplitude is

$$\text{Amp.} \equiv W(0, R; M_1, M_2) = L(0; M_1) L(R; M_2) . \quad (652)$$

For weak fields one sets $g_{\mu\nu} = \delta_{\mu\nu} + h_{\mu\nu}$, with $h_{\mu\nu} \ll 1$, and therefore $g_{\mu\nu}(x) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 1 + h_{00}(x)$.

Then the amplitude reduces to

$$W(M_1, M_2) = \exp\left\{-M_1 \int_0^T d\tau \sqrt{1 + h_{00}(\tau)}\right\} \exp\left\{-M_2 \int_0^T d\tau' \sqrt{1 + h_{00}(\tau')}\right\} . \quad (653)$$

In perturbation theory the averaged amplitude can then be easily evaluated (Hamber and Williams, 1995)

$$\langle W(0, R; M_1, M_2) \rangle = \exp\left\{-T \left(M_1 + M_2 - G \frac{M_1 M_2}{R}\right) + \dots\right\} . \quad (654)$$

and the static potential has indeed the expected form, $V(R) = -G M_1 M_2 / R$. The contribution involving the sum of the two particle masses is R independent, and can therefore be subtracted, if the Wilson line correlation is divided by the averages of the individual single line contribution, as in

$$V(R) = - \lim_{T \rightarrow \infty} \frac{1}{T} \log \frac{\langle W(0, R; M_1, M_2) \rangle}{\langle L(0; M_1) \rangle \langle L(R; M_2) \rangle} \sim -G \frac{M_1 M_2}{R} . \quad (655)$$

If one is only interested in the spatial dependence of the potential, one can simplify things further and take $M_1 = M_2 = M$. To higher order in the weak field expansion one has to take into account multiple graviton exchanges, contributions from graviton loops and self-energy contributions due to other particles.

How does all this translate to the lattice theory? At this point, the prescription for computing the Newtonian potential for quantum gravity should be clear. For each metric configuration (which is a given configuration of edge lengths on the lattice) one chooses a geodesic that closes due to the lattice periodicity (and there might be many that have this property for the topology of a four-torus), with length T (see Fig. 31). One then enumerates all the geodesics that lie at a fixed

distance R from the original one, and computes the associated correlation between the Wilson lines. After averaging the Wilson line correlation over many metric configurations, one extracts the potential from the R dependence of the correlation of Eq. (655). In general since two geodesics will not be at a fixed geodesic distance from each other in the presence of curvature, one needs to introduce some notion of average distance, which then gives the spatial separations of the sources R .

On the lattice one can construct the analog of the Wilson line for one heavy particle,

$$L(x, y, z) = \exp\{-M \sum_i l_i\} , \quad (656)$$

where edges are summed in the “ t ” direction, and the path is closed by the periodicity of the lattice in the t direction. One can envision the simplicial lattice as divided up in hypercubes, in which case the points x, y, z can be taken as the remaining labels for the Wilson line.

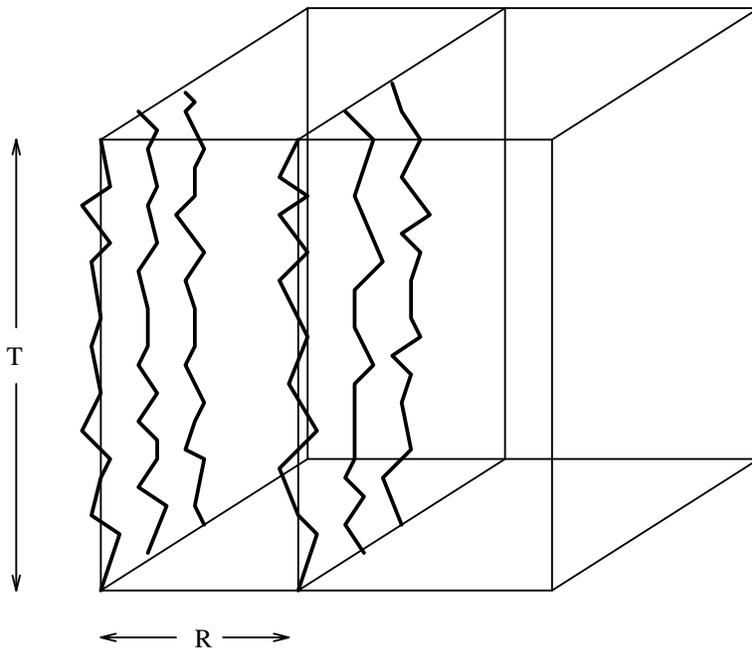


FIG. 31 Correlations between Wilson lines closed by the lattice periodicity.

For a single line one expects

$$\langle L(x, y, z) \rangle = \langle \exp\{-M \sum_i l_i\} \rangle \sim e^{-\tilde{M}T} , \quad (657)$$

where T is the linear size of lattice in the t direction and \tilde{M} the renormalized mass. The correlation between Wilson lines at average “distance” R is then given by

$$-\frac{1}{T} \log \left[\frac{\langle L(x, y, 0) L(x, y, R) \rangle}{\langle L(x, y, 0) \rangle \langle L(x, y, R) \rangle} \right] \underset{T \gg R}{\sim} V(R) . \quad (658)$$

Numerical studies suggest that the correct qualitative features of the potential emerge close to the critical point. In particular it was found that the potential is attractive close to the critical point, and for two equal mass particles of mass μ scales, as expected, like the mass squared. As for any correlation in gravity, the accurate determination of the potential as a function of distance R is a more difficult task, since at large distance the correlations are small and the statistical noise becomes large. Still, the first results (Hamber and Williams, 1995) suggest that the potential has more or less the expected classical form in the vicinity of the critical point, both as far as the mass dependence and perhaps even the distance dependence are concerned. In particular it is attractive.

The correlation of Wilson lines is not the only possible method to determine that non-perturbative gravity is still attractive. One alternate procedure involves the properties of correlation functions of scalar particles in the presence of gravity (de Bakker and Smit, 1995).

4. Scaling in the Vicinity of the Critical Point

In practice the correlation functions at fixed geodesic distance are difficult to compute numerically, and therefore not the best route to study the critical properties. But scaling arguments allow one to determine the scaling behavior of correlation functions from critical exponents characterizing the singular behavior of the *free energy* and various local averages in the vicinity of the critical point. In general a divergence of the correlation length ξ

$$\xi(k) \equiv \underset{k \rightarrow k_c}{\sim} A_\xi |k_c - k|^{-\nu} \quad (659)$$

signals the presence of a phase transition, and leads to the appearance of a singularity in the free energy $F(k)$. The scaling assumption for the free energy postulates that a divergent correlation length in the vicinity of the critical point at k_c leads to non-analyticities of the type

$$F \equiv -\frac{1}{V} \ln Z = F_{reg} + F_{sing} \quad (660)$$

$$F_{sing} \sim \xi^{-d}$$

where the second relationship follows simply from dimensional arguments (the free energy is an extensive quantity). The regular part F_{reg} is generally not determined from ξ by purely dimensional considerations, but as the name implies is a regular function in the vicinity of the critical point. Combining the definition of ν in Eq. (659) with the scaling assumption of Eq. (660) one obtains

$$F_{sing}(k) \underset{k \rightarrow k_c}{\sim} (\text{const.}) |k_c - k|^{d\nu} \quad (661)$$

The presence of a phase transition can then be inferred from non-analytic terms in invariant averages, such as the average curvature and its fluctuation. For the average curvature one obtains

$$\mathcal{R}(k) \underset{k \rightarrow k_c}{\sim} A_{\mathcal{R}} |k_c - k|^{d\nu-1} , \quad (662)$$

up to regular contributions (i.e. constant terms in the vicinity of k_c). An additive constant can be added, but numerical evidence so far points to this constant being consistent with zero. Similarly one has for the curvature fluctuation

$$\chi_{\mathcal{R}}(k) \underset{k \rightarrow k_c}{\sim} A_{\chi_{\mathcal{R}}} |k_c - k|^{-(2-d\nu)} . \quad (663)$$

At a critical point the fluctuation χ is in general expected to diverge, corresponding to the presence of a divergent correlation length. From such averages one can therefore in principle extract the correlation length exponent ν of Eq. (659) without having to compute a correlation function.

An equivalent result, relating the quantum expectation value of the curvature to the physical correlation length ξ , is obtained from Eqs. (659) and (662)

$$\mathcal{R}(\xi) \underset{k \rightarrow k_c}{\sim} \xi^{1/\nu-4} , \quad (664)$$

again up to an additive constant. Matching of dimensionalities in this last equation is restored by inserting an appropriate power of the Planck length $l_P = \sqrt{G}$ on the r.h.s..

One can relate the critical exponent ν to the scaling behavior of correlations at large distances. The curvature fluctuation is related to the connected scalar curvature correlator at zero momentum

$$\chi_{\mathcal{R}}(k) \sim \frac{\int dx \int dy \langle \sqrt{g} R(x) \sqrt{g} R(y) \rangle_c}{\langle \int dx \sqrt{g} \rangle} \quad (665)$$

A divergence in the above fluctuation is then indicative of long range correlations, corresponding to the presence of a massless particle. Close to the critical point one expects for large separations $l_0 \ll |x - y| \ll \xi$ a power law decay in the geodesic distance, as in Eq. (640),

$$\langle \sqrt{g} R(x) \sqrt{g} R(y) \rangle \underset{|x-y| \rightarrow \infty}{\sim} \frac{1}{|x - y|^{2n}} , \quad (666)$$

Inserting the above expression in Eq. (665) and comparing with Eq. (663) determines the n as $n = d - 1/\nu$. A priori one cannot exclude the possibility that some states acquire a mass away from the critical point, in which case the correlation functions would have the behavior of Eq. (641) for $|x - y| \gg \xi$.

5. Physical and Unphysical Phases

An important alternative to the analytic methods in the continuum is an attempt to solve quantum gravity directly via numerical simulations. The underlying idea is to evaluate the gravitational functional integral in the discretized theory Z by summing over a suitable finite set of representative field configurations. In principle such a method given enough configurations and a fine enough lattice can provide an arbitrarily accurate solution to the original quantum gravity theory.

In practice there are several important factors to consider, which effectively limit the accuracy that can be achieved today in a practical calculation. Perhaps the most important one is the enormous amounts of computer time that such calculations can use up. This is particularly true when correlations of operators at fixed geodesic distance are evaluated. Another practical limitation is that one is mostly interested in the behavior of the theory in the vicinity of the critical point at G_c , where the correlation length ξ can be quite large and significant correlations develop both between different lattice regions, as well as among representative field configurations, an effect known as critical slowing down. Finally, there are processes which are not well suited to a lattice study, such as problems with several different length (or energy) scales. In spite of these limitations, the progress in lattice field theory has been phenomenal in the last few years, driven in part by enormous advances in computer technology, and in part by the development of new techniques relevant to the problems of lattice field theories.

The starting point is the generation of a large ensemble of suitable edge length configurations. The edge lengths are updated by a straightforward Monte Carlo algorithm, generating eventually an ensemble of configurations distributed according to the action and measure of Eq. (421) (Hamber and Williams, 1984; Hamber, 1984; Berg, 1985); some more recent references are (Beirl et al, 1994; Riedler et al, 1999; Bittner et al, 2002). Further details of the method as applied to pure gravity can be found for example in the recent work (Hamber, 2000) and will not be repeated here.

As far as the lattice is concerned, one starts with the 4-d hypercube of Fig. 26 divided into simplices, and then stacks a number of such cubes in such a way as to construct an arbitrarily large lattice, as shown in Fig. 32. Other lattice structures are of course possible, including even a random lattice. The expectation is that for long range correlations involving distance scales much larger than the lattice spacing the precise structure of the underlying lattice structure will not matter.

The lattice sizes investigated typically range from 4^4 sites (3840 edges) to 32^4 sites (15,728,640

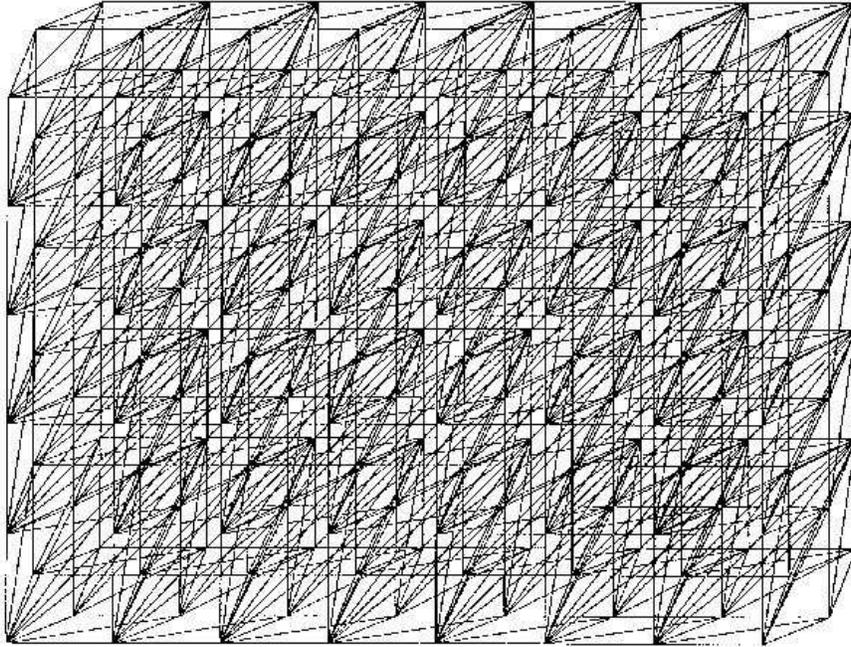


FIG. 32 Four-dimensional hypercubes divided into simplices and stacked to form a four-dimensional lattice.

edges). On a dedicated massively parallel supercomputer millions of consecutive edge length configurations can be generated for tens of values of k in a few months time.

Even though these lattices are not very large, one should keep in mind that due to the simplicial nature of the lattice there are many edges per hypercube with many interaction terms, and as a consequence the statistical fluctuations can be comparatively small, unless measurements are taken very close to a critical point, and at rather large separation in the case of the correlation functions or the potential. In addition, extrapolations to the infinite volume limit can be aided by finite size scaling methods, which exploit predictable renormalization group properties of finite size systems.

Usually the topology is restricted to a four-torus, corresponding to periodic boundary conditions. One can perform similar calculations with lattices employing different boundary conditions or topology, but one would expect the universal scaling properties of the theory to be determined exclusively by short-distance renormalization effects. Indeed the Feynman rules of perturbation theory do not depend in any way on boundary terms, although some momentum integrals might require an infrared cutoff.

Based on physical considerations it would seem reasonable to impose the constraint that the scale of the curvature be much smaller than the inverse of the average lattice spacing, but still considerably larger than the inverse of the overall system size. Equivalently, that in momentum

space physical scales should be much smaller than the ultraviolet cutoff, but much larger than the infrared cutoff. A typical requirement is therefore that

$$l_0 \ll \xi \ll L, \quad (667)$$

where L is the linear size of the system, ξ the correlation length related for example to the large scale curvature by $\mathcal{R} \sim 1/\xi^2$, and l_0 the lattice spacing. Contrary to ordinary lattice field theories, the lattice spacing in lattice gravity is a dynamical quantity. Thus the quantity $l_0 = \sqrt{\langle l^2 \rangle}$ only represents an average cutoff parameter.

Furthermore the bare cosmological constant λ_0 appearing in the gravitational action of Eq. (421) can be fixed at 1 in units of the cutoff, since it just sets the overall length scale in the problem. The higher derivative coupling a can be set to a value very close to 0 since one ultimately is interested in the limit $a \rightarrow 0$, corresponding to the pure Einstein theory.

One finds that for the measure in Eq. (407) this choice of parameters leads to a well behaved ground state for $k < k_c$ for higher derivative coupling $a \rightarrow 0$. The system then resides in the ‘smooth’ phase, with an effective dimensionality close to four. On the other hand for $k > k_c$ the curvature becomes very large and the lattice collapses into degenerate configurations with very long, elongated simplices. Fig. 33 shows an example of a typical edge length distribution in the well behaved strong coupling phase close to but below k_c .

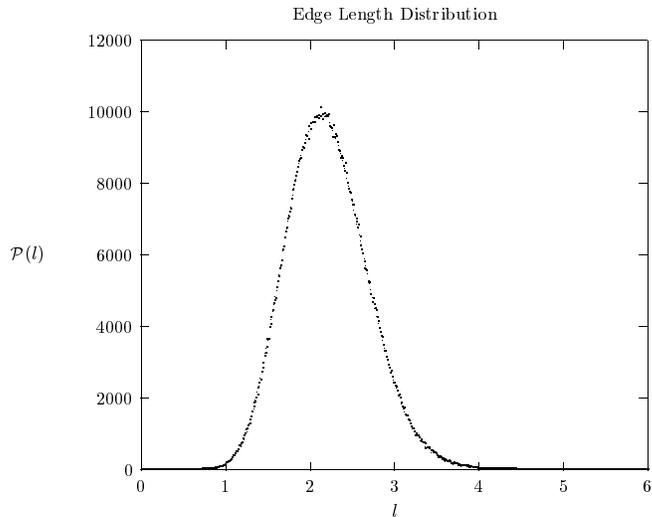


FIG. 33 A typical edge length distribution in the smooth phase for which $k < k_c$, or $G > G_c$. Note that the lattice gravitational measure of Eq. (407) cuts off the distribution at small edge lengths, while the cosmological constant term prevents large edge lengths from appearing.

Fig. 34 shows the corresponding curvature (δA or $\sqrt{g}R$) distribution.

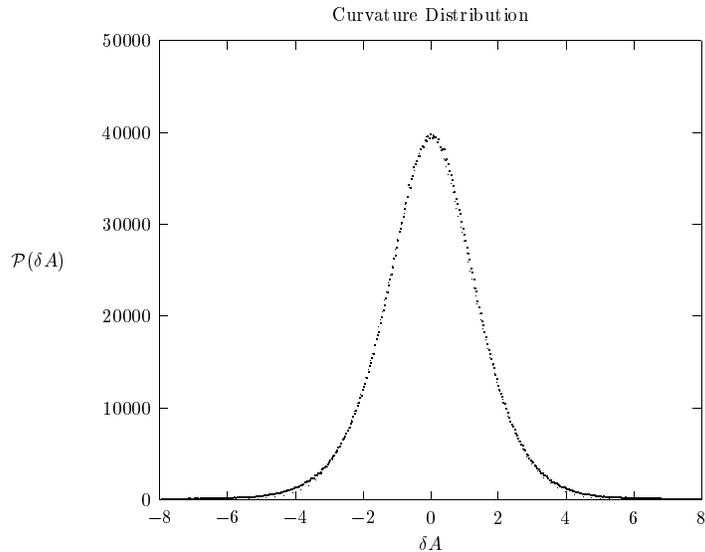


FIG. 34 A typical curvature distribution in the smooth phase for which $k < k_c$, or $G > G_c$. Note that distribution is peaked around close to zero curvature.

On one such edge length configuration one can compute the local curvatures, and then project the result on a two-dimensional plane. Using a red color coding for regions of increasingly positive curvature, and a blue color coding for regions of increasingly negative curvature, one obtains the picture in Fig. 35.

One finds that as k is varied, the average curvature \mathcal{R} is negative for sufficiently small k ('smooth' phase), and appears to go to zero continuously at some finite value k_c . For $k > k_c$ the curvature becomes very large, and the simplices tend to collapse into degenerate configurations with very small volumes ($\langle V \rangle / \langle l^2 \rangle^2 \sim 0$). This 'rough' or 'collapsed' phase is the region of the usual weak field expansion ($G \rightarrow 0$). In this phase the lattice collapses into degenerate configurations with very long, elongated simplices (Hamber, 1984; Hamber and Williams, 1985; Berg, 1985). This phenomenon is usually interpreted as a lattice remnant of the conformal mode instability of Euclidean gravity discussed in Sec. (II.B.4).

An elementary argument can be given to explain the fact that the collapsed phase for $k > k_c$ has an effective dimension of two. The instability is driven by the Euclidean Einstein term in the action, and in particular its unbounded conformal mode contribution. As the manifold during collapse reaches an effective dimension of two, the action effectively turns into a topological invariant,

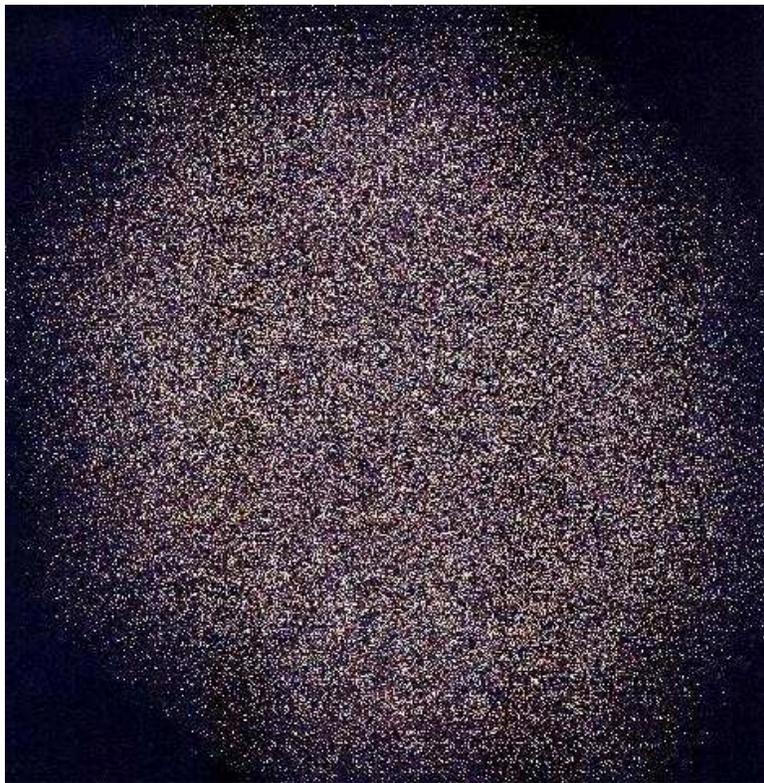


FIG. 35 Local curvatures projected on a plane, in the smooth phase $G > G_c$.

unable to drive the instability further to a still lower dimension ¹¹.

Accurate and reproducible curvature data can only be obtained for k below the instability point, since for $k > k_u \approx 0.053$ an instability develops, presumably associated with the unbounded conformal mode. Its signature is typical of a sharp first order transition, beyond which the system eventually tunnels into the rough, elongated phase which is two-dimensional in nature and has no physically acceptable continuum limit. The instability is caused by the appearance of one or more localized singular configuration, with a spike-like curvature singularity. At strong coupling such singular configurations are suppressed by a lack of phase space due to the functional measure, which imposes non-trivial constraints due to the triangle inequalities and their higher dimensional

¹¹ One way of determining coarse aspects of the underlying geometry is to compute the effective dimension in the scaling regime, for example by considering how the number of points within a thin shell of geodesic distance between τ and $\tau + \Delta$ scales with the geodesic distance itself. For distances a few multiples of the average lattice spacing one finds

$$N(\tau) \underset{\tau \rightarrow \infty}{\sim} \tau^{d_v}, \quad (668)$$

with $d_v = 3.1(1)$ for $G > G_c$ (the smooth phase) and $d_v \simeq 1.6(2)$ for $G < G_c$ (the rough phase). One concludes that in the rough phase the lattice tends to collapse into a degenerate tree-like configuration, whereas in the smooth phase the effective dimension of space-time is consistent with four. Higher derivative terms affect these results at very short distances, where they tend to make the geometry smoother.

analog. In other words, it seems that the measure regulates the conformal instability at sufficiently strong coupling.

As a consequence, the non-analytic behavior of the free energy (and its derivatives, which include for example the average curvature) has to be obtained by analytic continuation of the Euclidean theory into the metastable branch. This procedure, while perhaps unusual, is formally equivalent to the construction of the continuum theory exclusively from its strong coupling (small k , large G) expansion

$$\mathcal{R}(k) = \sum_{n=0}^{\infty} b_n k^n . \quad (669)$$

Ultimately it should be kept in mind that one is really only interested in the *pseudo-Riemannian* case, and not the Euclidean one for which an instability due to the conformal mode is to be expected. Indeed had such an instability not occurred for small enough G one might have wondered if the resulting theory still had any relationship to the original continuum theory.

6. Numerical Determination of the Scaling Exponents

One way to extract the critical exponent ν is to fit the average curvature to the form [see Eq. (662)]

$$\mathcal{R}(k) \underset{k \rightarrow k_c}{\sim} -A_{\mathcal{R}} (k_c - k)^{\delta} . \quad (670)$$

Using this set of procedures one obtains on lattices of up to 16^4 sites $k_c = 0.0630(11)$ and $\nu = 0.330(6)$. Note that the average curvature is negative at strong coupling up to the critical point: locally the parallel transport of vectors around infinitesimal loops seems to be described by a lattice version of Euclidean anti-de Sitter space.

Fig. 36 shows as an example a graph of the average curvature $\mathcal{R}(k)$ raised to the third power. One would expect to get a straight line close to the critical point if the exponent for $\mathcal{R}(k)$ is exactly $1/3$. The numerical data indeed supports this assumption, and in fact the linearity of the results close to k_c is quite striking. Using this procedure one obtains on the 16^4 -site lattice an estimate for the critical point, $k_c = 0.0639(10)$.

Often it can be advantageous to express results obtained in the cutoff theory in terms of physical (i.e. cutoff independent) quantities. By the latter one means quantities for which the cutoff dependence has been re-absorbed, or restored, in the relevant definition. As an example, an expression equivalent to Eq. (662), relating the vacuum expectation value of the local scalar curvature to the

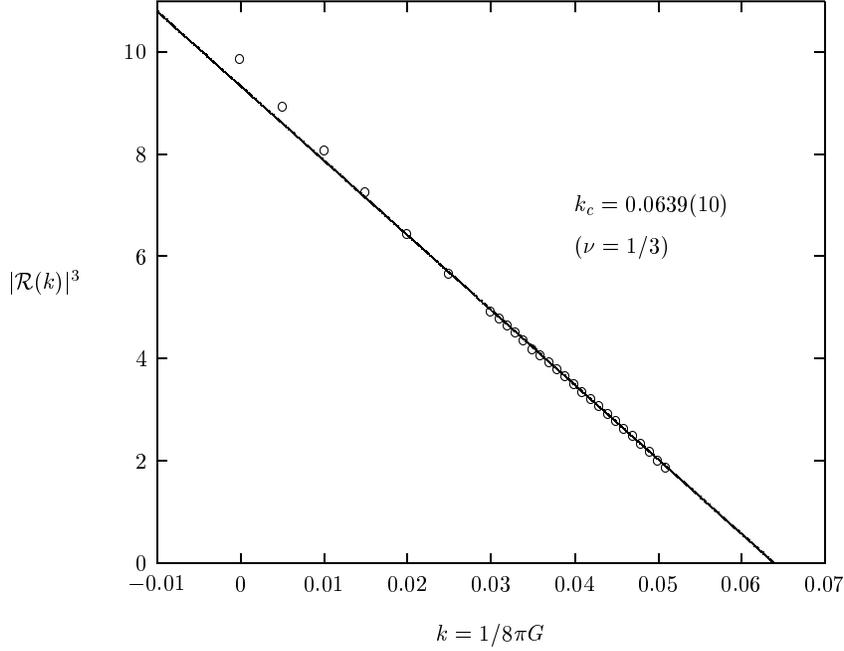


FIG. 36 Average curvature \mathcal{R} on a 16^4 lattice, raised to the third power. If $\nu = 1/3$, the data should fall on a straight line. The continuous line represents a linear fit of the form $A(k - k_c)$. The small deviation from linearity of the transformed data is quite striking.

physical correlation length ξ , is

$$\frac{\langle \int dx \sqrt{g} R(x) \rangle}{\langle \int dx \sqrt{g} \rangle} \underset{G \rightarrow G_c}{\sim} \text{const.} \left(l_P^2 \right)^{(d-2-1/\nu)/2} \left(\frac{1}{\xi^2} \right)^{(d-1/\nu)/2}, \quad (671)$$

which is obtained by substituting Eq. (659) into Eq. (662). The correct dimensions have been restored in this last equation by supplying appropriate powers of the Planck length $l_P = G_{phys}^{1/(d-2)}$, which involves the ultraviolet cutoff Λ . For $\nu = 1/3$ the result of Eq. (671) becomes particularly simple

$$\frac{\langle \int dx \sqrt{g} R(x) \rangle}{\langle \int dx \sqrt{g} \rangle} \underset{G \rightarrow G_c}{\sim} \text{const.} \frac{1}{l_P \xi} \quad (672)$$

Note that a naive estimate based on dimensional arguments would have suggested the incorrect result $\sim 1/l_P^2$. Instead the above expression actually vanishes at the critical point. This shows that ν plays the role of an anomalous dimension, determining the magnitude of deviations from naive dimensional arguments.

Since the critical exponents play such a central role in determining the existence and nature of the continuum limit, it appears desirable to have an independent way of estimating them, which either does not depend on any fitting procedure, or at least analyzes a different and complementary

set of data. By systematically studying the dependence of averages on the physical size of the system, one can independently estimate the critical exponents.

Reliable estimates for the exponents in a lattice field theory can take advantage of a comprehensive finite-size analysis, a procedure by which accurate values for the critical exponents are obtained by taking into account the linear size dependence of the result computed in a finite volume V .

In practice the renormalization group approach is brought in by considering the behavior of the system under scale transformations. Later the scale dependence is applied to the overall volume itself. The usual starting point for the derivation of the scaling properties is the renormalization group behavior of the free energy $F = -\frac{1}{V} \log Z$

$$F(t, \{u_j\}) = F_{reg}(t, \{u_j\}) + b^{-d} F_{sing}(b^{y_t} t, \{b^{y_j} u_j\}) , \quad (673)$$

where F_{sing} is the singular, non-analytic part of the free energy, and F_{reg} is the regular part. b is the block size in the RG transformation, while y_t and y_j ($j \geq 2$) are the relevant eigenvalues of the RG transformation, and t the reduced temperature variable that gives the distance from criticality. One denotes here by $y_t > 0$ the relevant eigenvalue, while the remaining eigenvalues $y_j \leq 0$ are associated with either marginal or irrelevant operators. Usually the leading critical exponent y_t^{-1} is called ν , while the next subleading exponent y_2 is denoted $-\omega$. For more details on the method we have to refer to the comprehensive review in (Cardy, 1988).

The correlation length ξ determines the asymptotic decay of correlations, in the sense that one expects for example for the two-point function at large distances

$$\langle \mathcal{O}(x) \mathcal{O}(y) \rangle \underset{|x-y| \gg \xi}{\sim} e^{-|x-y|/\xi} . \quad (674)$$

The scaling equation for the correlation length itself

$$\xi(t) = b \xi(b^{y_t} t) \quad (675)$$

implies for $b = t^{-1/y_t}$ that $\xi \sim t^{-\nu}$ with a correlation length exponent

$$\nu = 1/y_t . \quad (676)$$

Derivatives of the free energy F with respect to t then determine, after setting the scale factor $b = t^{-1/y_t}$, the scaling properties of physical observables, including corrections to scaling. Thus for example, the second derivative of the free energy with respect to t yields the specific heat exponent $\alpha = 2 - d/y_t = 2 - d\nu$,

$$\frac{\partial^2}{\partial t^2} F(t, \{u_j\}) \sim t^{-(2-d\nu)} . \quad (677)$$

In the gravitational case one identifies the scaling field t with $k_c - k$, where $k = 1/16\pi G$ involves the bare Newton's constant. The appearance of singularities in physical averages, obtained from appropriate derivatives of F , is rooted in the fact that close to the critical point at $t = 0$ the correlation length diverges.

The above results can be extended to the case of a finite lattice of volume V and linear dimension $L = V^{1/d}$. The volume-dependent free energy is then written as

$$F(t, \{u_j\}, L^{-1}) = F_{reg}(t, \{u_j\}) + b^{-d} F_{sing}(b^{y_t} t, \{b^{y_j} u_j\}, b/L) \quad . \quad (678)$$

For $b = L$ (a lattice consisting of only one point) one obtains the Finite Size Scaling (FSS) form of the free energy [see for example (Brezin and Zinn-Justin, 1985) for a field-theoretic justification]. After taking derivatives with respect to the fields t and $\{u_j\}$, the FSS scaling form for physical observables follows,

$$O(L, t) = L^{x_O/\nu} \left[\tilde{f}_O(L t^\nu) + \mathcal{O}(L^{-\omega}) \right] \quad , \quad (679)$$

where x_O is the scaling dimension of the operator O , and $\tilde{f}_O(x)$ an arbitrary function.

As an example, consider the average curvature \mathcal{R} . From Eq. (679), with $t \sim k_c - k$ and $x_O = 1 - 4\nu$, one has

$$\mathcal{R}(k, L) = L^{-(4-1/\nu)} \left[\tilde{\mathcal{R}}\left((k_c - k) L^{1/\nu}\right) + \mathcal{O}(L^{-\omega}) \right] \quad (680)$$

where $\omega > 0$ is a correction-to-scaling exponent. If scaling involving k and L holds according to Eq. (679), then all points should lie on the same universal curve.

Fig. 37 shows a graph of the scaled curvature $\mathcal{R}(k) L^{4-1/\nu}$ for different values of $L = 4, 8, 16$, versus the scaled coupling $(k_c - k)L^{1/\nu}$. If scaling involving k and L holds according to Eq. (680), with $x_O = 1 - 4\nu$ the scaling dimension for the curvature, then all points should lie on the same universal curve. The data is in good agreement with such behavior, and provides a further test on the exponent, which seems consistent within errors with $\nu = 1/3$.

As a second example consider the curvature fluctuation $\chi_{\mathcal{R}}$. From the general Eq. (679) one expects in this case, for $t \sim k_c - k$ and $x_O = 2 - 4\nu$,

$$\chi_{\mathcal{R}}(k, L) = L^{2/\nu-4} \left[\tilde{\chi}_{\mathcal{R}}\left((k_c - k) L^{1/\nu}\right) + \mathcal{O}(L^{-\omega}) \right] \quad , \quad (681)$$

where $\omega > 0$ is again a correction-to-scaling exponent. If scaling involving k and L holds according to Eq. (679), then all points should lie on the same universal curve.

Fig. 38 shows a graph of the scaled curvature fluctuation $\chi_{\mathcal{R}}(k)/L^{2/\nu-4}$ for different values of $L = 4, 8, 16$, versus the scaled coupling $(k_c - k)L^{1/\nu}$. If scaling involving k and L holds according

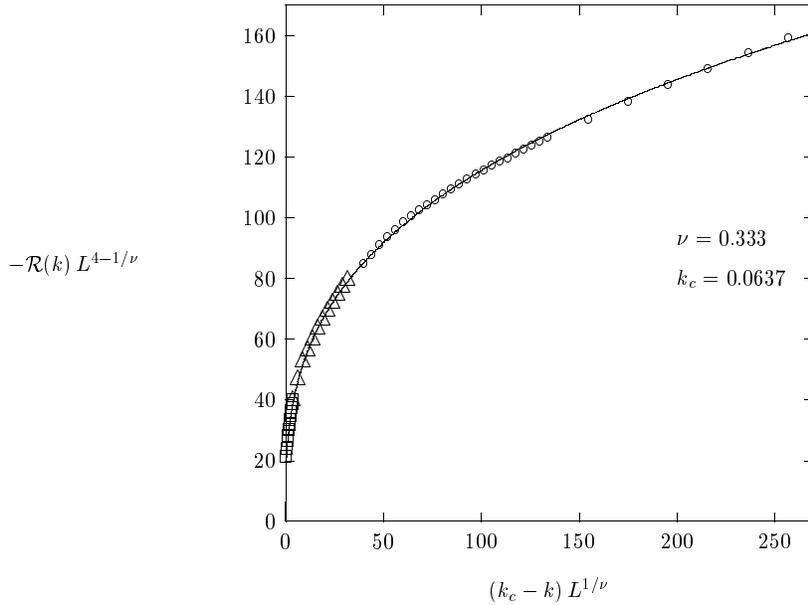


FIG. 37 Finite size scaling behavior of the scaled curvature versus the scaled coupling. Here $L = 4$ for the lattice with 4^4 sites (\square), $L = 8$ for a lattice with 8^4 sites (\triangle), and $L = 16$ for the lattice with 16^4 sites (\circ). Statistical errors are comparable to the size of the dots. The continuous line represents a best fit of the form $a + bx^c$. Finite size scaling predicts that all points should lie on the same universal curve. At $k_c = 0.0637$ the scaling plot gives the value $\nu = 0.333$.

to Eq. (681), then all points should lie on the same universal curve. Again the data supports such scaling behavior, and provides a further estimate on the value for ν .

The value of k_c itself should depend on the size of the system. One expects

$$k_c(L) \underset{L \rightarrow \infty}{\sim} k_c(\infty) + c L^{-1/\nu} + \dots \quad (682)$$

where $k_c(\infty)$ is the infinite-volume limit critical point.

The previous discussion applies to continuous, second order phase transitions. First order phase transitions are driven by instabilities, and are in general not governed by any renormalization group fixed point. The underlying reason is that the correlation length does not diverge at the first order transition point, and thus the system never becomes scale invariant. In the simplest case, a first order transition develops as the system tunnels between two neighboring minima of the free energy. In the metastable branch the free energy acquires a small complex part with a very weak essential singularity in the coupling at the first order transition point (Langer, 1967; Fisher, 1967; Griffiths, 1969). As a consequence, such a singularity is not generally visible from the stable branch, in the sense that a power series expansion in the temperature is unaffected by such a weak singularity.

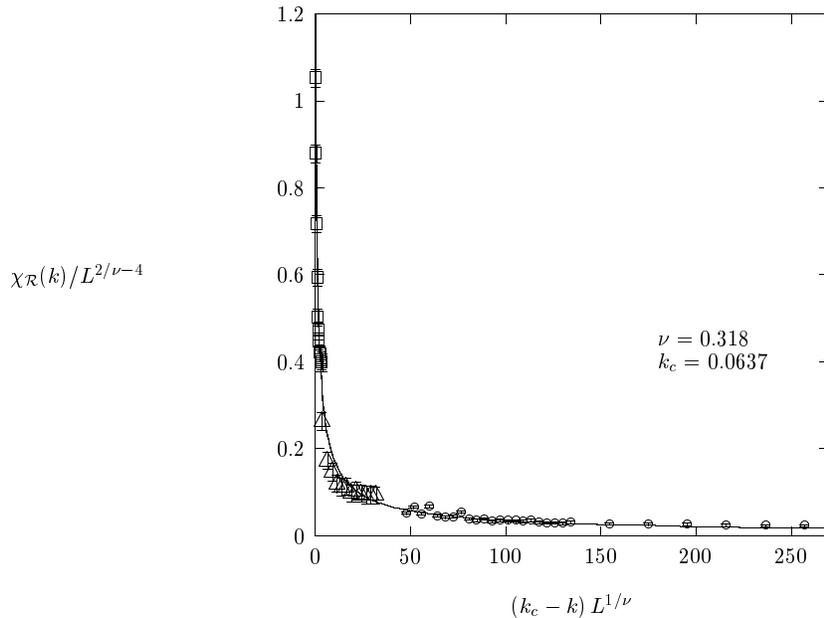


FIG. 38 Finite size scaling behavior of the scaled curvature fluctuation versus the scaled coupling. Here $L = 4$ for the lattice with 4^4 sites (\square), $L = 8$ for the lattice with 8^4 sites (\triangle), and $L = 16$ for the lattice with 16^4 sites (\circ). The continuous line represents a best fit of the form $1/(a + bx^c)$. Finite size scaling predicts that all points should lie on the same universal curve. At $k_c = 0.0637$ the scaling plot gives the value $\nu = 0.318$.

Indeed in the infinite volume limit the singularities associated with a first order transition at T_u become infinitely sharp, a θ - or δ -function type singularity. While the singularity in the free energy at the endpoint of the metastable branch (at say T_c) cannot be explored directly, it can be reached by a suitable analytic continuation from the stable side of the free energy branch. A similar situation arises in the case of lattice QCD with fermions, where zero fermion mass (chiral) limit is reached by extrapolation (Hamber and Parisi, 1983).

From the best data (with the smallest statistical uncertainties and the least systematic effects) one concludes

$$k_c = 0.0636(11) \quad \nu = 0.335(9) \quad , \quad (683)$$

which suggests $\nu = 1/3$ for pure quantum gravity. Note that at the critical point the gravitational coupling is not weak, $G_c \approx 0.626$ in units of the ultraviolet cutoff. It seems that the value $\nu = 1/3$ does not correspond to any known field theory or statistical mechanics model in four dimensions. For a perhaps related system, namely dilute branched polymers, it is known that $\nu = 1/2$ in three dimensions, and $\nu = 1/4$ at the upper critical dimension $d = 8$, so one would expect a value close

Method	ν^{-1} in $d = 3$	ν^{-1} in $d = 4$
lattice	1.67(6)	-
lattice	-	2.98(7)
$2 + \epsilon$	1.6	4.4
truncation	1.2	2.666
exact ?	1.5882	3

Table I: Direct determinations of the critical exponent ν^{-1} for quantum gravitation, using various analytical and numerical methods in three and four space-time dimensions.

to $1/3$ somewhere in between. On the other hand for a scalar field one would have obtained $\nu = 1$ in $d = 2$ and $\nu = \frac{1}{2}$ for $d \geq 4$, which seems excluded.

Table I provides a short summary of the critical exponents for quantum gravitation as obtained by various perturbative and non-perturbative methods in three and four dimensions. The $2 + \epsilon$ and the truncation method results were discussed previously in Secs. II.C.4 and II.C.5, respectively. The lattice model of Eq. (421) in four dimensions gives for the critical point $G_c \approx 0.626$ in units of the ultraviolet cutoff, and $\nu^{-1} = 2.98(7)$ which is used for comparison in Table I. In three dimensions the numerical results are consistent with the universality class of the interacting scalar field. The same set of results are compared graphically in Fig. 39 and Fig. 40 below.

The direct numerical determinations of the critical point $k_c = 1/8\pi G_c$ in $d = 3$ and $d = 4$ space-time dimensions are in fact quite close to the analytical prediction of the lattice $1/d$ expansion given previously in Eq. (599),

$$\frac{k_c}{\lambda_0^{1-2/d}} = \frac{2^{1+2/d}}{d^3} \left[\frac{\Gamma(d)}{\sqrt{d+1}} \right]^{2/d}. \quad (684)$$

The above expression gives for a bare cosmological constant $\lambda_0 = 1$ the estimate $k_c = \sqrt{3}/(16 \cdot 5^{1/4}) = 0.0724$ in $d = 4$, to be compared with the numerical result $k_c = 0.0636(11)$ in (Hamber, 2000). Even in $d = 3$ one has $k_c = 2^{5/3}/27 = 0.118$, to be compared with the direct determination $k_c = 0.112(5)$ from (Hamber and Williams, 1993). These estimates are compared below in Fig. 40.

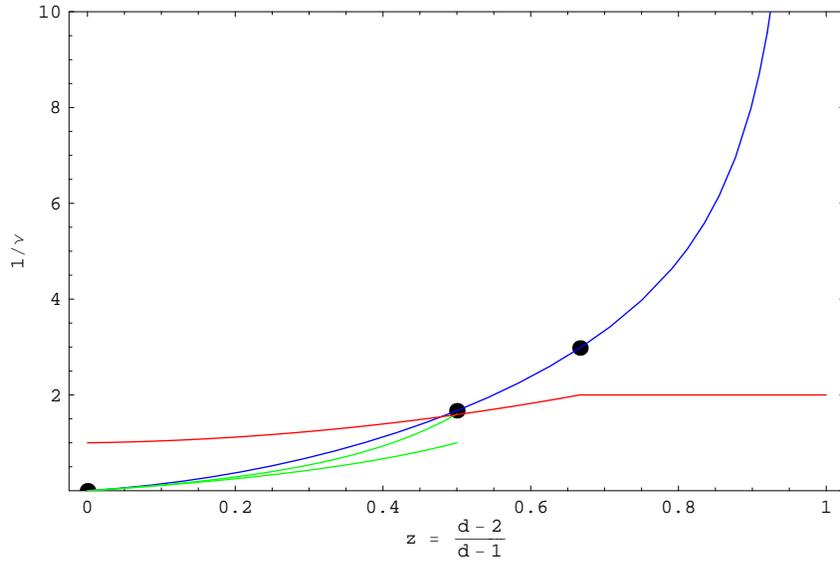


FIG. 39 Universal gravitational exponent $1/\nu$ as a function of the dimension. The abscissa is $z = (d - 2)/(d - 1)$, which maps $d = 2$ to $z = 0$ and $d = \infty$ to $z = 1$. The larger circles at $d = 3$ and $d = 4$ are the lattice gravity results, interpolated (continuous curve) using the exact lattice results $1/\nu = 0$ in $d = 2$, and $\nu = 0$ at $d = \infty$ [from Eq. (608)]. The two curves close to the origin are the $2 + \epsilon$ expansion for $1/\nu$ to one loop (lower curve) and two loops (upper curve). The lower almost horizontal line gives the value for ν expected for a scalar field theory, for which it is known that $\nu = 1$ in $d = 2$ and $\nu = \frac{1}{2}$ in $d \geq 4$.

7. Renormalization Group and Lattice Continuum Limit

The discussion in the previous sections points to the existence of a phase transition in the lattice gravity theory, with divergent correlation length in the vicinity of the critical point, as in Eq. (659)

$$\xi(k) \underset{k \rightarrow k_c}{\sim} A_\xi |k_c - k|^{-\nu} \quad (685)$$

As described previously, the existence of such a correlation length is usually inferred indirectly by scaling arguments, from the presence of singularities in the free energy $F_{latt} = -\frac{1}{V} \ln Z_{latt}$ as a function of the lattice coupling k . Equivalently, ξ could have been computed directly from correlation functions at fixed geodesic distance using the definition in Eq. (641), or even from the correlation of Wilson lines associated with the propagation of two heavy spinless particles. The outcome of such large scale numerical calculations is eventually a determination of the quantities ν , $k_c = 1/8\pi G_c$ and A_ξ from first principles, to some degree of numerical accuracy.

In either case one expects the scaling result of Eq. (685) close to the fixed point, which we choose to rewrite here in terms of the inverse correlation length $m \equiv 1/\xi$

$$m = \Lambda A_m |k - k_c|^\nu . \quad (686)$$

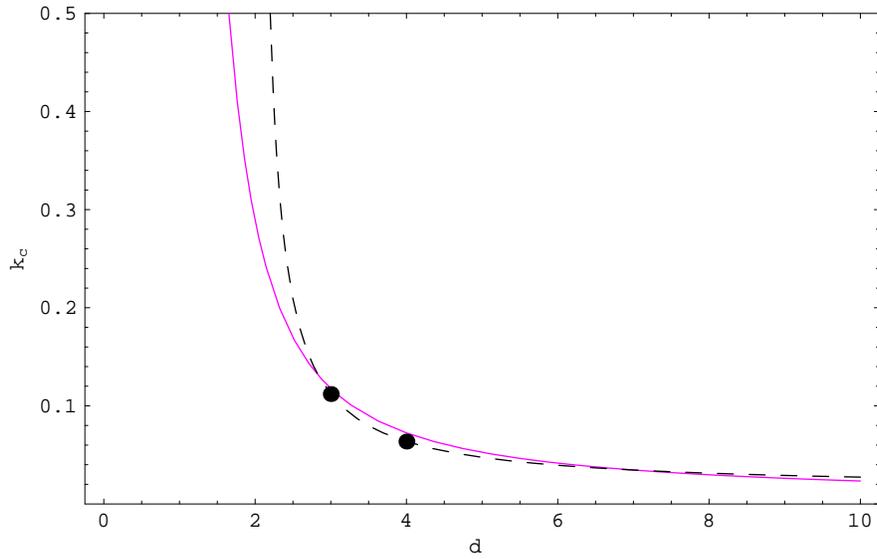


FIG. 40 Critical point $k_c = 1/8\pi G_c$ in units of the ultraviolet cutoff as a function of dimension d . The circles at $d = 3$ and $d = 4$ are the lattice results, suitably interpolated (dashed curve) using the additional lattice result $1/k_c = 0$ in $d = 2$. The lower continuous curve is the analytical large- d lattice result of Eq. (599).

Note that in the above expression the correct dimension for m (inverse length) has been restored by inserting explicitly on the r.h.s. the ultraviolet cutoff Λ . Here k and k_c are of course still dimensionless quantities, and correspond to the bare microscopic couplings at the cutoff scale, $k \equiv k(\Lambda) \equiv 1/(8\pi G(\Lambda))$. A_m is a calculable numerical constant, related to A_ξ in Eq. (659) by $A_m = A_\xi^{-1}$. It is worth pointing out that the above expression for $m(k)$ is almost identical in structure to the one for the non-linear σ -model in the $2 + \epsilon$ expansion, Eq. (221) and in the large N limit, Eqs. (244), (245) and (249). It is of course also quite similar to $2 + \epsilon$ result for continuum gravity, Eq. (304).

The lattice continuum limit corresponds to the large cutoff limit taken at *fixed* m ,

$$\Lambda \rightarrow \infty, \quad k \rightarrow k_c, \quad m \text{ fixed}, \quad (687)$$

which shows that the continuum limit is reached in the vicinity of the ultraviolet fixed point. Phrased equivalently, one takes the limit in which the lattice spacing $a \approx 1/\Lambda$ is sent to zero at *fixed* $\xi = 1/m$, which requires an approach to the non-trivial UV fixed point $k \rightarrow k_c$. The quantity m is supposed to be a renormalization group invariant, a physical scale independent of the scale at which the theory is probed. In practice, since the cutoff ultimately determines the physical value of Newton's constant G , Λ cannot be taken to ∞ . Instead a very large value will suffice, $\Lambda^{-1} \sim 10^{-33} \text{cm}$, for which it will still be true that $\xi \gg \Lambda$ which is all that is required for the continuum limit.

For discussing the renormalization group behavior of the coupling it will be more convenient to write the result of Eq. (686) directly in terms of Newton's constant G as

$$m = \Lambda \left(\frac{1}{a_0} \right)^\nu \left[\frac{G(\Lambda)}{G_c} - 1 \right]^\nu, \quad (688)$$

with the dimensionless constant a_0 related to A_m by $A_m = 1/(a_0 k_c)^\nu$. Note that the above expression only involves the dimensionless ratio $G(\Lambda)/G_c$, which is the only relevant quantity here. The lattice theory in principle completely determines both the exponent ν and the amplitude a_0 for the quantum correction. Thus from the knowledge of the dimensionless constant A_m in Eq. (686) one can estimate from first principles the value of a_0 in Eqs. (693). Lattice results for the correlation functions at fixed geodesic distance give a value for $A_m \approx 0.72$ with a significant uncertainty, which, when combined with the values $k_c \simeq 0.0636$ and $\nu \simeq 0.335$ given above, gives $a_0 = 1/(k_c A_m^{1/\nu}) \simeq 42$. The rather surprisingly large value for a_0 appears here as a consequence of the relatively small value of the lattice k_c in four dimensions.

The renormalization group invariance of the physical quantity m requires that the running gravitational coupling $G(\mu)$ varies in the vicinity of the fixed point in accordance with the above equation, with $\Lambda \rightarrow \mu$, where μ is now an arbitrary scale,

$$m = \mu \left(\frac{1}{a_0} \right)^\nu \left[\frac{G(\mu)}{G_c} - 1 \right]^\nu, \quad (689)$$

The latter is equivalent to the renormalization group invariance requirement

$$\mu \frac{d}{d\mu} m(\mu, G(\mu)) = 0 \quad (690)$$

provided $G(\mu)$ is varied in a specific way. Indeed this type of situation was already encountered before, for example in Eqs. (247) and (305). Eq. (690) can therefore be used to obtain, if one so wishes, a β -function for the coupling $G(\mu)$ in units of the ultraviolet cutoff,

$$\mu \frac{\partial}{\partial \mu} G(\mu) = \beta(G(\mu)), \quad (691)$$

with $\beta(G)$ given in the vicinity of the non-trivial fixed point, using Eq. (689), by

$$\beta(G) \equiv \mu \frac{\partial}{\partial \mu} G(\mu) \Big|_{G \rightarrow G_c} \simeq -\frac{1}{\nu} (G - G_c) + \dots \quad (692)$$

The above procedure is in fact in complete analogy to what was done for the non-linear σ -model, for example in Eq. (249). But the last two steps are not really necessary, for one can obtain the scale dependence of the gravitational coupling directly from Eq. (689), by simply solving for $G(\mu)$,

$$G(\mu) = G_c \left[1 + a_0 (m^2/\mu^2)^{1/2\nu} + \dots \right] \quad (693)$$

This last expression can be compared directly to the $2 + \epsilon$ result of Eq. (300), as well as to the σ -model result of Eq. (207). The physical dimensions of G can be restored by multiplying the above expression on both sides by the ultraviolet cutoff Λ , if one so desires. One concludes that the above result physically implies gravitational anti-screening: the gravitational coupling G increases with distance.

Note that the last equation only involves the dimensionless ratio $G(\mu)/G_c$, and is therefore unaffected by whether the coupling G is dimensionful (after inserting an appropriate power of the cutoff Λ) or dimensionless. It simply relates the gravitational coupling at one scale to the coupling at a different scale,

$$\frac{G(\mu_1)}{G(\mu_2)} \approx \frac{1 + a_0(m/\mu_1)^{1/\nu} + \dots}{1 + a_0(m/\mu_2)^{1/\nu} + \dots} \quad (694)$$

In conclusion, the lattice result for $G(\mu)$ in Eq. (693) and the β -function in Eq. (692) are qualitatively similar to what one finds both in the $2 + \epsilon$ expansion for gravity and in the non-linear σ -model *in the strong coupling phase*.

But there are also significant differences. Besides the existence of a phase transition between two geometrically rather distinct phases, one major new aspect provided by non-perturbative lattice studies is the fact that the weakly coupled small G phase turns out to be pathological, in the sense that the theory becomes unstable, with the four-dimensional lattice collapsing into a tree-like two-dimensional structure for $G < G_c$. While in continuum perturbation theory both phases, and therefore both signs for the coupling constant evolution in Eq. (297), seem acceptable (giving rise to both a ‘‘Coulomb’’ phase, and a strong coupling phase), the Euclidean lattice results rule out the small $G < G_c$ branched polymer phase. The collapse eventually stops at $d = 2$ because the gravitational action then becomes a topological invariant.

It appears difficult therefore to physically characterize the weak coupling phase based on just the lattice results, which only seem to make sense in the strong coupling phase $G > G_c$. One could envision an approach wherein such a weak coupling phase would be discussed in the framework of some sort of analytic continuation from the strong coupling phase, which would seem possible at least for some lattice results, such as the gravitational β -function of Eq. (692). The latter clearly makes sense on both sides of the transition, just as is the case for Eq. (249) for the non-linear σ -model. In particular Eq. (692) implies that the coupling will flow towards the Gaussian fixed point $G = 0$ for $G < G_c$. The scale dependence in this phase will be such that one expects gravitational screening: the coupling $G(\mu)$ gets increasingly weaker at larger distances. But how to remove the geometric collapse to a two-dimensional manifold remains a major hurdle; one could

envision an approach where one introduces one more cutoff on the edges at short distances, so that each simplex cannot go below a certain fatness. But if the results so far can be used as a guide, the gradual removal of such a cutoff would then plunge the theory back into a degenerate two-dimensional, and therefore physically unacceptable, geometry.

8. Curvature Scales

As can be seen from Eqs. (264) and (630) the path integral for pure quantum gravity can be made to depend on the gravitational coupling G and the cutoff Λ only: by a suitable rescaling of the metric, or the edge lengths in the discrete case, one can set the cosmological constant to unity in units of the cutoff. The remaining coupling G should then be viewed more appropriately as the gravitational constant *in units of the cosmological constant* λ .

The renormalization group running of $G(\mu)$ in Eq. (693) involves an invariant scale $\xi = 1/m$. At first it would seem that this scale could take any value, including very small ones based on the naive estimate $\xi \sim l_P$, which would preclude any observable quantum effects in the foreseeable future. But the result of Eqs. (671) and (672) suggest otherwise, namely that the non-perturbative scale ξ is in fact related to *curvature*. From astrophysical observation the average curvature is very small, so one would conclude from Eq. (672) that ξ is very large, and possibly macroscopic. But the problem with Eq. (672) is that it involves the lattice Ricci scalar, a quantity related curvature probed by parallel transporting vectors around infinitesimal loops with size comparable to the lattice cutoff Λ^{-1} . What one would like is instead a relationship between ξ and quantities which describe the geometry on larger scales.

In many ways the quantity m of Eq. (689) behaves as a dynamically generated mass scale, quite similar to what happens in the non-linear σ -model case [Eq. (245)], or in the $2 + \epsilon$ gravity case [Eq. (301)]. Indeed in the weak field expansion, presumably appropriate for slowly varying fields on very large scales, a mass-like term does appear, as in Eq. (64), with $\mu^2 = 16\pi G\lambda_0 \equiv 2\lambda$ where λ is the scaled cosmological constant. From the classical field equation $R = 4\lambda$ one can relate the above λ , and therefore the mass-like parameter m , to curvature, which leads to the identification

$$\lambda_{obs} \simeq \frac{1}{\xi^2} \tag{695}$$

with λ_{obs} the observed small but non-vanishing cosmological constant.

A further indication that the identification of the observed scaled cosmological constant with a mass-like - and therefore renormalization group invariant - term makes sense beyond the weak field

limit can be seen for example by comparing the structure of the three classical field equations

$$\begin{aligned}
R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \lambda g_{\mu\nu} &= 8\pi G T_{\mu\nu} \\
\partial^\mu F_{\mu\nu} + \mu^2 A_\nu &= 4\pi e j_\nu \\
\partial^\mu \partial_\mu \phi + m^2 \phi &= \frac{g}{3!} \phi^3
\end{aligned} \tag{696}$$

for gravity, QED (massive via the Higgs mechanism) and a self-interacting scalar field, respectively.

A third argument suggesting the identification of the scale ξ with large scale curvature and therefore with the observed scaled cosmological constant goes as follows. Observationally the curvature on large scale can be determined by parallel transporting vectors around very large loops, with typical size much larger than the lattice cutoff l_P . In gravity, curvature is detected by parallel transporting vectors around closed loops. This requires the calculation of a path dependent product of Lorentz rotations \mathbf{R} , in the Euclidean case elements of $SO(4)$, as discussed in Sec. III.A.2. On the lattice, the above rotation is directly related to the path-ordered (\mathcal{P}) exponential of the integral of the lattice affine connection $\Gamma_{\mu\nu}^\lambda$ via

$$R^\alpha_\beta = \left[\mathcal{P} e^{\int_{\text{path between simplices}} \Gamma^\lambda dx_\lambda} \right]^\alpha_\beta . \tag{697}$$

Now, in the strongly coupled gravity regime ($G > G_c$) large fluctuations in the gravitational field at short distances will be reflected in large fluctuations of the \mathbf{R} matrices. Deep in the strong coupling regime it should be possible to describe these fluctuations by a uniform (Haar) measure. Borrowing from the analogy with Yang-Mills theories, and in particular non-Abelian lattice gauge theories with compact groups [see Eq. (328)], one would therefore expect an exponential decay of near-planar Wilson loops with area A of the type

$$W(\Gamma) \sim \text{tr} \exp \left[\int_{S(C)} R \cdot_{\mu\nu} A_C^{\mu\nu} \right] \sim \exp(-A/\xi^2) \tag{698}$$

where A is the minimal physical area spanned by the near-planar loop. A derivation of this standard result for non-Abelian gauge theories can be found, for example, in the textbook (Peskin and Schroeder, 1995).

In summary, the Wilson loop in gravity provides potentially a measure for the magnitude of the large-scale, averaged curvature, operationally determined by the process of parallel-transporting test vectors around very large loops, and which therefore, from the above expression, is computed to be of the order $R \sim 1/\xi^2$. One would expect the power to be universal, but not the amplitude, leaving open the possibility of having both de Sitter or anti-de Sitter space at large distances (as discussed previously in Sec. IV.A.6, the average curvature describing the parallel transport

of vectors around *infinitesimal* loops is described by a lattice version of Euclidean anti-de Sitter space). A recent explicit lattice calculation indeed suggests that the de Sitter case is singled out, at least for sufficiently strong coupling (Hamber and Williams, 2007). Furthermore one would expect, based on general scaling arguments, that such a behavior would persist throughout the whole strong coupling phase $G > G_c$, all the way up to the non-trivial fixed point. From it then follows the identification of the correlation length ξ with a measure of large scale curvature, the most natural candidate being the scaled cosmological constant λ_{phys} , as in Eq. (695). This relationship, taken at face value, implies a very large, cosmological value for $\xi \sim 10^{28} cm$, given the present bounds on λ_{phys} . Other closely related possibilities may exist, such as an identification of ξ with the Hubble constant (as measured today), determining the macroscopic expansion rate of the universe via the correspondence $\xi \simeq 1/H_0$. Since this quantity is presumably time-dependent, a possible scenario would be one in which $\xi^{-1} = H_\infty = \lim_{t \rightarrow \infty} H(t)$, with $H_\infty^2 = \frac{\lambda}{3}$, for which the horizon radius is $R_\infty = H_\infty^{-1}$.

Since, as pointed out in Sec. II.C.4 and Sec. IV.A.1, the gravitational path integral only depends in a non-trivial way on the dimensionless combination $G \sqrt{\lambda_0}$, the physical Newton's constant itself G can be decomposed into non-running and running parts as

$$G = \frac{1}{G \lambda_0} \cdot G^2 \lambda_0 \rightarrow \xi^2 \cdot \left[(G \sqrt{\lambda_0}) (\mu^2) \right]^2 \quad (699)$$

where we have used $1/G\lambda_0 \sim \xi^2$. The running of the second, dimensionless term in square brackets can be directly deduced from either Eqs. (689) or (693). Note that there λ_0 does not appear there explicitly, since originally it was set equal to one by scaling the metric (or the edge lengths).

In conclusion, the modified Einstein equations, incorporating the quantum running of G , read

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \lambda g_{\mu\nu} = 8\pi G(\mu) T_{\mu\nu} \quad (700)$$

with $\lambda \simeq \frac{1}{\xi^2}$, and only $G(\mu)$ on the r.h.s. scale-dependent in accordance with Eq. (693). The precise meaning of $G(\mu)$ in a covariant framework will be given later in Sec. V.A.

9. Gravitational Condensate

In strongly coupled gravity there appears to be a deep relationship, already encountered previously in non-Abelian gauge theories, between the non-perturbative scale ξ appearing in Eqs. (693), and the non-perturbative vacuum condensate of Eqs. (671) and (695), which is a measure of curvature. The inescapable conclusion of the results of Eqs. (671) and (397) is that the scale ξ appearing

in Eq. (693) is related to curvature, and must be *macroscopic* for the lattice theory to be consistent. How can quantum effects propagate to such large distances and give such drastic modifications to gravity? The answer to this paradoxical question presumably lies in the fact that gravitation is carried by a massless particle whose interactions cannot be screened, on any length scale.

It is worth pointing out here that the gravitational vacuum condensate, which only exists in the strong coupling phase $G > G_c$, and which is proportional to the curvature, is genuinely non-perturbative. One can summarize the result of Eq. (695) as

$$\mathcal{R}_{obs} \simeq (10^{-30}eV)^2 \sim \xi^{-2} \quad (701)$$

where the condensate is, according to Eq. (688), non-analytic at $G = G_c$. A graviton vacuum condensate of order $\xi^{-1} \sim 10^{-30}eV$ is of course extraordinarily small compared to the QCD color condensate ($\Lambda_{\overline{MS}} \simeq 220 MeV$) and the electro-weak Higgs condensate ($v \simeq 250 GeV$). One can pursue the analogy with non-Abelian gauge theories further by stating that the quantum gravity theory cannot provide a value for the non-perturbative curvature scale ξ : it needs to be fixed by some sort of phenomenological input, either by Eq. (693) or by Eq. (695). But one important message is that the scale ξ in those two equations is one and the same.

Can the above physical picture be used to provide further insight into the nature of the phase transition, and more specifically the value for ν ? We will mention here a simple geometric argument which can be given to support the exact value $\nu = 1/3$ for pure gravity (Hamber and Williams, 2004). First one notices that the vacuum polarization induced scale dependence of the gravitational coupling $G(r)$ as given in Eq. (693) implies the following general structure for the quantum corrected static gravitational potential,

$$V(r) = -G(r) \frac{mM}{r} \approx -G(0) \frac{mM}{r} \left[1 + c(r/\xi)^{1/\nu} + \mathcal{O}((r/\xi)^{2/\nu}) \right] \quad (702)$$

for a point source of mass M located at the origin and for intermediate distances $l_p \ll r \ll \xi$. One can visualize the above result by stating that virtual graviton loops cause an effective anti-screening of the primary gravitational source M , giving rise to a quantum correction to the potential proportional to $r^{1/\nu-1}$. But only for $\nu = 1/3$ can the additional contribution be interpreted as being due to a close to uniform mass distribution surrounding the original source, of strength

$$\rho_\xi(M) = \frac{3cM}{4\pi\xi^3} \quad (703)$$

Such a simple geometric interpretation fails unless the exponent ν for gravitation is exactly one third. In fact in dimensions $d \geq 4$ one would expect based on the geometric argument that $\nu =$

$1/(d-1)$ if the quantum correction to the gravitational potential arises from such a virtual graviton cloud. These arguments rely of course on the lowest order result $V(r) \sim \int d^{d-1}k e^{ik \cdot x} / k^2 \sim r^{3-d}$ for single graviton exchange in $d > 3$ dimensions.

Equivalently, the running of G can be characterized as being in part due to a tiny non-vanishing (and positive) non-perturbative gravitational vacuum contribution to the cosmological constant, with

$$\lambda_0(M) = \frac{3cM}{\xi^3} \quad (704)$$

and therefore an associated effective classical average curvature of magnitude $R_{class} \sim G\lambda_0 \sim GM/\xi^3$. It is amusing that for a very large mass distribution M , the above expression for the curvature can only be reconciled with the naive dimensional estimate $R_{class} \sim 1/\xi^2$, provided for the gravitational coupling G itself one has $G \sim \xi/M$, which is reminiscent of Mach's principle and its connection with the Lense-Thirring effect (Lense and Thirring, 1918; Sciama, 1953; Feynman, 1962).

V. SCALE DEPENDENT GRAVITATIONAL COUPLING

Non-perturbative studies of quantum gravity suggest the possibility that gravitational couplings might be weakly scale dependent due to nontrivial renormalization group effects. This would introduce a new gravitational scale, unrelated to Newton's constant, required in order to parametrize the gravitational running in the infrared region. If one is willing to accept such a scenario, then it seems difficult to find a compelling theoretical argument for why the non-perturbative scale entering the coupling evolution equations should be very small, comparable to the Planck length. One possibility is that the relevant non-perturbative scale is related to the curvature and therefore macroscopic in size, which could have observable consequences. One key ingredient in this argument is the relationship, in part supported by Euclidean lattice results combined with renormalization group arguments, between the scaling violation parameter and the scale of the average curvature.

A. Effective Field Equations

1. Scale Dependence of G

To summarize the results of the previous section, the result of Eq. (693) implies for the running gravitational coupling in the vicinity of the ultraviolet fixed point

$$G(k^2) = G_c \left[1 + a_0 \left(\frac{m^2}{k^2} \right)^{\frac{1}{2\nu}} + O\left((m^2/k^2)^{\frac{1}{\nu}} \right) \right] \quad (705)$$

with $m = 1/\xi$, $a_0 > 0$ and $\nu \simeq 1/3$. Since ξ is expected to be very large, the quantity G_c in the above expression should now be identified with the laboratory scale value $\sqrt{G_c} \sim 1.6 \times 10^{-33} \text{cm}$. Quantum corrections on the r.h.s are therefore quite small as long as $k^2 \gg m^2$, which in real space corresponds to the “short distance” regime $r \ll \xi$.

The interaction in real space is often obtained by Fourier transform, and the above expression is singular as $k^2 \rightarrow 0$. The infrared divergence needs to be regulated, which can be achieved by utilizing as the lower limit of momentum integration $m = 1/\xi$. Alternatively, as a properly infrared regulated version of the above expression one can use

$$G(k^2) \simeq G_c \left[1 + a_0 \left(\frac{m^2}{k^2 + m^2} \right)^{\frac{1}{2\nu}} + \dots \right] \quad (706)$$

The last form for $G(k^2)$ will only be necessary in the regime where k is small, so that one can avoid unphysical results. From Eq. (706) the gravitational coupling then approaches at very large distances $r \gg \xi$ the finite value $G_\infty = (1 + a_0 + \dots)G_c$. Note though that in Eqs. (705) or (706) the cutoff no longer appear explicitly, it is absorbed into the definition of G_c . In the following we will be mostly interested in the regime $l_P \ll r \ll \xi$, for which Eq. (705) is completely adequate.

The first step in analyzing the consequences of a running of G is to re-write the expression for $G(k^2)$ in a coordinate-independent way. The following methods are not new, and have found over the years their fruitful application in gauge theories and gravity, for example in the discussion of non-local effective actions (Vilkovisky, 1984; Barvinsky and Vilkovisky, 1985). Since in going from momentum to position space one usually employs $k^2 \rightarrow -\square$, to obtain a quantum-mechanical running of the gravitational coupling one should make the replacement

$$G \rightarrow G(\square) \quad (707)$$

and therefore from Eq. (705)

$$G(\square) = G_c \left[1 + a_0 \left(\frac{1}{\xi^2 \square} \right)^{\frac{1}{2\nu}} + \dots \right]. \quad (708)$$

In general the form of the covariant d'Alembertian operator \square depends on the specific tensor nature of the object it is acting on,

$$\square T^{\alpha\beta\dots}_{\gamma\delta\dots} = g^{\mu\nu} \nabla_\mu \left(\nabla_\nu T^{\alpha\beta\dots}_{\gamma\delta\dots} \right) \quad (709)$$

Only on scalar functions one has the fairly simple result

$$\square S(x) = \frac{1}{\sqrt{g}} \partial_\mu g^{\mu\nu} \sqrt{g} \partial_\nu S(x) \quad (710)$$

whereas on second rank tensors one has the already significantly more complicated expression $\square T_{\alpha\beta} \equiv g^{\mu\nu} \nabla_\mu (\nabla_\nu T_{\alpha\beta})$.

The running of G is expected to lead to a non-local gravitational action, for example of the form

$$I = \frac{1}{16\pi G} \int dx \sqrt{g} \left(1 - a_0 \left(\frac{1}{\xi^2 \square} \right)^{1/2\nu} + \dots \right) R \quad (711)$$

Due to the fractional exponent in general the covariant operator appearing in the above expression, namely

$$A(\square) = a_0 \left(\frac{1}{\xi^2 \square} \right)^{1/2\nu} \quad (712)$$

has to be suitably defined by analytic continuation from positive integer powers. The latter can be done for example by computing \square^n for positive integer n and then analytically continuing to $n \rightarrow -1/2\nu$. Alternatively one can make use of the identity

$$\frac{1}{\square^n} = \frac{(-1)^n}{\Gamma(n)} \int_0^\infty ds s^{n-1} \exp(i s \square) \quad (713)$$

and later perform the relevant integrals with $n \rightarrow 1/2\nu$. Other procedures can be used to define $A(\square)$, for example based on an integral representation involving the scalar propagator (Lopez Nacir and Mazzitelli, 2007).

It should be stressed here that the action in Eq. (711) should be treated as a *classical* effective action, with dominant radiative corrections at short distances ($r \ll \xi$) already automatically built in, and for which a restriction to generally smooth field configurations does make some sense. In particular one would expect that in most instances it should be possible, as well as meaningful, to neglect terms involving large numbers of derivatives of the metric in order to compute the effects of the new contributions appearing in the effective action.

Had one *not* considered the action of Eq. (711) as a starting point for constructing the effective theory, one would naturally be led (following Eq. (707)) to consider the following effective field

equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \lambda g_{\mu\nu} = 8\pi G (1 + A(\square)) T_{\mu\nu} \quad (714)$$

the argument again being the replacement $G \rightarrow G(\square) \equiv G(1 + A(\square))$. Being manifestly covariant, these expressions at least satisfy some of the requirements for a set of consistent field equations incorporating the running of G . The above effective field equation can in fact be re-cast in a form similar to the classical field equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \lambda g_{\mu\nu} = 8\pi G \tilde{T}_{\mu\nu} \quad (715)$$

with $\tilde{T}_{\mu\nu} = (1 + A(\square)) T_{\mu\nu}$ defined as an effective, or gravitationally dressed, energy-momentum tensor. Just like the ordinary Einstein gravity case, in general $\tilde{T}_{\mu\nu}$ might not be covariantly conserved a priori, $\nabla^\mu \tilde{T}_{\mu\nu} \neq 0$, but ultimately the consistency of the effective field equations demands that it be exactly conserved, in consideration of the Bianchi identity satisfied by the Riemann tensor (a similar problem arises in other non-local modifications of gravity (Barvinsky, 2003)). The ensuing new covariant conservation law

$$\nabla^\mu \tilde{T}_{\mu\nu} \equiv \nabla^\mu [(1 + A(\square)) T_{\mu\nu}] = 0 \quad (716)$$

can be then be viewed as a *constraint* on $\tilde{T}_{\mu\nu}$ (or $T_{\mu\nu}$) which, for example, in the specific case of a perfect fluid, will imply again a definite relationship between the density $\rho(t)$, the pressure $p(t)$ and the RW scale factor $a(t)$, just as it does in the standard case. Then the requirement that the bare energy momentum-tensor be conserved would imply that the quantum contribution $A(\square) T_{\mu\nu}$ itself be separately conserved. That this is indeed attainable can be shown in a few simple cases, such as the static isotropic solution discussed below. There a "vacuum fluid" is introduced to account for the vacuum polarization contribution, whose energy momentum tensor can be shown to be covariantly conserved. That the procedure is consistent in general is not clear, in which case the present approach should perhaps be limited to phenomenological considerations.

Let us make a few additional comments regarding the above effective field equations, in which we will set the cosmological constant $\lambda = 0$ from now on. One simple observation is that the trace equation only involves the (simpler) scalar d'Alembertian, acting on the trace of the energy-momentum tensor

$$R = -8\pi G (1 + A(\square)) T_\mu{}^\mu \quad (717)$$

Finally, to the order one is working here, the above effective field equations should be equivalent to

$$\left(1 - A(\square) + O(A(\square)^2)\right) \left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R\right) = 8\pi G T_{\mu\nu} \quad (718)$$

where the running of G has been moved over to the “gravitational” side.

2. Poisson's Equation and Vacuum Polarization Cloud

One of the simplest cases to analyze is of course the static case. The non-relativistic, static Newtonian potential is defined as usual as

$$\phi(r) = (-M) \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} G(\mathbf{k}^2) \frac{4\pi}{\mathbf{k}^2} \quad (719)$$

and therefore proportional to the 3 – d Fourier transform of

$$\frac{4\pi}{\mathbf{k}^2} \rightarrow \frac{4\pi}{\mathbf{k}^2} \left[1 + a_0 \left(\frac{m^2}{\mathbf{k}^2} \right)^{\frac{1}{2\nu}} + \dots \right] \quad (720)$$

But, as already mentioned, for small \mathbf{k} proper care has to be exercised in providing a properly infrared regulated version of the above expression, which, from Eq. (706), reads

$$\frac{4\pi}{(\mathbf{k}^2 + \mu^2)} \rightarrow \frac{4\pi}{(\mathbf{k}^2 + \mu^2)} \left[1 + a_0 \left(\frac{m^2}{\mathbf{k}^2 + m^2} \right)^{\frac{1}{2\nu}} + \dots \right] \quad (721)$$

where the limit $\mu \rightarrow 0$ should be taken at the end of the calculation.

Given the running of G from either Eq. (706), or Eq. (705) in the large \mathbf{k} limit, the next step is naturally an attempt at finding a solution to Poisson's equation with a point source at the origin, so that one can determine the structure of the quantum corrections to the static gravitational potential in real space. There are in principle two equivalent ways to compute the potential $\phi(r)$, either by inverse Fourier transform of Eq. (720), or by solving Poisson's equation $\Delta\phi = 4\pi\rho$ with the source term $\rho(r)$ given by the inverse Fourier transform of the correction to $G(k^2)$, as given below in Eq. (724). The zero-th order term then gives the standard Newtonian $-MG/r$ term, while the correction in general is given by a rather complicated hypergeometric function. But for the special case $\nu = 1/2$ the Fourier transform of Eq. (721) is easy to do, the integrals are elementary and the running of $G(r)$ so obtained is particularly transparent,

$$G(r) = G_\infty \left(1 - \frac{a_0}{1 + a_0} e^{-mr} \right) \quad (722)$$

where we have set $G_\infty \equiv (1 + a_0)G$ and $G \equiv G(0)$. G therefore increases slowly from its value G at small r to the larger value $(1 + a_0)G$ at infinity. Fig. 41 illustrates the anti-screening effect of the virtual graviton cloud. Fig. 42 gives a schematic illustration of the behavior of G as a function of r .

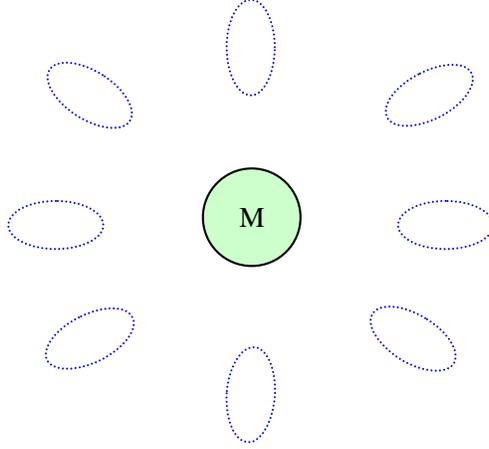


FIG. 41 A virtual graviton cloud surrounds the point source of mass M , leading to an anti-screening modification of the static gravitational potential.

Another possible procedure to obtain the static potential $\phi(r)$ is to solve directly the radial Poisson equation for $\phi(r)$. This will give a density $\rho(r)$ which can later be used to generalize to the relativistic case. In the $a_0 \neq 0$ case one needs to solve $\Delta\phi = 4\pi\rho$, or in the radial coordinate for $r > 0$

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\phi}{dr} \right) = 4\pi G \rho_m(r) \quad (723)$$

with the source term ρ_m determined from the inverse Fourier transform of the correction term in Eq. (721). The latter is given by

$$\rho_m(r) = \frac{1}{8\pi} c_\nu a_0 M m^3 (mr)^{-\frac{1}{2}(3-\frac{1}{\nu})} K_{\frac{1}{2}(3-\frac{1}{\nu})}(mr) \quad (724)$$

with the constant

$$c_\nu \equiv \frac{2^{\frac{1}{2}(5-\frac{1}{\nu})}}{\sqrt{\pi} \Gamma(\frac{1}{2\nu})} . \quad (725)$$

One can verify that the vacuum polarization density ρ_m has the property

$$4\pi \int_0^\infty r^2 dr \rho_m(r) = a_0 M \quad (726)$$

where the standard integral $\int_0^\infty dx x^{2-n} K_n(x) = 2^{-n} \sqrt{\pi} \Gamma\left(\frac{3}{2} - n\right)$ has been used. Note that the gravitational vacuum polarization distribution is singular close to $r = 0$, just as in QED, Eq. (321).

The $r \rightarrow 0$ result for $\phi(r)$ (discussed in the following, as an example, for $\nu = 1/3$) can then be obtained by solving the radial equation for $\phi(r)$,

$$\frac{1}{r} \frac{d^2}{dr^2} [r \phi(r)] = \frac{2 a_0 M G m^3}{\pi} K_0(m r) \quad (727)$$

where the (modified) Bessel function is expanded out to lowest order in r , $K_0(m r) = -\gamma - \ln\left(\frac{m r}{2}\right) + O(m^2 r^2)$, giving

$$\phi(r) = -\frac{M G}{r} + a_0 M G m^3 \frac{r^2}{3 \pi} \left[-\ln\left(\frac{m r}{2}\right) - \gamma + \frac{5}{6} \right] + O(r^3) \quad (728)$$

where the two integration constants are matched to the general large r solution

$$\phi(r) \underset{r \rightarrow \infty}{\sim} -\frac{M G}{r} \left[1 + a_0 \left(1 - c_l(m r)^{\frac{1}{2\nu}-1} e^{-m r} \right) \right] \quad (729)$$

with $c_l = 1/(\nu 2^{\frac{1}{2\nu}} \Gamma(\frac{1}{2\nu}))$. Note again that the vacuum polarization density $\rho_m(r)$ has the expected normalization property

$$4 \pi \int_0^\infty r^2 dr \frac{a_0 M m^3}{2 \pi^2} K_0(m r) = \frac{2 a_0 M m^3}{\pi} \cdot \frac{\pi}{2 m^3} = a_0 M \quad (730)$$

so that the total enclosed additional gravitational charge is indeed just $a_0 M$, and $G_\infty = G_0(1 + a_0)$.

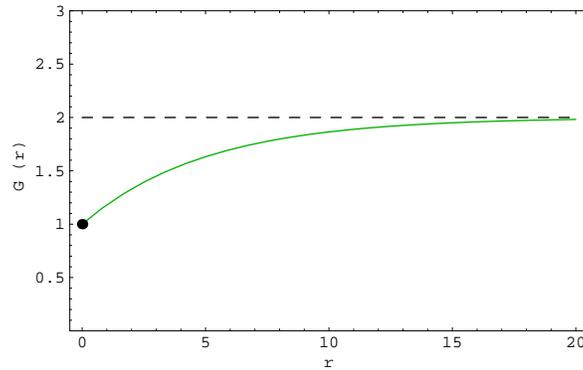


FIG. 42 Schematic scale dependence of the gravitational coupling $G(r)$, from Eq. (722) valid for $\nu = 1/2$. The gravitational coupling rises initially like a power of r , and later approaches the asymptotic value $G_\infty = (1 + a_0)G$ for large r . The behavior for other values of $\nu > 1/3$ is similar.

3. Static Isotropic Solution

The discussion of the previous section suggests that the quantum correction due to the running of G can be described, at least in the non-relativistic limit of Eq. (706) as applied to Poisson's

equation, in terms of a vacuum energy density $\rho_m(r)$, distributed around the static source of strength M in accordance with the result of Eqs. (724) and Eq. (726).

In general a manifestly covariant implementation of the running of G , via the $G(\square)$ given in Eq. (708), will induce a non-vanishing effective pressure term. It is natural therefore to attempt to represent the vacuum polarization cloud by a relativistic perfect fluid, with energy-momentum tensor

$$T_{\mu\nu} = (p + \rho) u_\mu u_\nu + g_{\mu\nu} p \quad (731)$$

which in the static isotropic case reduces to

$$T_{\mu\nu} = \text{diag} [B(r) \rho(r), A(r) p(r), r^2 p(r), r^2 \sin^2 \theta p(r)] \quad (732)$$

and gives a trace $T = 3p - \rho$. The tt , rr and $\theta\theta$ components of the field equations then read

$$-\lambda B(r) + \frac{A'(r)B(r)}{rA(r)^2} - \frac{B(r)}{r^2 A(r)} + \frac{B(r)}{r^2} = 8\pi G B(r) \rho(r) \quad (733)$$

$$\lambda A(r) - \frac{A(r)}{r^2} + \frac{B'(r)}{rB(r)} + \frac{1}{r^2} = 8\pi G A(r) p(r) \quad (734)$$

$$-\frac{B'(r)^2 r^2}{4A(r)B(r)^2} + \lambda r^2 - \frac{A'(r)B'(r)r^2}{4A(r)^2 B(r)} + \frac{B''(r)r^2}{2A(r)B(r)} - \frac{A'(r)r}{2A(r)^2} + \frac{B'(r)r}{2A(r)B(r)} = 8G\pi r^2 p(r) \quad (735)$$

with the $\varphi\varphi$ component equal to $\sin^2 \theta$ times the $\theta\theta$ component. Covariant energy conservation $\nabla^\mu T_{\mu\nu} = 0$ implies

$$[p(r) + \rho(r)] \frac{B'(r)}{2B(r)} + p'(r) = 0 \quad (736)$$

and forces a definite relationship between $B(r)$, $\rho(r)$ and $p(r)$. The three field equations and the energy conservation equation are, as usual, not independent, because of the Bianchi identity. It seems reasonable to attempt to solve the above equations (usually considered in the context of relativistic stellar structure (Misner Thorne Wheeler 1972)) with the density $\rho(r)$ given by the $\rho_m(r)$ of Eq. (724). This of course raises the question of how the relativistic pressure $p(r)$ should be chosen, an issue that the non-relativistic calculation did not have to address. One finds that covariant energy conservation in fact completely determines the pressure in the static case, leading to consistent equations and solutions (note that in particular it would not be consistent to take $p(r) = 0$).

Since the function $B(r)$ drops out of the tt field equation, the latter can be integrated immediately, giving

$$A(r)^{-1} = 1 - \frac{2MG}{r} - \frac{\lambda}{3}r^2 - \frac{8\pi G}{r} \int_0^r dx x^2 \rho(x) \quad (737)$$

It is natural to identify $c_1 = -2MG$, which of course corresponds to the solution with $a_0 = 0$ ($p = \rho = 0$). Next, the rr field equation can be solved for $B(r)$,

$$B(r) = \exp \left\{ c_2 - \int_{r_0}^r dy \frac{1 + A(y) (\lambda y^2 - 8\pi G y^2 p(y) - 1)}{y} \right\} \quad (738)$$

with the constant c_2 again determined by the requirement that the above expression for $B(r)$ reduce to the standard Schwarzschild solution for $a_0 = 0$ ($p = \rho = 0$), giving $c_2 = \ln(1 - 2MG/r_0 - \lambda r_0^2/3)$.

The last task left therefore is the determination of the pressure $p(r)$. One needs to solve the equation

$$p'(r) + \frac{\left(8\pi G r^3 p(r) + 2MG - \frac{2}{3}\lambda r^3 + 8\pi G \int_{r_0}^r dx x^2 \rho(x) \right) (p(r) + \rho(r))}{2r \left(r - 2MG - \frac{\lambda}{3}r^3 - 8\pi G \int_0^r dx x^2 \rho(x) \right)} = 0 \quad (739)$$

which is usually referred to as the equation of hydrostatic equilibrium. From now on we will focus only the case $\lambda = 0$. The last differential equation can be solved for $p(r)$,

$$p_m(r) = \frac{1}{\sqrt{1 - \frac{2MG}{r}}} \left(c_3 - \int_{r_0}^r dz \frac{MG\rho(z)}{z^2 \sqrt{1 - \frac{2MG}{z}}} \right) \quad (740)$$

where the constant of integration has to be chosen so that when $\rho(r) = 0$ (no quantum correction) one has $p(r) = 0$ as well. Because of the singularity in the integrand at $r = 2MG$, we will take the lower limit in the integral to be $r_0 = 2MG + \epsilon$, with $\epsilon \rightarrow 0$.

To proceed further, one needs the explicit form for $\rho_m(r)$, which was given in Eq. (724),

$$\rho_m(r) = \frac{1}{8\pi} c_\nu a_0 M m^3 (mr)^{-\frac{1}{2}(3-\frac{1}{\nu})} K_{\frac{1}{2}(3-\frac{1}{\nu})}(mr) \quad (741)$$

The required integrands involve for general ν the modified Bessel function $K_n(x)$, which can lead to rather complicated expressions for the general ν case. To determine the pressure, one supposes that it as well has a power dependence on r in the regime under consideration, $p_m(r) = c_p A_0 r^\gamma$, where c_p is a numerical constant, and then substitute $p_m(r)$ into the pressure equation Eq. (739). This gives, past the horizon $r \gg 2MG$ the algebraic condition

$$(2\gamma - 1) c_p M G r^{\gamma-1} - c_p \gamma r^\gamma - M G r^{1/\nu-4} \simeq 0 \quad (742)$$

giving the same power $\gamma = 1/\nu - 3$ as for $\rho(r)$, $c_p = -1$ and surprisingly also $\gamma = 0$, implying that in this regime only $\nu = 1/3$ gives a consistent solution.

The case $\nu = 1/3$ can be dealt with separately, starting from the expression for $\rho_m(r)$ for $\nu = 1/3$

$$\rho_m(r) = \frac{1}{2\pi^2} a_0 M m^3 K_0(m r) \quad (743)$$

One has for small r

$$\rho_m(r) = -\frac{a_0}{2\pi^2} M m^3 \left(\ln \frac{m r}{2} + \gamma \right) + \dots \quad (744)$$

and consequently

$$A^{-1}(r) = 1 - \frac{2 M G}{r} + \frac{4 a_0 M G m^3}{3 \pi} r^2 \ln(m r) + \dots \quad (745)$$

From Eq. (739) one can then obtain an expression for the pressure $p_m(r)$, and one finds again in the limit $r \gg 2MG$

$$p_m(r) = \frac{a_0}{2\pi^2} M m^3 \ln(m r) + \dots \quad (746)$$

After performing the required r integral in Eq. (738), and evaluating the resulting expression in the limit $r \gg 2MG$, one obtains

$$B(r) = 1 - \frac{2 M G}{r} + \frac{4 a_0 M G m^3}{3 \pi} r^2 \ln(m r) + \dots \quad (747)$$

It is encouraging to note that in the solution just obtained the running of G is the *same* in $A(r)$ and $B(r)$. The expressions for $A(r)$ and $B(r)$ are consistent with a gradual slow increase in G with distance, in accordance with the formula

$$G \rightarrow G(r) = G \left(1 + \frac{a_0}{3 \pi} m^3 r^3 \ln \frac{1}{m^2 r^2} + \dots \right) \quad (748)$$

in the regime $r \gg 2 M G$, and therefore of course in agreement with the original result of Eqs. (705) or (706), namely that the classical laboratory value of G is obtained for $r \ll \xi$. Note that the correct relativistic small r correction of Eq. (748) agrees roughly in magnitude (but not in sign) with the approximate non-relativistic, Poisson equation result of Eq. (728).

There are similarities, as well as some rather substantial differences, with the corresponding QED result of Eq. (321). In the gravity case, the correction vanishes as r goes to zero: in this limit one is probing the bare mass, unobstructed by its virtual graviton cloud. On the other hand, in the QED case, as one approaches the source one is probing the bare charge, whose magnitude diverges logarithmically for small r .

Finally it should be recalled that neither function $A(r)$ or $B(r)$ are directly related to the relativistic potential for particle orbits, which is given instead by the combination

$$V_{eff}(r) = \frac{1}{2A(r)} \left[\frac{l^2}{r^2} - \frac{1}{B(r)} + 1 \right] \quad (749)$$

where l is proportional to the orbital angular momentum of the test particle, as discussed for example in (Hartle 2005).

The running G term acts in a number of ways as a local cosmological constant term, for which the r dependence of the vacuum solution for small r is fixed by the nature of the Schwarzschild solution with a cosmological constant term. One can therefore wonder what the solutions might look like in d dimensions. In $d \geq 4$ dimensions the Schwarzschild solution to Einstein gravity with a cosmological term is (Myers and Perry 1986)

$$A^{-1}(r) = B(r) = 1 - 2MGc_d r^{3-d} - \frac{2\lambda}{(d-2)(d-1)} r^2 \quad (750)$$

with $c_d = 4\pi\Gamma(\frac{d-1}{2})/(d-2)\pi^{\frac{d-1}{2}}$, which would suggest, in analogy with the results for $d = 4$ given above that in $d \geq 4$ dimensions only $\nu = 1/(d-1)$ is possible. This last result would also be in agreement with the exact value $\nu = 0$ found at $d = \infty$ in Sec. III.B.3.

4. Cosmological Solutions

A scale dependent Newton's constant will lead to small modifications of the standard cosmological solutions to the Einstein field equations. Here we will provide a brief discussion of what modifications are expected from the effective field equations on the basis of $G(\square)$, as given in Eq. (707), which itself originates in Eqs. (706) and (705).

One starts therefore from the quantum effective field equations of Eq. (714),

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \lambda g_{\mu\nu} = 8\pi G(1 + A(\square))T_{\mu\nu} \quad (751)$$

with $A(\square)$ defined in Eq. (712). In the Friedmann-Robertson-Walker (FRW) framework these are applied to the standard homogenous isotropic metric

$$ds^2 = -dt^2 + a^2(t) \left\{ \frac{dr^2}{1 - k r^2} + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2) \right\} \quad (752)$$

It should be noted that there are *two* quantum contributions to the above set of effective field equations. The first one arises because of the presence of a non-vanishing cosmological constant $\lambda \simeq 1/\xi^2$ caused by the non-perturbative vacuum condensate of Eq. (695). As in the case of

standard FRW cosmology, this is expected to be the dominant contributions at large times t , and gives an exponential (for $\lambda > 0$ or cyclic (for $\lambda < 0$) expansion of the scale factor.

The second contribution arises because of the running of G in the effective field equations,

$$G(\square) = G(1 + A(\square)) = G \left[1 + a_0 \left(\xi^2 \square \right)^{-\frac{1}{2\nu}} + \dots \right] \quad (753)$$

for for $t \ll \xi$, with $\nu \simeq 1/3$ and $a_0 > 0$ a calculable coefficient of order one [see Eqs. (705) and (706)].

As a first step in solving the new set of effective field equations, consider first the *trace* of the field equation in Eq. (751), written as

$$\left(1 - A(\square) + O(A(\square)^2) \right) R = 8\pi G T_\mu{}^\mu \quad (754)$$

where R is the scalar curvature. Here the operator $A(\square)$ has been moved over on the gravitational side, so that it now acts on functions of the metric only, using the binomial expansion of $1/(1 + A(\square))$. To proceed further, one needs to compute the effect of $A(\square)$ on the scalar curvature. The d'Alembertian operator acting on scalar functions $S(x)$ is given by

$$\frac{1}{\sqrt{g}} \partial_\mu g^{\mu\nu} \sqrt{g} \partial_\nu S(x) \quad (755)$$

and for the Robertson-Walker metric, acting on functions of t only, one obtains a fairly simple result in terms of the scale factor $a(t)$

$$- \frac{1}{a^3(t)} \frac{\partial}{\partial t} \left[a^3(t) \frac{\partial}{\partial t} \right] F(t) \quad (756)$$

As a next step one computes the action of \square on the scalar curvature R , which gives

$$- 6 \left[-2k \ddot{a}(t) - 5 \dot{a}^2(t) \ddot{a}(t) + a(t) \ddot{a}^2(t) + 3 a(t) \dot{a}(t) a^{(3)}(t) + a^2(t) a^{(4)}(t) \right] / a^3(t) \quad (757)$$

and then \square^2 on R etc. Since the resulting expressions are of rapidly escalating complexity, one sets $a(t) = r_0 t^\alpha$, in which case one has first for the scalar curvature itself

$$R = 6 \left[\frac{k}{r_0^2 t^{2\alpha}} + \frac{\alpha(-1 + 2\alpha)}{t^2} \right] \quad (758)$$

Acting with \square^n on the above expression gives for $k = 0$ and arbitrary power n

$$c_n 6 \alpha (-1 + 2\alpha) t^{-2-2n} \quad (759)$$

with the coefficient c_n given by

$$c_n = 4^n \frac{\Gamma(n+1) \Gamma(\frac{3\alpha-1}{2})}{\Gamma(\frac{3\alpha-1}{2} - n)} \quad (760)$$

Here use has been made of the relationship

$$\left(\frac{d}{dz}\right)^\alpha (z - c)^\beta = \frac{\Gamma(\beta + 1)}{\Gamma(\beta - \alpha + 1)} (z - c)^{\beta - \alpha} \quad (761)$$

to analytically continue the above expressions to negative fractional n (Samko et al 1993; Zavada 1998). For $n = -1/2\nu$ the correction on the scalar curvature term R is therefore of the form

$$\left[1 - a_0 c_\nu (t/\xi)^{1/\nu}\right] \cdot 6\alpha (-1 + 2\alpha) t^{-2} \quad (762)$$

with

$$c_\nu = 2^{-\frac{1}{\nu}} \frac{\Gamma(1 - \frac{1}{2\nu})\Gamma(\frac{3\alpha-1}{2})}{\Gamma(\frac{3\alpha-1}{2} + \frac{1}{2\nu})} \quad (763)$$

Putting everything together, one then obtains for the trace part of the effective field equations

$$\left[1 - a_0 c_\nu \left(\frac{t}{\xi}\right)^{1/\nu} + O\left((t/\xi)^{2/\nu}\right)\right] \frac{6\alpha (2\alpha - 1)}{t^2} = 8\pi G \rho(t) \quad (764)$$

The new term can now be moved back over to the matter side in accordance with the structure of the original effective field equation of Eq. (751), and thus avoids the problem of having to deal with the binomial expansion of $1/(1 + A(\square))$. One then has

$$\frac{6\alpha (2\alpha - 1)}{t^2} = 8\pi G \left[1 + a_0 c_\nu \left(\frac{t}{\xi}\right)^{1/\nu} + O\left((t/\xi)^{2/\nu}\right)\right] \rho(t) \quad (765)$$

which is the Robertson-Walker metric form of Eq. (751). If one assumes for the matter density $\rho(t) \sim \rho_0 t^\beta$, then matching powers when the new term starts to take over at larger distances gives the first result

$$\beta = -2 - 1/\nu \quad (766)$$

Thus the density decreases *faster* in time than the classical value ($\beta = -2$) would indicate. The expansion appears therefore to be accelerating, but before reaching such a conclusion one needs to determine the time dependence of the scale factor $a(t)$ (or α) as well.

One can alternatively pursue the following exercise in order to check the overall consistency of the approach. Consider \square^n acting on $T_\mu{}^\mu = -\rho(t)$ instead, as in the trace of the effective field equation Eq. (751)

$$R = -8\pi G (1 + A(\square)) T_\mu{}^\mu \quad (767)$$

for $\lambda = 0$ and $p(t) = 0$. For $\rho(t) = \rho_0 t^\beta$ and $a(t) = r_0 t^\alpha$ one finds in this case

$$\square^n (-\rho(t)) \rightarrow 4^n (-1)^{n+1} \frac{\Gamma(\frac{\beta}{2} + 1)\Gamma(\frac{\beta+3\alpha+1}{2})}{\Gamma(\frac{\beta}{2} + 1 - n)\Gamma(\frac{\beta+3\alpha+1}{2} - n)} \rho_0 t^{\beta-2n} \quad (768)$$

which again implies for $n \rightarrow -1/2\nu$ the value $\beta = -2 - 1/\nu$ as in Eq. (766) for large(r) times, when the quantum correction starts to become important (since the left hand side of Einstein's equation always goes like $1/t^2$, no matter what the value for α is, at least for $k=0$).

The next step is to examine the full effective field equations (as opposed to just their trace part) of Eq. (717) with cosmological constant $\lambda = 0$,

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G (1 + A(\square)) T_{\mu\nu} \quad (769)$$

Here the d'Alembertian operator

$$\square = g^{\mu\nu} \nabla_\mu \nabla_\nu \quad (770)$$

acts on a second rank tensor,

$$\nabla_\nu T_{\alpha\beta} = \partial_\nu T_{\alpha\beta} - \Gamma_{\alpha\nu}^\lambda T_{\lambda\beta} - \Gamma_{\beta\nu}^\lambda T_{\alpha\lambda} \equiv I_{\nu\alpha\beta}$$

$$\nabla_\mu (\nabla_\nu T_{\alpha\beta}) = \partial_\mu I_{\nu\alpha\beta} - \Gamma_{\nu\mu}^\lambda I_{\lambda\alpha\beta} - \Gamma_{\alpha\mu}^\lambda I_{\nu\lambda\beta} - \Gamma_{\beta\mu}^\lambda I_{\nu\alpha\lambda} \quad (771)$$

and would thus seem to require the calculation of 1920 terms, of which fortunately many vanish by symmetry. Next one assumes again that $T_{\mu\nu}$ has the perfect fluid form, for which one obtains from the action of \square on $T_{\mu\nu}$

$$\begin{aligned} (\square T_{\mu\nu})_{tt} &= 6 [\rho(t) + p(t)] \left(\frac{\dot{a}(t)}{a(t)} \right)^2 - 3 \dot{\rho}(t) \frac{\dot{a}(t)}{a(t)} - \ddot{\rho}(t) \\ (\square T_{\mu\nu})_{rr} &= \frac{1}{1 - k r^2} \left\{ 2 [\rho(t) + p(t)] \dot{a}(t)^2 - 3 \dot{p}(t) a(t) \dot{a}(t) - \ddot{p}(t) a(t)^2 \right\} \\ (\square T_{\mu\nu})_{\theta\theta} &= r^2 (1 - k r^2) (\square T_{\mu\nu})_{rr} \\ (\square T_{\mu\nu})_{\varphi\varphi} &= r^2 (1 - k r^2) \sin^2 \theta (\square T_{\mu\nu})_{rr} \end{aligned} \quad (772)$$

with the remaining components equal to zero. Note that a non-vanishing pressure contribution is generated in the effective field equations, even if one assumes initially a pressureless fluid, $p(t) = 0$. As before, repeated applications of the d'Alembertian \square to the above expressions leads to rapidly escalating complexity, which can only be tamed by introducing some further simplifying assumptions. In the following we will therefore assume that $T_{\mu\nu}$ has the perfect fluid form appropriate for non-relativistic matter, with a power law behavior for the density, $\rho(t) = \rho_0 t^\beta$, and $p(t) = 0$. Thus all components of $T_{\mu\nu}$ vanish in the fluid's rest frame, except the tt one, which is simply $\rho(t)$. Setting $k = 0$ and $a(t) = r_0 t^\alpha$ one then finds

$$\begin{aligned} (\square T_{\mu\nu})_{tt} &= (6\alpha^2 - \beta^2 - 3\alpha\beta + \beta) \rho_0 t^{\beta-2} \\ (\square T_{\mu\nu})_{rr} &= 2r_0^2 t^{2\alpha} \alpha^2 \rho_0 t^{\beta-2} \end{aligned} \quad (773)$$

which again shows that the tt and rr components get mixed by the action of the \square operator, and that a non-vanishing rr component gets generated, even though it was not originally present.

Higher powers of the d'Alembertian \square acting on $T_{\mu\nu}$ can then be computed as well. But a comparison with the left hand (gravitational) side of the effective field equation, which always behaves like $\sim 1/t^2$ for $k = 0$, shows that in fact a solution can only be achieved at order \square^n provided the exponent β satisfies $\beta = -2 + 2n$, or

$$\beta = -2 - 1/\nu \quad (774)$$

as was found previously from the trace equation, Eqs. (751) and (766). As a result one obtains a much simpler set of expressions, which for general n read

$$(\square^n T_{\mu\nu})_{tt} \rightarrow c_{tt}(\alpha, \nu) \rho_0 t^{-2} \quad (775)$$

for the tt component, and similarly for the rr component

$$(\square^n T_{\mu\nu})_{rr} \rightarrow c_{rr}(\alpha, \nu) r_0^2 t^{2\alpha} \rho_0 t^{-2} \quad (776)$$

But remarkably one finds for the two coefficients the simple identity

$$c_{rr}(\alpha, \nu) = \frac{1}{3} c_{tt}(\alpha, \nu) \quad (777)$$

as well as $c_{\theta\theta} = r^2 c_{rr}$ and $c_{\varphi\varphi} = r^2 \sin^2 \theta c_{rr}$. The identity $c_{rr} = \frac{1}{3} c_{tt}$ implies, from the consistency of the tt and rr effective field equations at large times,

$$\alpha = \frac{1}{2} \quad (778)$$

One can find a closed form expression for the coefficients c_{tt} and $c_{rr} = c_{tt}/3$ as functions of ν which are not particularly illuminating, except for providing an explicit proof that they exist.

As a result, in the simplest case, namely for a universe filled with non-relativistic matter ($p=0$), the effective Friedmann equations then have the following appearance

$$\begin{aligned} \frac{k}{a^2(t)} + \frac{\dot{a}^2(t)}{a^2(t)} &= \frac{8\pi G(t)}{3} \rho(t) + \frac{1}{3\xi^2} \\ &= \frac{8\pi G}{3} \left[1 + c_\xi (t/\xi)^{1/\nu} + \dots \right] \rho(t) + \frac{1}{3} \lambda \end{aligned} \quad (779)$$

for the tt field equation, and

$$\frac{k}{a^2(t)} + \frac{\dot{a}^2(t)}{a^2(t)} + \frac{2\ddot{a}(t)}{a(t)} = -\frac{8\pi G}{3} \left[c_\xi (t/\xi)^{1/\nu} + \dots \right] \rho(t) + \lambda \quad (780)$$

for the rr field equation. The running of G appropriate for the Robertson-Walker metric, and appearing explicitly in the first equation, is given by

$$G(t) = G \left[1 + c_\xi \left(\frac{t}{\xi} \right)^{1/\nu} + \dots \right] \quad (781)$$

with c_ξ of the same order as a_0 of Eq. (705). Note that the running of $G(t)$ induces as well an effective pressure term in the second (rr) equation.¹² One has therefore an effective density given by

$$\rho_{eff}(t) = \frac{G(t)}{G} \rho(t) \quad (782)$$

and an effective pressure

$$p_{eff}(t) = \frac{1}{3} \left(\frac{G(t)}{G} - 1 \right) \rho(t) \quad (783)$$

with $p_{eff}(t)/\rho_{eff}(t) = \frac{1}{3}(G(t) - G)/G(t)$. Strictly speaking, the above results can only be proven if one assumes that the pressure's time dependence is given by a power law. In the more general case, the solution of the above equations for various choices of ξ and a_0 has to be done numerically. Within the FRW framework, the gravitational vacuum polarization term behaves therefore in some ways (but not all) like a positive pressure term, with $p(t) = \omega \rho(t)$ and $\omega = 1/3$, which is therefore characteristic of radiation. One could therefore visualize the gravitational vacuum polarization contribution as behaving like ordinary radiation, in the form of a dilute virtual graviton gas: a radiative fluid with an equation of state $p = \frac{1}{3}\rho$. But this would overlook the fact that the relationship between density $\rho(t)$ and scale factor $a(t)$ is quite different from the classical case.

The running of $G(t)$ in the above equations follows directly from the basic result of Eq. (705), following the more or less unambiguously defined sequence $G(k^2) \rightarrow G(\square) \rightarrow G(t)$. For large times $t \gg \xi$ the form of Eq. (705), and therefore Eq. (781), is no longer appropriate, due to the spurious infrared divergence of Eq. (705) at small k^2 . Indeed from Eq. (706), the infrared regulated version of the above expression should read instead

$$G(t) \simeq G \left[1 + c_\xi \left(\frac{t^2}{t^2 + \xi^2} \right)^{\frac{1}{2\nu}} + \dots \right] \quad (784)$$

For very large times $t \gg \xi$ the gravitational coupling then approaches a constant, finite value $G_\infty = (1 + a_0 + \dots) G_c$. The modification of Eq. (784) should apply whenever one considers times

¹² We wish to emphasize that we are *not* talking here about models with a time-dependent value of G . Thus, for example, the value of $G \simeq G_c$ at laboratory scales should be taken to be constant throughout most of the evolution of the universe.

for which $t \ll \xi$ is not valid. But since $\xi \sim 1/\sqrt{\lambda}$ is of the order the size of the visible universe, the latter regime is largely of academic interest.

It should also be noted that the effective Friedmann equations of Eqs. (779) and (780) also bear a superficial degree of resemblance to what might be obtained in some scalar-tensor theories of gravity, where the gravitational Lagrangian is postulated to be some singular function of the scalar curvature (Capozziello et al 2003; Carroll et al 2004; Flanagan 2004). Indeed in the Friedmann-Robertson-Walker case one has, for the scalar curvature in terms of the scale factor,

$$R = 6 \left(k + \dot{a}^2(t) + a(t) \ddot{a}(t) \right) / a^2(t) \quad (785)$$

and for $k = 0$ and $a(t) \sim t^\alpha$ one has

$$R = \frac{6 \alpha (2 \alpha - 1)}{t^2} \quad (786)$$

which suggests that the quantum correction in Eq. (779) is, at this level, nearly indistinguishable from an inverse curvature term of the type $(\xi^2 R)^{-1/2\nu}$, or $1/(1 + \xi^2 R)^{1/2\nu}$ if one uses the infrared regulated version. The former would then correspond to an effective gravitational action

$$I_{eff} \simeq \frac{1}{16\pi G} \int dx \sqrt{g} \left(R + \frac{f \xi^{-\frac{1}{\nu}}}{|R|^{\frac{1}{2\nu}-1}} - 2\lambda \right) \quad (787)$$

with f a numerical constant of order one, and $\lambda \simeq 1/\xi^2$. But this superficial resemblance is seen here more as an artifact, due to the particularly simple form of the Robertson-Walker metric, with the coincidence of several curvature invariants not expected to be true in general. In particular in Eqs. (779) and (780) it would seem artificial and in fact inconsistent to take $\lambda \sim 1/\xi^2$ to zero while keeping the ξ in $G(t)$ finite.

VI. CONCLUSIONS AND OUTLOOK

While it is certainly possible that traditional quantum-field theoretic approaches to quantum gravitation might ultimately fail, it is not clear yet from the evidence so far that such a conclusion is warranted. There are certainly aspects of gravity that make it unique among theories of fundamental interactions, such as the geometric interpretation and its possible connection with the ultimate nature of space-time. These aspects might or might not play a fundamental role in establishing the theoretical consistency of the quantum theory. In a more traditional field-theoretic investigation of gravity the scope might have to be limited as well: the theory might not provide any new deep insights into fundamental questions (such as why the gravitational couplings take

on particular values), as is the case already in *QED*. Indeed the following discussion is not incompatible with the belief that present gravitational theories should eventually be replaced by more fundamental ones at distances comparable to the ultraviolet cutoff scale.

The theory of non-renormalizable interactions shows that the usual approach to quantum gravity, based on straight perturbation theory in the gravitational coupling applied to four dimensions is essentially flawed, and leads to fundamentally misleading answers. The $2 + \epsilon$ expansion approach to gravity provides entirely new insights into what is most likely the true ground state of the theory, but faces a tremendous challenge in reaching the physical case of dimensions four, and furthermore gives few hints as to what the true nature of the strong coupling ground state arising for $G > G_c$ might be.

Due to its inherent complexity, the lattice approach is generally not particularly well suited for analytical investigations; even the simple task of connecting the lattice weak field limit with the corresponding continuum result represents a small algebraic tour de force. Yet the lattice theory appears to provide a concrete and constructive proof for the existence of the gravitational functional integral, at least in the regime where it exists, and perhaps the only reliable framework in which the issue of the gravitational measure can be properly addressed.

The existence of a small set of clear, unambiguous analytical results from the $2 + \epsilon$ expansions provides the opportunity, as is the case already in the non-linear sigma model, of using the lattice theory to address a set of basic issues directly in four dimensions, relying at the same time on a pre-existing limited theoretical framework of scaling relations and running coupling scenarios. The natural underlying assumption is therefore that the four-dimensional results are qualitatively similar to the analytical results, but with somewhat different amplitudes and exponents. The numerical evidence so far suggests that such an identification is warranted, and is in fact at least up to now completely consistent.

At the same time, the lattice theory provides new essential ingredients, such as the non-existence (in the Euclidean field theory framework) of the weak coupling phase, and the appearance of a renormalization group invariant gravitational correlation length associated with large scale curvature. Furthermore the specific values of the critical exponents in four dimensions suggest possible scenarios (such as a non-local effective theory) by which new non-trivial analytical results in four dimensions might be obtained. In any case, there seems to be a clear prediction that gravitational couplings will be scale dependent. One can in fact re-phrase the last sentence even more strongly, to the effect that it would seem very difficult to accommodate in a quantum theory of gravity couplings that are not scale dependent.

One might view the non-perturbative lattice results in many respects as quite unsatisfactory. Since the lattice theory does not fix the correlation length ξ , the latter remains completely undetermined and has to be fixed by physical considerations. Since it relates to curvature, it is natural to take it to be very large so as to recover agreement with observation. Consequently there is no clear explanation for the smallness of the cosmological constant, which instead has to be tuned to its physical value. Furthermore, since short-distance cutoff physics essentially decouples from universal long distance physics, there is no indication of what specific cutoff mechanism might be operative at short distances.

There is also the possibility of clear disagreements between theoretical predictions and observation. In the strong coupling, physical phase the average curvature for infinitesimal loops is negative, corresponding to an apparent Euclidean anti-de Sitter phase at very short distances, comparable to the ultraviolet cutoff. What happens for large loops is a more difficult and largely open questions, although there are indications that for large loops the average curvature is positive instead. In any case it is not clear yet that the sign of the curvature is a truly universal quantity. Since the recent distant supernova observations suggest a positive cosmological constant, there is potential for serious conflict. The analysis gets clouded further by the fact that in the strong coupling phase a growing gravitational coupling at large distance (corresponding to gravitational anti-screening) leads to cosmic acceleration and would therefore in part mimic the effects of a positive cosmological constant.

Could then the weak coupling phase still be physical, in spite of the fact that it appears pathological in the Euclidean lattice theory? In principle, as a last measure, such a phase could be reached by analytic continuation from the strong coupling phase of the Euclidean theory (it is easy to see that for example the β -function is expected to be analytic at the fixed point), and the same exponent ν would apply to both sides of the transition, as in the non-linear sigma model. In this gravitational “Coulomb” phase the correlation length ξ would be infinite, the cosmological constant presumably identically zero, and one would expect gravitational screening on some scale. How one would avoid the conclusion that in this phase space-time is essentially two-dimensional is unclear, at least to the author.

Acknowledgments

I would like to thank first of all Ruth Williams for her collaboration on the many topics covered in this review. In addition I have benefited over the years from discussions with James Bjorken,

Thibault Damour, Stanley Deser, James Hartle, Giorgio Parisi, Mike Peskin, Tullio Regge, Gabriele Veneziano and Jean Zinn-Justin on topics related to the contents of this review, and many other colleagues. I wish to thank in particular Hermann Nicolai, Robert Schrader and Stefan Theisen for their comments on the manuscript.

This work was supported in part by a research grant from the Max Planck Gesellschaft zur Förderung der Wissenschaften. The author wishes to thank Hermann Nicolai and the Max Planck Institut für Gravitationsphysik (Albert-Einstein-Institut) in Potsdam for a very warm hospitality.

References

- [1] Abers, E. S. , 2004, *Quantum Mechanics*, Prentice Hall, New Jersey.
- [2] Abbott, L., 1981, *Introduction to the Background Field Method*, CERN-Th 3113.
- [3] Aida, T. , Y. Kitazawa, H. Kawai and M. Ninomiya, 1994, *Nucl. Phys. B* **427**, 158.
- [4] Aida, T. , and Y. Kitazawa, 1997, *Nucl. Phys.* **B491** 427.
- [5] Albeverio, S. A. , and Hoegh-Krohn R. J., 1976, *Mathematical Theory of Feynmann Path Integrals*, Lecture Notes in Mathematics, No. 523, Springer-Verlag, Berlin.
- [6] Ambjorn, J. , M. Carfora and A. Marzuoli, 1997, ‘*The Geometry of Dynamical Triangulations*’, Lecture Notes in Physics (Springer Verlag, Berlin).
- [7] Aoki, S. , *et al.* [CP-PACS Collaboration], 2003, *Phys. Rev. D* **67**, 034503.
- [8] Arnowitt, R. , S. Deser and C. W. Misner, 1960, *Phys. Rev.* **117**, 1595.
- [9] Arnowitt, R. , S. Deser and C. W. Misner, 1962, *The Dynamics Of General Relativity*, in *Gravitation: An Introduction to Current Research*, Louis Witten ed. (Wiley 1962), chapter 7, pp. 227-265 (gr-qc/0405109).
- [10] Avis, S. J. , C. J. Isham and D. Storey, 1978, *Phys. Rev. D* **18**, 3565.
- [11] Avramidy, I. G. , and A. O. Barvinsky, 1985, *Phys. Lett.* **159B** 269.
- [12] Bakeyev, T. D. , and A. A. Slavnov, 1996 *Mod. Phys. Lett. A* **11**, 1539.
- [13] Balian, R. , J. M. Drouffe and C. Itzykson, 1975, *Phys. Rev. D* **11**, 2104, [Erratum-ibid. **D 19**, 2514 (1979)].
- [14] Bardeen, W. A. , B. W. Lee and R. E. Shrock, 1976, *Phys. Rev.* **D14**, 985.
- [15] Barrett, J. W. , M. Galassi, W. A. Miller, R. D. Sorkin, P. A. Tuckey and R. M. Williams, 1997, *Int. J. Theor. Phys.* **36**, 815.
- [16] Barth, N. H. , and S. M. Christensen, 1983, *Phys. Rev.* **D28**, 1876.
- [17] Barvinsky, A. O. , 2003, *Phys. Lett.* **B572** 109.
- [18] Barvinsky, A. O. , 2007, *Phys. Rev. Lett.* **99** 071301.
- [19] Barvinsky, A. O. , and G. A. Vilkovisky, 1985, *Phys. Rept.* **119** 1.
- [20] Beneke, M. , 1999, *Renormalons*, *Physics Reports* **317** 1.

- [21] Berg, B. , 1985, *Phys. Rev. Lett.* **55** 904; *Phys. Lett.* **B176** 39 (1986).
- [22] Beirl, W. , E. Gerstenmayer, H. Markum and J. Riedler, 1994, *Phys. Rev.* **D49** 5231.
- [23] Bernreuther, W. and F. J. Wegner, 1986, *Phys. Rev. Lett.* **57**, 1383.
- [24] Bern, Z. , L. J. Dixon and R. Roiban, 2007, *Phys. Lett. B* **644**, 265.
- [25] Bittner, E. , W. Janke and H. Markum, 2002, *Phys. Rev.* **D66**, 024008.
- [26] Brans, C. , and R. H. Dicke, 1961, *Phys. Rev.* **124** 925.
- [27] Brewin, L. C. , and A. P. Gentle, 2001, *Class. Quant. Grav.* **18**, 517.
- [28] Brezin, E. and J. Zinn-Justin, 1976, *Phys. Rev. Lett.* **36**, 691.
- [29] Brezin, E. , J. Zinn-Justin and J. C. Le Guillou, 1976, *Phys. Rev. D* **14**, 2615.
- [30] Brezin, E. , and S. Hikami, 1996, LPTENS-96-64 (Dec 1996), cond-mat/9612016.
- [31] Brezin, E. , J. C. Le Guillou, J. Zinn-Justin, 1976, in *Phase Transitions and Critical Phenomena*, Vol. 6, eds. C. Domb and M. S. Green, Academic Press, New York.
- [32] Brezin, E. , and J. Zinn-Justin, 1985, *Nucl. Phys. B* **257**, 867.
- [33] Boulware, D. G. , and S. Deser, 1975, *Annals Phys.* **89**, 193.
- [34] Callan, C. G. . , 1970, *Phys. Rev. D* **2**, 1541.
- [35] Campostrini, M. , M. Hasenbusch, A. Pelissetto, P. Rossi and E. Vicari, 2001, *Phys. Rev.* **B63**, 214503.
- [36] Capper, D. M. , G. Leibbrandt and M. Ramon Medrano, 1973, *Phys. Rev. D* **8**, 4320.
- [37] Caracciolo, S. , and A. Pelissetto, 1988, *Nucl. Phys. B* **299**, 693.
- [38] Cardy, J. L. (ed.), 1988, *Finite Size Scaling*, Current Physics - Sources and Comments, North Holland.
- [39] Cardy, J. L. , 1997, *Scaling and the Renormalization Group in Statistical Physics*, Cambridge University press.
- [40] Cardy, J. L. and H. W. Hamber, 1980, *Phys. Rev. Lett.* **45**, 499.
- [41] Capozziello, S. , S. Carloni and A. Troisi, 2003, in *Recent Research Developments in Astronomy and Astrophysics*, RSP/AA/21-03, astro-ph/0303041.
- [42] Carroll, S. M. , V. Duvvuri, M. Trodden and M. S. Turner, 2004, *Phys. Rev. D* **70**, 043528.
- [43] Caselle, M. , A. D'Adda and L. Magnea, 1989, *Phys. Lett. B* **232**, 457.
- [44] Cheeger, J. , W. Müller and R. Schrader, 1982, in *Unified Theories Of Elementary Particles*, Heisenberg Symposium, München 1981, Springer Lecture Notes in Physics, P. Breitlohner and H. P. Dürr eds., New York 1982.
- [45] Cheeger, J., W. Müller and R. Schrader, 1984, *Comm. Math. Phys.* **92**, 405.
- [46] Christ, N. H. , R. Friedberg and T. D. Lee, 1982, *Nucl. Phys. B* **202**, 89.
- [47] Christensen, S. M. , and M.J. Duff, 1978, *Phys. Lett.* **B79** 213.
- [48] Christensen, S. M. , M. J. Duff, G. W. Gibbons and M. Roček, 1980, *Phys. Rev. Lett.* **45**, 161.
- [49] Coxeter, H. , 1948, *Regular Polytopes*, Methuen and Co. Ltd., London.
- [50] Coxeter, H. , 1974, *Regular Complex Polytopes*, Cambridge University press.
- [51] Cremmer, E. , and J. Scherk, 1977, *Nucl. Phys. B* **127**, 259.
- [52] Cremmer, E. , B. Julia and J. Scherk, 1978, *Phys. Lett. B* **76**, 409.

- [53] Cremmer, E. , and B. Julia, 1978, *Phys. Lett. B* **80**, 48; *Nucl. Phys. B* **159**, 141 (1979).
- [54] Collins, P. A. , and R. M. Williams, 1972, *Phys. Rev.* **D5**, 1908; **D7** (1973) 965; **D10** (1974) 353.
- [55] Damour, T. , 2006, *Experimental Tests of Gravitational Theory*, in the 2006 Review of Particle Properties, W. M. Yao et al., *J. Phys.* **G 33**, 1.
- [56] Das, A. , 1977, *Phys. Rev. D* **15**, 2805.
- [57] Das, A., M. Kaku and P.K. Townsend, 1979, *Phys. Lett.* **81B** 11.
- [58] David, F. , 1982, *Nucl. Phys. B* **209**, 433.
- [59] D'Eath, P. D. , 1984, *Phys. Rev. D* **29**, 2199 [Erratum-ibid. *D 32*, 1593 (1985)].
- [60] D'Eath, P. D. , 1994, *Phys. Lett. B* **321**, 368.
- [61] de Bakker, B. , and J. Smit, 1995, *Nucl. Phys.* **B439** 239.
- [62] de Calan, C. , P. A. Faria da Veiga, J. Magnen, R. Seneor, 1991, *Phys. Rev. Lett.* **66** 3233.
- [63] Deser, S. , and P. van Nieuwenhuizen, 1974, *Phys. Rev.* **D10** 401, 411.
- [64] Deser, S. , and B. Zumino, 1976, *Phys. Lett. B* **62**, 335.
- [65] Deser, S. , J. H. Kay and K. S. Stelle, 1977, *Phys. Rev. Lett.* **38**, 527.
- [66] Deser, S. , R. Jackiw and S. Templeton, 1982, *Annals Phys.* **140**, 372; *ibid.* **185**, 406 (1988).
- [67] Deser, S. , R. Jackiw and G. 't Hooft, 1984, *Annals Phys.* **152**, 220.
- [68] Deser, S. , and R. Jackiw, 1984, *Annals Phys.* **153** 405.
- [69] DeWitt, B. S. , 1962, *J. Math. Phys.* **3**, 1073.
- [70] DeWitt, B. S. , 1964, *Dynamical Theory of Groups and Fields*, in *Relativity, Groups and Topology*, 1963 Les Houches Lectures, eds. B. de Witt and C. de Witt, Gordon and Breach, New York 1964.
- [71] DeWitt, B. S. , 1967, *Phys. Rev.* **160**, 1113.
- [72] DeWitt, B. S. , 1967, *Phys. Rev.* **162**, 1195.
- [73] DeWitt, B. S. , 1967, *Phys. Rev.* **162**, 1239.
- [74] Dirac, P. A. M. , 1958, *Proc. Roy. Soc. (London)* **A246**, 333.
- [75] Domb, C. , and H. S. Green, 1976, '*Phase Transitions and Critical Phenomena*', (Academic Press, London 1976), Vol. 3.
- [76] Drouffe, J. M. , G. Parisi and N. Surlas, 1979, *Nucl. Phys.* **B 161**, 397.
- [77] Drummond, I. T. , 1986, *Nucl. Phys. B* **273**, 125.
- [78] Dubal, M. R. , 1989, *Class. Quant. Grav.* **6**, 141; **6**, 1925; **7**, 371 (1990).
- [79] Faddeev, L. , and V. Popov, 1974, *Sov. Phys. Usp.* **16** 777 (1974); *Usp. Fiz. Nauk.* **109** 427 (1974).
- [80] Feynman, R. P. , 1948, *Rev. Mod. Phys.* **20**, 367 - 387.
- [81] Feynman, R. P. , 1950, *Phys. Rev.* } **80**, 440.
- [82] Feynman, R. P. , 1962, *Lectures on Gravitation, 1962-1963*, edited by F. B. Morinigo and W. G. Wagner, *California Institute of Technology (Pasadena, 1971)*.
- [83] Feynman, R. P. , 1963, *Act. Phys. Pol.* **24** 697.
- [84] Feynman, R. P. , and A. Hibbs, 1965, *Quantum Mechanics and Path integrals*, (McGraw-Hill, New York).

- [85] Fisher, M. E. , 1967, *Physics* **3** 255.
- [86] Fischer, P. , and D. F. Litim, 2006, *Phys. Lett. B* **638**, 497.
- [87] Flanagan, E. E. , 2004, *Phys. Rev. Lett.* **92**, 071101.
- [88] Frampton, P. , 2000, ‘Gauge Field Theories’, (Wiley, New York).
- [89] Fradkin, E. S. , and G. A. Vilkovisky, 1973, *Phys. Rev.* **D8**, 4241.
- [90] Fradkin, E. S. , and G. A. Vilkovisky, 1977, *CERN preprint TH-2332*.
- [91] Fradkin, E. S. , and A. A. Tseytlin, 1981, *Phys. Lett.* **104B** 377; *Nucl. Phys.* **B201** (1982) 469.
- [92] Freedman, D. Z. , P. van Nieuwenhuizen and S. Ferrara, 1976, *Phys. Rev. D* **13**, 3214.
- [93] Forger, K. , B. A. Ovrut, S. J. Theisen and D. Waldram, 1996, *Phys. Lett. B* **388**, 512.
- [94] Fröhlich, J. , 1981, *I. H. E. S. preprint 1981*, in *Non-Perturbative Quantum Field Theory: Mathematical Aspects and Applications, Selected Papers* (World Scientific, Singapore, 1992), pp. 523-545.
- [95] Fujikawa, K. , 1983, *Nucl. Phys.* **B226** 437.
- [96] Galassi, M. , 1993, *Phys. Rev. D* **47**, 3254-64.
- [97] Gallavotti, G. , 1985, *Rev. Mod. Phys.* **57**, 471.
- [98] Gastmans, R. , R. Kallosh and C. Truffin, 1978, *Nucl. Phys.* **B133** 417.
- [99] Gawedzki, K. , and A. Kupiainen, 1985, *Phys. Rev. Lett.* **54** 2191; **55** 363.
- [100] Gell-Mann, M. , and F. E. Low, 1954, *Phys. Rev.* **95**, 1300.
- [101] Gentle, A. P. , 2002, *Gen. Rel. Grav.* **34**, 1701.
- [102] Gentle, A. P. , and W. A. Miller, 1998, *Class. Quant. Grav.* **15**, 389.
- [103] Gibbons, G. W. , and S. W. Hawking, 1977, *Phys. Rev. D* **15**, 2752.
- [104] Gibbons, G. W. , S. W. Hawking and M. J. Perry, 1978, *Nucl. Phys. B* **138**, 141.
- [105] Gibbons, G. W. , and M. J. Perry, 1978, *Nucl. Phys. B* **146**, 90.
- [106] Giddings, S. B. , D. Marolf and J. B. Hartle, 2006, *Phys. Rev. D* **74**, 064018.
- [107] Gilkey, P. , 1975, *J. Diff. Geom.* **10** (1975), 601-618.
- [108] Glimm, J. , and A. Jaffe, 1981, *Quantum Mechanics - A Functional Integral Point of View*, Springer-Verlag, Berlin, 1981.
- [109] Glimm, J. , and A. Jaffe, 1974, *Phys. Rev. Lett.* **33**, 440 - 442.
- [110] Goroff, M. H. , and A. Sagnotti, 1985, *Phys. Lett. B* **160**, 81; *Nucl. Phys. B* **266**, 709 (1986).
- [111] Goldberg, J. N. , 1958, *Phys. Rev.* **111**, 315.
- [112] Griffiths, R. B. , 1969, *Phys. Rev. Lett.* **23** 17.
- [113] Grisaru, M. T. , 1977, *Phys. Lett. B* **66**, 75.
- [114] Grisaru, M. T. , P. van Nieuwenhuizen and J. A. M. Vermaseren, 1976, *Phys. Rev. Lett.* **37**, 1662.
- [115] Gross, D. J. , and A. Neveu, 1974, *Phys. Rev. D* **10**, 3235.
- [116] Gross, D. J. , 1984, *Phys. Lett. B* **138**, 185.
- [117] Guida, R. and J. Zinn-Justin, 1998, *J. Phys. A* **31**, 8103.
- [118] Halliwell, J. J. , and J. B. Hartle, 1990, *Phys. Rev. D* **41**, 1815.
- [119] Halliwell, J. J. , and J. Louko, 1989, *Phys. Rev. D* **39**, 2206.

- [120] Halliwell, J. J. , and J. Louko, 1989, *Phys. Rev. D* **40**, 1868.
- [121] Halliwell, J. J. , and J. Louko, 1989, *Phys. Rev. D* **42**, 3997.
- [122] Hamber, H. W. , 1981, *Phys. Rev.* **D24**, 941.
- [123] Hamber, H. W. , 1984, *Simplicial Quantum Gravity*, in *Critical Phenomena, Random Systems, Gauge Theories*, 1984 Les Houches Summer School, Session XLIII, edited by K. Osterwalder and R. Stora (North Holland, Amsterdam, 1986).
- [124] Hamber, H. W. , 1993, *Nucl. Phys. B* **400**, 347.
- [125] Hamber, H. W. , 1994, *Phys. Rev.* **D50** 3932.
- [126] Hamber, H. W. , 2000, *Phys. Rev.* **D61** 124008; *Phys. Rev.* **D45** 507 (1992).
- [127] Hamber, H. W. , and S. Liu, 1995, *Phys. Lett.* **B357** 51.
- [128] Hamber, H. W. , and G. Parisi, 1983, *Phys. Rev. D* **27**, 208.
- [129] Hamber, H. W. , and G. Kagal, 2004, *Class. Quant. Grav.* **21**, 5915.
- [130] Hamber, H. W. , and R. M. Williams, 1984, *Nucl. Phys. B* **248**, 392 [Erratum-*ibid.* *B* **260**, 747 (1985)]; *Nucl. Phys. B* **269**, 712 (1986).
- [131] Hamber, H. W. , and R. M. Williams, 1985, *Phys. Lett. B* **157**, 368.
- [132] Hamber, H. W. , and R. M. Williams, 1993, *Phys. Rev.* **D47** 510.
- [133] Hamber, H. W. , and R. M. Williams, 1995, *Nucl. Phys.* **B435** 361.
- [134] Hamber, H. W. , and R. M. Williams, 1999, *Phys. Rev.* **D59** 064014.
- [135] Hamber, H. W. , and R. M. Williams, 2005, *Phys. Rev. D* **72**, 044026.
- [136] Hamber, H. W. , and R. M. Williams, 2006, *Phys. Rev. D* **73**, 044031.
- [137] Hamber, H. W. , and R. M. Williams, 2007, *Phys. Rev. D* **75**, 084014.
- [138] Hamber, H. W. , and R. M. Williams, 2007, *Phys. Rev. D* **76**, 084008.
- [139] Hartle, J. B. , 1984, *Phys. Rev.* **D** **29**, 2730.
- [140] Hartle, J. B. , 1985, *J. Math. Phys.* **26**, 804.
- [141] Hartle, J. B. , 1986, *J. Math. Phys.* **27**, 287.
- [142] Hartle, J. B. , 1989, *J. Math. Phys.* **30**, 452.
- [143] Hartle, J. B. , and S. W. Hawking, 1976, *Phys. Rev. D* **13**, 2188.
- [144] Hartle, J. B. , and K. V. Kuchar, 1986, *Phys. Rev. D* **34**, 2323.
- [145] Hartle, J. B. , W. A. Miller and R. M. Williams, 1997, *Class. and Quant. Gravity* **14**, 2137.
- [146] Hasslacher, B. , and M. J. Perry, 1981, *Phys. Lett. B* **103**, 21.
- [147] Hawking, S. W. , 1977, *Commun. Math. Phys.* **55**, 133.
- [148] Hawking, S. W. , 1978, *Phys. Rev. D* **18**, 1747.
- [149] Hawking, S. W. , 1978, *Nucl. Phys. B* **144**, 349.
- [150] Hawking, S. W. , 1979, in ‘General Relativity - An Einstein Centenary Survey’, edited by S. W. Hawking and W. Israel, (Cambridge University Press, 1979).
- [151] Hawking, S. W. , 1984, *Phys. Lett.* **134B**, 403.
- [152] Hawking, S. W. , and T. Hertog, 2002, *Phys. Rev. D* **65**, 103515.

- [153] Heisenberg, W. , 1938, *Z. Phys.* **110**, 251; *ibid.* **113**, 61 (1939).
- [154] Hikami, S. , and E. Brezin, 1978, *J. Phys. A* **11**, 1141.
- [155] Hikami, S. , 1983, *Nucl. Phys.* **B215** [FS7], 555.
- [156] Hughes, R. J. , 1980, *Phys. Lett. B* **97**, 246; *Nucl. Phys. B*, 186, 376 (1981).
- [157] Immirzi, G. , 1996, *Class. Quantum Grav.* **13** 2385-2393..
- [158] Itzykson, C. , 1983, *Fields On A Random Lattice*, in *Progress in Gauge Field Theory*, Cargèse, France, Sep 1-15, 1983.
- [159] Itzykson, C. , and M. Bander, 1983, in *Non-linear equations in field theory*, Meudon Seminar (Springer-Verlag, 1983).
- [160] Itzykson, C. , and Zuber J. -B. , 1980, *Quantum Field Theory*, (McGraw-Hill, New York, 1980).
- [161] Jevicki, A. , and M. Ninomiya, 1985, *Phys. Lett. B* **150**, 115.
- [162] Julve, J. , and M. Tonin, 1978, *Nuovo Cimento* **46B**, 137.
- [163] Kawai, H. , and M. Ninomiya, 1990, *Nucl. Phys.* **B336** 115.
- [164] Kawai, H. , Y. Kitazawa and M. Ninomiya, 1993, *Nucl. Phys.* **B393** 280; **B404** 684 (1993).
- [165] Khatsymovsky, V. , 1991, *Class. Quantum Grav.* **8** 1205-1216 (1991).
- [166] Kibble, T. W. B. , 1961, *J. Math. Phys.* **2**, 212.
- [167] Kivel, N. A. , A. S. Stepanenko and A. N. Vasiliev, 1994, *Nucl. Phys. B* **424**, 619.
- [168] Kleinert, H. , 2006, *Path Integrals in Quantum Mechanics, Statistics and Polymer Physics*, World Scientific.
- [169] Kleinert, H. , 2000, *Phys. Lett. A* **277**, 205.
- [170] Kondo, K. , 1984, *Prog. Theor. Phys.* **72**, 841.
- [171] Konopleva, N. P. , and V. N. Popov, 1979, 'Gauge Fields', (Harwood Academic Publishers, New York 1979).
- [172] Kosterlitz, J. M. , and D. J. Thouless, 1973, *J. Phys. CC* **6**, 1181.
- [173] Kuchař, K. , 1970, *Jour. Math. Phys.* **11**, 3322.
- [174] Lafontaine, J. , 1986, *Seminaire Bourbaki*, 38^eme annee, 1985-86, n. 664.
- [175] Lanczos, C. , 1938, *The Annals of Mathematics* **39**, 842.
- [176] Langer, J. S. , 1967, *Phys. Rev. Lett.* **21**, 973; *Ann. Phys.* **41**, 108 (1967).
- [177] Lauscher, O. , and M. Reuter, 2002, *Class. Quant. Grav.* **19** 483.
- [178] Lee, T. D. , 1983, in 'Discrete Mechanics', 1983 *Erice International School of Subnuclear Physics*, vol. 21 (Plenum Press, New York 1985); R. Friedberg and T. D. Lee, *Nucl. Phys.* **B242** 145 (1984).
- [179] Le Guillou, J. C. , and J. Zinn-Justin, 1980, *Phys. Rev.* **B21** 3976; *J. Phys. (France)* **50** 1395 (1989).
- [180] Lehto, M. , H. B. Nielsen and M. Ninomiya, 1987, *Nucl. Phys. B* **289**, 684.
- [181] Lense, J. , and H. Thirring, 1918, *Physik. Zeitsch.* **19** 156.
- [182] Leutwyler, H. , 1964, *Phys. Rev.* **134**, 1155.
- [183] Lipa, J. A. et al., 2003, *Phys. Rev.* **B68**, 174518.
- [184] Litim, D. , 2004, *Phys. Rev. Lett.* **92** 201301.

- [185] Loll, R. , W. Westra and S. Zohren, 2006 *Nucl. Phys. B* **751**, 419.
- [186] Lopez Nacir, D. , and F. D. Mazzitelli, 2007, *Phys. Rev. D* **75**, 024003.
- [187] Lubensky, T. C. , and J. Isaacson, 1979, *Phys. Rev.* **A20** 2130.
- [188] Lund, F. , and T. Regge, 1974, *Princeton preprint (unpublished)*.
- [189] Lüscher, M. , 1987, ‘Solution of Lattice ϕ^4 Theory in 4 Dimensions’, in *Non-Perturbative Quantum Field Theory, Cargèse 1987*.
- [190] MacDowell, S. W. , and F. Mansouri, 1977, *Phys. Rev. Lett.* **38**, 739 [Erratum-*ibid.* **38**, 1376 (1977)].
- [191] Mannion, C. , and J. G. Taylor, 1981, *Phys. Lett.* **100B**, 261.
- [192] Meijering, J. L. , 1953, *Philips Research Reports* **8**, 270.
- [193] Miller, W. A. , and J. A. Wheeler, 1985, *Nuovo Cimento* **8**, 418-34.
- [194] Misner, C. W. , 1957, *Rev. Mod. Phys.* **29**, 497.
- [195] Modanese, G. , 1993, *Phys. Rev. D* **47**, 502; *Phys. Rev. D* **49**, 6534 (1994).
- [196] Modanese, G. , 1995, *Nucl. Phys. B* **434**, 697.
- [197] Morris, T. R. , 1994, *Phys. Lett. B* **329**, 241; *Phys. Lett. B* **334**, 355.
- [198] Morris, T. R. , and M. D. Turner, 1998, *Nucl. Phys. B* **509**, 637.
- [199] Myers, R. C. , and M. J. Perry, 1986, *Annals Phys.* **172**, 304.
- [200] Nicolai, H. , 1984, *Supersymmetric Functional Integration Measures, Lectures given at 1984 NATO ASI Inst. on Supersymmetry, Bonn West Germany, Aug. 20-31*.
- [201] Nicolai, H. , and P. K. Townsend, 1981, *Phys. Lett. B* **98**, 257.
- [202] Nielsen, N. K. , 1981, *Am. J. Phys.* **49**, 1171.
- [203] Nishimoto, K. , 1990, (ed.), *Fractional Calculus and Applications, in the Proceedings of the 3rd Conference on Fractional Calculus, Tokyo 1989, Nihon University, Japan, 1990*.
- [204] Osterwalder, K. , and R. Schrader, 1972, *Phys. Rev. Lett.* **29**, 1423-1425.
- [205] Osterwalder, K. , and R. Schrader, 1973, *Comm. Math. Phys.*, **31**, 83-112; *Comm. Math. Phys.* , **42** (1975), 281-305.
- [206] Parisi, G. , 1973, *Lett. Nuovo Cim.* **6S2**, 450.
- [207] Parisi, G. , 1975, *Nucl. Phys.* **B100** 368.
- [208] Parisi, G. , 1976, ‘On NonRenormalizable Interactions’, in *the proceedings of the Cargèse Summer Institute (1976)*.
- [209] Parisi, G. , 1981, *Statistical Field Theory, Benjamin Cummings, New York, 1981*.
- [210] Parisi, G. , 1985, *Nucl. Phys.* **254** 58 (1985).
- [211] Pauli, W. , 1958, *Teoria della Relatività, (Boringhieri, Torino 1958) (English edition: Theory of Relativity, Dover Publications 1956)*.
- [212] Peskin, M. E. , and D. V. Schroeder, 1995, ‘An Introduction to Quantum Field Theory’, *sec. 22.1 (Addison Wesley, Reading, Massachusetts, 1995)*.
- [213] Piran, T. , and R. M. Williams, 1986, *Phys. Rev. D* **33**, 1622.
- [214] Polyakov, A. M. , 1975, *Phys. Lett.* **59B**, 79.

- [215] Polyakov, A. M. , 1989, 'Gauge Fields and Strings', (Oxford University Press, 1989).
- [216] Ponzano, G. and T. Regge, 1968, in Spectroscopic and Group Theoretical Methods in Physics, eds. F. Bloch, S. G. Cohen, A. De-Shalit, S. Sambursky, I. Talmi, (North-Holland Amsterdam, 1968), p. 1-58.
- [217] Porter, J. D. , 1987, *Class. Quantum Grav.* **4**, 375(1987); **4**, 391; **4**, 651.
- [218] Regge, T. , 1961, *Nuovo Cimento* **19** 558.
- [219] Regge, T. , and R. M. Williams, 2000, *J. Math. Phys.* **41**, 3964.
- [220] Reuter, M. , 1998, *Phys. Rev. D* **57**, 971.
- [221] Reuter, M. , and F. Saueressig, 2002, *Phys. Rev.* **D65** 065016.
- [222] Richardson, J. L. , 1979, *Phys. Lett. B* **82**, 272.
- [223] Riedler, J. , W. Beirl, E. Bittner, A. Hauke, P. Homolka and H. Markum, 1999, *Class. Quant. Grav.* **16**, 1163.
- [224] Roček, M. , and R. M. Williams, 1981, *Phys. Lett.* **104B** 31; *Z. Phys.* **C21** 371 (1984).
- [225] Roček, M. , and R. M. Williams, 1985, *Class. Quantum Grav.* **2** 701-706.
- [226] Sakata, S. , H. Umezawa and S. Kamefuchi, 1952, *Prog. Theor. Phys.* **7**, 377.
- [227] Salam, A. , and J. Strathdee, 1978, *Phys. Rev. D* **18**, 4480.
- [228] Samko, S. G. , A. A. Kilbas and O. I. Marichev, 1993, *Fractional Integrals and Derivatives: Theory and Application*, (Gordon and Breach, 1993).
- [229] Schleich, K. , 1985, *Phys. Rev. D* **32** 1889.
- [230] Schleich, K. , 1987, *Phys. Rev. D* **36**, 2342.
- [231] Sciamia, D. W. , 1953, *Monthly Notices of the Royal Astronomical Society* **113**, 34-42
- [232] Singer, I. M. , and J. A. Thorpe, 1967, *Lecture Notes in Elementary Topology and Geometry*, Scott, Foresman and Co., Glenview, Ill. 1967.
- [233] Slavnov, A. A. , 1972, *Teor. Mat. Fiz.* **13**, 174; **33**, 210 (1977).
- [234] Smolin, L. , 1979, *Nucl. Phys.* **B148**, 333.
- [235] Sorkin, R. D. , 1975, *Phys. Rev. D* **12**, 385-96.
- [236] Sorkin, R. D. , 1975, *J. Math. Phys.* **16**, 2432 [Erratum-*ibid.* **19**, 1800 (1978)].
- [237] Stelle, K. S. , 1977, *Phys. Rev. D* **16**, 953.
- [238] Stelle, K. S. , 1978, *Gen. Rel. Grav.* **9**, 353.
- [239] Streater, R. F. , and A. S. Wightman, 1964, *PCT, Spin and Statistics and All That*, Princeton University Press, 2000.
- [240] Stückelberg, E. C. , and A. Petermann, 1953, *Helv. Phys. Acta* **26**, 499-520.
- [241] Symanzik, K. , 1968, *Euclidean Quantum Field Theory, Varenna Lectures 1968*, ed. R. Jost (Academic Press, New York, 1969).
- [242] Symanzik, K. , 1970, *Commun. Math. Phys.* **18**, 227.
- [243] Symanzik, K. , 1983, *Nucl. Phys. B* **226**, 187; *ibid.* 205.
- [244] Strominger, A. , 1981, *Phys. Rev. D* **24**, 3082.

- [245] 't Hooft, G. , and M. Veltman, 1974, *Ann. Inst. Poincaré* **20**, 69.
- [246] 't Hooft, G. , 1978, in 'Recent Developments in Gravitation', *Cargèse Lecture notes 1978*, M. Levy and S. Deser eds., *Nato Science Series*, Springer 1979.
- [247] 't Hooft, G. , 2002, *Perturbative Quantum Gravity*, *Erice Lecture Notes*, *Subnuclear Physics Series* Vol. 40, ed. A. Zichichi, World Scientific, Singapore.
- [248] Tomboulis, E. T. , 1984, *Phys. Rev. Lett.* **52**, 1173.
- [249] Townsend, P. K. , 1977, *Phys. Rev. D* **15**, 2795; *Phys. Rev. D* **15** (1977) 2802.
- [250] Tuckey, P. A. , 1989, *Class. Quantum Grav.* **6** 1-21; **7** 2055-2071 (1990).
- [251] Utiyama, R. , and B. DeWitt, 1962, *J. Math. Phys.* **3**, 608.
- [252] van de Ven, A. E. M., 1992 *Nucl. Phys.***B** 378, 309.
- [253] van Nieuwenhuizen, P. , 1973, *Nucl. Phys.* **B60** 478.
- [254] Veltman, M. , 1975, *Quantum Theory of Gravitation*, in 'Methods in Field Theory', *Les Houches Lecture notes session XXVIII (1975)*.
- [255] Vilkovisky, G. A. , 1984, *Nucl. Phys. B* **234**, 125.
- [256] Voronoi, M. G. , 1908, *J. Reine Angew. Math.* 134, 198-287.
- [257] Warner, N. P. , 1982, *Proc. Roy. Soc. Lond. A* **383**, 359.
- [258] Wegner, F. J. , 1989, *Nucl. Phys.* **B316**, 663.
- [259] Weinberg, S. , 1979, *Ultraviolet divergences in quantum gravity*, in 'General Relativity - An Einstein Centenary Survey', edited by S. W. Hawking and W. Israel, (Cambridge University Press, 1979).
- [260] Weinberg, S. , 1997, *Phys. Rev.* **D56** 2303.
- [261] West, P. , 1978, *Phys. Lett. B* **76**, 569.
- [262] Wetzels, W. , 1985, *Phys. Lett. B* **153**, 297.
- [263] Weyl, H. , 1918, *Raum, Zeit, Materie - Vorlesungen ber die Allgemeine Relativittstheorie*, (Berlin : Springer, 1918); (English edition: 'Space, Time, Matter, Dover, New York 1952).
- [264] Wheeler, J. A. , 1964, *Geometrodynamics and the Issue of the Final State*, in *Relativity, Groups and Topology*, eds. B. DeWitt and C. DeWitt (New York, Gordon and Breach, 1964) 463-500.
- [265] Whitt, B. , 1984, *Phys. Lett. B* **145**, 176.
- [266] Williams, R. M. and G. F. R. Ellis, 1981, *Gen. Rel. Grav.* **13** 361; **16** (1984) 1003.
- [267] Williams, R. M. , 1986, *Class. Quantum Grav.* **3** 853-869.
- [268] Wilson, K. G. , 1971, *Phys. Rev. B* **4**, 3174; *ibid.* 3184.
- [269] Wilson, K. G. , 1972, *Phys. Rev. Lett.* **28**, 548.
- [270] Wilson, K. G. , and M. E. Fisher, 1972, *Phys. Rev. Lett.* **28**, 240.
- [271] Wilson, K. G. , 1973, *Phys. Rev. D* **7**, 2911.
- [272] Wilson, K. G. , 1974, *Phys. Rev. D* **10**, 2445.
- [273] Wilson, K. G. , 1975, *Rev. Mod. Phys.* **47** 773; **55** 583 (1983).
- [274] Wong, C. Y. , 1971, *J. Math. Phys.* **12**, 70.
- [275] Wulf, M. , 1997, *Int. J. Mod. Phys. D* **6**, 107.

- [276] Zamolodchikov, A. B. , and A. B. Zamolodchikov, 1978, *Phys. Lett. B* **72**, 481.
- [277] Zavada, P. , 1998, *Commun. Math. Phys.* **192**, 261 [[funct-an/9608002](#)].
- [278] Zinn-Justin, J. , 2002, *Quantum Field Theory and Critical Phenomena*, Oxford University Press, fourth edition, 2002.
- [279] Zinn-Justin, J. , 2005, *Path Integrals in Quantum Mechanics*, Oxford University Press, 2005.

