Institute for Theoretical Physics
Heidelberg University

Quantum Gravity and the Renormalisation Group

Lecturer:

Prof. Dr. Astrid Eichhorn

About these lecture notes

These lecture notes are a typeset version of the handwritten notes provided by Astrid Eichhorn for the course “Quantum Gravity and the Renormalisation Group”. The typesetting was completed by Nawder Stokes.

Last edited on: February 4, 2025

Please write to stokes@thphys.uni-heidelberg.de for any corrections or to report typographical errors.

Contents

Chapter 1

Motivation for Quantum Gravity

There are two main lines of argument to motivate why we need a quantum theory of gravity. The first relies on the observation that matter, which serves as a source for spacetime curvature in General Relativity, has quantum properties. The second is based on the fact that some solutions of the field equations of General Relativity are singular and geodesically incomplete. Those solutions (including, e.g., black holes) are relevant to describe astrophysical observations and thus General Relativity is insufficient to properly account for all properties of the gravitational interaction in our universe.

1.1 *Argument 1: Matter is quantum, so gravity should be as well*

Our starting point is the Einstein equations:

$$G_{\mu\nu} = 8\pi G\, T_{\mu\nu}\,, \tag{1.1}$$

in units where $c = 1$. (We will also mostly set $\hbar = 1$ in the following.) The LHS is the Einstein tensor $G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$ and the RHS contains Newton's gravitational constant G and the energy-momentum tensor $T_{\mu\nu}$. The energy-momentum tensor arises from the Standard-Model Lagrangian, which underlies a quantum field theory for matter. As a consequence, there is an inconsistency between the LHS and the RHS of the Einstein equations, which can be thought about in two complementary ways:

- The equations are mathematically inconsistent. The LHS is built from functions on spacetime whereas the RHS is constructed from operators.
- The equations are physically inconsistent. The RHS exhibits quantum uncertainty (e.g., position vs. momentum), but there is no corresponding uncertainty on the LHS.

Following the second point, we could ask ourselves, what is the gravitational field sourced by "fuzzy" distributions of matter? One possible answer may be semi-classical gravity.

Semi-classical gravity

The idea underlying semi-classical gravity is to make the Einstein equations consistent by inserting expectation values on the RHS. In other words, we use that $\langle T_{\mu\nu}\rangle$ is a function on spacetime. Hence, we can replace the Einstein equations with

$$G_{\mu\nu} = 8\pi G \langle T_{\mu\nu}\rangle \tag{1.2}$$

With this ansatz, we now have

- matter described by a quantum field theory on curved spacetime,
- a gravitational field sourced by the expectation value of energy-momentum tensor.

We may now ask, whether this is a sufficient theory to describe all possible situations with quantum-mechanical matter but classical spacetime consistently? To shed light on this, let us consider the following thought experiment.

Example: Semi-classical superposition of two masses

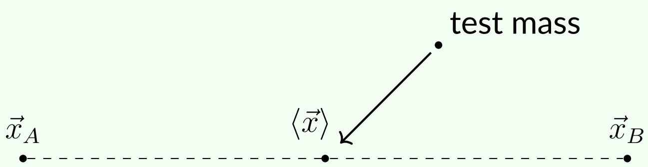
A massive, gravitating body is in a superposition of states centered about two locations, $\vec{x}_A$ and $\vec{x}_B$. The wavefunction of this body at some fixed time t is given by

$$|\psi(x,t)\rangle = \frac{1}{\sqrt{2}}\left(|\delta(\vec{x}-\vec{x}_A)\rangle + |\delta(\vec{x}-\vec{x}_B)\rangle\right) . \tag{1.3}$$

(Here, we are not concerned with the future evolution of this wavefunction, under which the delta-functions will broaden. As long as $\vec{x}_A$ and $\vec{x}_B$ are sufficiently separated from each other, it is for our purposes sufficient to write the wave-function in the above form.) The expected position is then

$$\langle \vec{x}\,\rangle = \frac{\vec{x}_A + \vec{x}_B}{2} . \tag{1.4}$$

By the semiclassical Einstein equations (1.2), the gravitational field is sourced at $\langle \vec{x}\,\rangle$, which is in-between the two masses:



A test mass was added to illustrate how the gravitational force acts on $\langle \vec{x}\,\rangle$, where it is sourced. Upon measurement, the wavefunction collapses to one of the two locations $\vec{x}_A$ or $\vec{x}_B$.
If we consider $\vec{x}_A$ and $\vec{x}_B$ sufficiently far separated from each other, a paradoxical situation arises: a gravitational field is sourced at a location $\frac{\vec{x}_A+\vec{x}_B}{2}$, far from either of the two locations at which the gravitating body may be found upon measurement. We do not expect that this is a situation that can actually be realized in nature, but rather that it arises because we have pushed the semi-classical theory beyond the regime of its validity.

To put such a thought experiment into practice has been attempted before (e.g. by Page and Geilker 1981, see however the critical discussion Ballentine 1982). The general challenge is to achieve a quantum-mechanical superposition for an object that is massive enough for its gravitational field to be measurable. Currently, new experimental efforts in this direction are being undertaken.

The above thought experiment illustrates that we expect that the semi-classical theory holds in regimes where quantum fluctuations on the RHS are actually small. In situations where this is not the case, we need a fully quantum treatment of the system. In analogy to the other fundamental forces in nature, we expect that the response of the gravitational field to a massive superposition as discussed above is to go into the corresponding superposition.

1.2 Argument 2: General Relativity signals its own breakdown

General Relativity signals its own breakdown by harboring curvature singularities and incomplete geodesics in physically relevant spacetimes, e.g. black hole spacetimes such as the Schwarzschild (or more generally the Kerr) spacetime. The line-element for the Schwarzschild spacetime in Schwarzschild coordinates is given by

$$\mathrm{d}s^2 = -\left(1 - \frac{2GM}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2GM}{r}\right)^{-1}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\,, \tag{1.5}$$

where $\mathrm{d}\Omega^2$ is the line-element of the two-sphere. This is a vacuum solution of General Relativity, i.e. a solution of

$$R = 0 \implies R_{\mu\nu} = 0\,, \tag{1.6}$$

which is obtained by multiplying the Einstein equations with the inverse metric $g^{\mu\nu}$. However, the fact that the Ricci tensor $R_{\mu\nu}$ vanishes, does not signify that curvature is zero, as the Riemann tensor $R_{\beta\mu\kappa\lambda} \neq 0$. In particular, we can build a curvature invariant, namely the Kretschmann scalar

$$K = R_{\beta\nu\kappa\lambda}R^{\beta\nu\kappa\lambda} = 48\frac{G^2M^2}{r^6}\,, \tag{1.7}$$

which diverges for $r \to 0$. This signifies infinite curvature at $r = 0$, suggesting infinite tidal forces, which is clearly an unphysical result. Also, if one calculates the proper time τ_0 it takes for a massive particle to reach $r = 0$ when starting from $R_{\text{max}} > 2GM$, one finds

$$\tau_0 = \frac{\pi R_{\text{max}}^{3/2}}{2^{3/2}\sqrt{M}}\,. \tag{1.8}$$

This expression is finite and, hence, signifies that geodesics terminate in finite proper time.
We conclude that General Relativity signals its own breakdown. A more complete theory does not necessarily need to be a quantum theory – it could also be a classical, modified theory of gravity – but it is a “minimal” assumption. Therefore, one typically requires of quantum gravity theories that they can resolve curvature singularities and render spacetimes geodesically complete.

Chapter 2

Perturbative quantization of gravity

Our goal is to quantize the gravitational field. To do so, we first consider a perturbative quantization of gravity. We can think of this as a quantization of gravitational waves, which are small fluctuations on top of a Minkowski background spacetime.

There are (at least) two reasons to start with this setting. First, perturbative quantization is successful for the other fundamental forces that we know of and that make up the Standard Model of particle physics. Thus, it is well motivated to first test whether gravity can be quantized within the same formalism. Second, gravity is a priori different than other fields, because other fields exist *on top* of a spacetime geometry, whereas the metric field *determines* the spacetime geometry. By quantizing gravity perturbatively, we actually quantize small fluctuations on top of a fixed background spacetime geometry and thus we can then use the same formalism as for the other fields.

As a result of perturbative quantization, we will encounter gravitons, which are the analogue of photons, i.e., they are the (massless spin-two) quanta of the gravitational field.

2.1 *Recap of gravitational waves*

The Einstein field equations[i]

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu}\,, \tag{2.1}$$

here with the Einstein tensor written out, are highly non-linear. The non-linearity becomes clear when we take into account that the Ricci scalar R contains the inverse metric, because $R = g^{\mu\nu} R_{\mu\nu}$. In addition, the Ricci tensor also contains the inverse metric within the Christoffel symbols that it is constructed from. One can think of these non-linearities as self-interactions of the metric (and

[i]There are several, classically equivalent formulations of General Relativity that use different fields, e.g., unimodular gravity, Palatini gravity and others. Classical equivalence is not sufficient to guarantee equivalence at the quantum level, because classical equivalence means that the solutions to the equations of motion (i.e., the *on-shell* configurations) are the same. However, in a quantum theory, the *off-shell* configurations also enter observable quantities (through loop contributions). We will start from the classical formulation in terms of the metric and the Christoffel symbol for the perturbative quantization.

upon quantization, the gravitons). To quantize perturbatively, we consider only the linear theory, i.e., we neglect self-interactions. To linearize, we expand in perturbations $h_{\mu\nu}(x)$ around a background geometry $\bar{g}_{\mu\nu}$. Here, we expand about a flat background $\bar{g}_{\mu\nu} = \eta_{\mu\nu}$, i.e. we set

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \,. \tag{2.2}$$

This is a good approximation as long as the perturbation is small, i.e. $|h_{\mu\nu}| \ll 1$. A small perturbation means that the gravitational field created by the perturbation $h_{\mu\nu}$ is small enough or, in other words, self-interactions of the gravitational field are negligible.

In the next step, we derive the linearization of R and $R_{\mu\nu}$ to obtain the linearized Einstein equations. For this step, we need to know that

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} + \mathcal{O}(h^2) \,. \tag{2.3}$$

This can be checked by requiring that $g_{\mu\nu}g^{\nu\kappa} = \delta^\kappa_\mu$, i.e., $g^{\mu\nu}$ is the inverse of $g_{\mu\nu}$. From this, we can derive that

$$\Gamma_\beta{}^\lambda{}_\nu = \frac{1}{2}\eta^{\lambda\tau}\left(\partial_\beta h_{\tau\nu} + \partial_\nu h_{\tau\rho} - \partial_\tau h_{\beta\nu}\right) + \mathcal{O}(h^2) \,, \tag{2.4}$$

$$R_{\mu\nu\rho\sigma} = \frac{1}{2}\left(\partial_\beta\partial_\sigma h_{\nu\rho} - \partial_\beta\partial_\rho h_{\nu\sigma} - \partial_\nu\partial_\sigma h_{\beta\rho} + \partial_\nu\partial_\rho h_{\beta\sigma}\right) + \mathcal{O}(h^2) \,, \tag{2.5}$$

$$R_{\beta\nu} = R^\alpha{}_{\beta\alpha\nu} = \frac{1}{2}\left(\partial_\alpha\partial_\nu h_\beta{}^\alpha - \partial^2 h_{\beta\nu} - \partial_\beta\partial_\nu h + \partial_\beta\partial_\alpha h^\alpha{}_\nu\right) + \mathcal{O}(h^2) \,, \tag{2.6}$$

where $\mathcal{O}(h^2)$ contain terms that are quadratic in the perturbation and its derivatives, $\partial^2 := \partial^\alpha\partial_\alpha$ is the d'Alembert operator and $h := h^\alpha{}_\alpha$ is the trace of the perturbation metric. Using the expression for the Ricci tensor, one finds for the Ricci scalar

$$R = g^{\beta\nu}R_{\beta\nu} = -h^{\beta\nu}R_{\beta\nu} + \partial_\alpha\partial_\beta h^{\beta\alpha} - \partial^2 h = \partial_\alpha\partial_\beta h^{\beta\alpha} - \partial^2 h \,, \tag{2.7}$$

where the first term vanishes on a flat background. Note that if we had expanded the above expressions to higher order, we would obtain **interaction** terms of $h_{\mu\nu}$. These become important in the strong-field regime, where gravitational waves are generated, e.g. through mergers of black holes. In the **propagation** of gravitational waves far away from the source, because their amplitude is so small, the interactions are negligible.

With the above expressions, we find the linearized Einstein equations

$$\partial^2 h_{\mu\nu} - \left(\partial_\mu\partial^\rho h_{\rho\nu} + \partial_\nu\partial^\rho h_{\rho\mu}\right) + \partial_\mu\partial_\nu h + \eta_{\mu\nu}\partial_\rho\partial_\sigma h^{\rho\sigma} - \eta_{\mu\nu}\partial^2 h = -16\pi G T_{\mu\nu} \,. \tag{2.8}$$

This can be written in a more compact form by defining

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h \,. \tag{2.9}$$

With this, the linearized Einstein equations are

$$\partial^2 \bar{h}_{\mu\nu} - \left(\partial_\mu\partial^\rho \bar{h}_{\rho\nu} + \partial_\nu\partial^\rho \bar{h}_{\rho\mu}\right) + \eta_{\mu\nu}\partial_\rho\partial_\sigma \bar{h}^{\rho\sigma} = -16\pi G T_{\mu\nu} \,. \tag{2.10}$$

This already resembles a wave equation for $\bar{h}_{\mu\nu}$, but there are also extra terms. In the following, we will find out how to remove these terms using a symmetry of the equations.

We make the following observation concerning the linearized Einstein equations: they have an infinite-dimensional kernel consisting of fields of the form

$$h_{\mu\nu} = \partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu \,, \tag{2.11}$$

or

$$\bar{h}_{\mu\nu} = \partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu - \eta_{\mu\nu} \partial_\lambda \epsilon^\lambda \,. \tag{2.12}$$

This means that if $\bar{h}_{\mu\nu}{}^{(1)}$ is a solution of the linearized Einstein equation, then

$$\bar{h}_{\mu\nu}{}^{(2)} = \bar{h}_{\mu\nu}{}^{(1)} + \partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu - \eta_{\mu\nu} \partial_\lambda \epsilon^\lambda \,, \tag{2.13}$$

is also a solution. Thus, not all components of $\bar{h}_{\mu\nu}$ are **physical**. The cause of this redundancy is the diffeomorphism invariance of the gravitational action. To see this, consider the transformation

$$x^\alpha \to x^\alpha + \epsilon^\alpha = x'^\alpha \,, \tag{2.14}$$

where ϵ^μ is some vector with small magnitude. Then the metric transforms as

$$\begin{aligned} g_{\mu\nu} = \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} g'_{\alpha\beta} \iff \eta_{\mu\nu} + h_{\mu\nu} &= \left(\frac{\partial \epsilon^\alpha}{\partial x^\mu} + \delta^\alpha{}_\mu \right) \left(\frac{\partial \epsilon^\beta}{\partial x^\nu} + \delta^\beta{}_\nu \right) \left(\eta_{\alpha\beta} + h'_{\alpha\beta} \right) \\ \iff h_{\mu\nu} &= h'_{\mu\nu} + \partial_\mu \epsilon_\nu + \partial_\nu \epsilon_\mu + \mathcal{O}(\epsilon^2) \,. \end{aligned} \tag{2.15}$$

Just as in gauge theories, this is an **unphysical** symmetry in the sense that it is really an over-parameterization of the physical degrees of freedom: we are introducing more components of $h_{\mu\nu}$ than there are propagating degrees of freedom. The source of this symmetry is that we have introduced unphysical quantities – in this case, coordinates. This implies that $h_{\mu\nu}$ and $h'_{\mu\nu}$ are **physically equivalent**. The difference between them cannot be probed by any experiment.

To get rid of this unphysical redundancy, we **gauge-fix** by imposing an additional condition on $\bar{h}_{\mu\nu}$. A gauge condition removes unphysical degrees of freedom, i.e., gauge degrees of freedom. We will choose a condition that also simplifies the equations of motion. To keep the term $\partial^2 \bar{h}_{\mu\nu}$, but remove the terms $\sim \partial_\mu \bar{h}^{\mu\nu}$, we would like to choose de Donder gauge

$$\partial_\mu \bar{h}^{\mu\nu} = 0 \iff \partial_\mu h^{\mu\nu} = \frac{1}{2} \partial^\nu h \,. \tag{2.16}$$

However, we can of course not simply demand a condition, but need to check that the condition can be fulfilled by an appropriate choice of ϵ^μ. Otherwise, the condition would actually be a condition on the physical degrees of freedom and not just on the gauge degrees of freedom. To check that de Donder gauge is a viable gauge condition, we proceed as follows: given an $h_{\mu\nu}$ which does not satisfy the de Donder gauge condition $\partial_\mu \bar{h}^{\mu\nu} = 0$, we search for an ϵ_μ such that $h'_{\mu\nu} = h_{\mu\nu} + \partial_\mu\epsilon_\nu + \partial_\mu\epsilon_\mu$ fulfills it. We may write

$$\begin{aligned} 0 = \partial_\mu \bar{h}'^{\mu\nu} &= \partial_\mu \left(\bar{h}^{\mu\nu} + \partial^\mu\epsilon^\nu + \partial^\nu\epsilon^\mu - \eta^{\mu\nu}\partial_\lambda\epsilon^\lambda\right) \\ &= \partial_\mu \bar{h}^{\mu\nu} + \partial_\mu\partial^\mu\epsilon^\nu + \partial_\mu\partial^\nu\epsilon^\mu - \eta^{\mu\nu}\partial_\mu\partial_\lambda\epsilon^\lambda \\ &= \partial_\mu \bar{h}^{\mu\nu} + \partial^2\epsilon^\nu\,, \end{aligned} \tag{2.17}$$

by first inserting the expression (2.12) for the transformation of $\bar{h}'_{\mu\nu}$ under infinitesimal coordinate change. Hence, we find that $\epsilon_\nu(x)$ must satisfy

$$\partial_\mu \bar{h}^{\mu\nu} + \partial^2\epsilon^\nu = 0\,, \tag{2.18}$$

which always admits a solution. Thus, de Donder gauge is a viable gauge condition to impose. In fact, the solution is only determined up to solutions of the homogeneous equation

$$\partial^2\epsilon^\nu = 0\,. \tag{2.19}$$

This implies that the de Donder gauge condition does not completely fix the gauge freedom and there is a some residual gauge freedom left. We will make use of this freedom later.

Linearized Einstein equations in de Donder gauge

In the de Donder gauge condition $\partial_\mu \bar{h}_{\mu\nu} = 0$, where $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$, the linearized Einstein field equations are

$$\partial^2 \bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu}\,. \tag{2.20}$$

This is a wave equation for $\bar{h}_{\mu\nu}$ with a source term.

To arrive at the field modes that we want to quantize, we must solve (2.20) in the vacuum. We start by first proving the following

Claim: $\partial^2 \bar{h}_{\mu\nu} = 0 \iff \partial^2 h_{\mu\nu} = 0$.

Proof. $\Longrightarrow$: Assuming that $\bar{h}_{\mu\nu}$ satisfies the de Donder gauge condition, we can write

$$0 = \partial^2 \bar{h}_{\mu\nu} = \partial^2 h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\partial^2 h\,. \tag{2.21}$$

Taking the trace of this equation yields

$$0 = \partial^2 h\,. \tag{2.22}$$

Inserting this back into the first equation leads to the desired result.
$\Longleftarrow$: Assuming now that $h_{\mu\nu}$ satisfies the de Donder gauge condition, we can write

$$\partial^2 h_{\mu\nu} = 0\,. \tag{2.23}$$

By taking the trace, we find

$$\partial^2 h = 0 \implies -\frac{1}{2}\eta_{\mu\nu}\partial^2 h = 0\,, \tag{2.24}$$

and thus that we can add this term to $\partial^2 h_{\mu\nu}$ above without changing the RHS, given us the desired result.

□

We can thus concentrate on solving the equation $\partial^2 h_{\mu\nu} = 0$. This equation has a general solution that can be written as a superposition of plane waves:

$$h_{\mu\nu}(x) = \int \mathrm{d}^3 p\ \left[\tilde{h}_{\mu\nu}(p)e^{ip_\alpha x^\alpha} + \tilde{h}^*_{\mu\nu}(p)e^{-ip_\alpha x^\alpha}\right] \qquad \text{and} \qquad p^2 = 0\,. \tag{2.25}$$

We will consider a single Fourier mode

$$h_{\mu\nu}(x) = \Pi_{\mu\nu}e^{ip_\alpha x^\alpha} + \Pi^*_{\mu\nu}e^{-ip_\alpha x^\alpha}\,, \tag{2.26}$$

where $\Pi_{\mu\nu}$ is the polarisation tensor and the star denotes complex conjugation. Note also that to be a solution of the wave-equation, $p^2 = 0$ has to hold, i.e. gravitons are massless in General Relativity. We will now count the degrees of freedom in the polarisation tensor in order to arrive at the physical degrees of freedom that we would like to quantize. The gauge condition requires the four conditions

$$p_\mu \bar{\Pi}^{\mu\nu} := p_\mu \left(\Pi^{\mu\nu} - \frac{1}{2}\Pi\eta^{\mu\nu}\right) = 0\,. \tag{2.27}$$

We now use the residual gauge freedom to impose further conditions on $\bar{\Pi}^{\mu\nu}$. We propose that we can impose the condition

$$U^\mu \bar{\Pi}_{\mu\nu} = 0. \tag{2.28}$$

As before, we have to check that this is a viable condition to impose, i.e., that there is always a choice of ϵ such that the condition can be achieved. We focus on the Fourier mode of $\epsilon^\mu(x)$ with the same momentum p_μ as the solution $h_{\mu\nu}(x)$ we are considering,

$$\epsilon_\mu(x) = \tilde{\epsilon}_\mu e^{ip_\alpha x^\alpha} + \tilde{\epsilon}_\mu e^{-ip_\alpha x^\alpha}\,. \tag{2.29}$$

Under the gauge transformation (2.12), we know

$$\bar{\Pi}'_{\mu\nu} = \bar{\Pi}_{\mu\nu} + i\left(p_\mu\epsilon_\nu + p_\nu\epsilon_\mu - \eta_{\mu\nu}p^\lambda\epsilon_\lambda\right)\,. \tag{2.30}$$

Dotting in a constant vector U^μ into both sides of this equation and setting the RHS equal to zero leads to

$$U^\mu\bar{\Pi}'_{\mu\nu} = U^\mu\bar{\Pi}_{\mu\nu} + i\left(U^\mu p_\mu\epsilon_\nu + p_\nu U^\mu\epsilon_\mu - U_\nu p^\lambda\epsilon_\lambda\right) = 0\,. \tag{2.31}$$

These are four coupled linear equations for the four components of ϵ_μ; they always have a (complex) solution. Hence, we can impose the condition Eq. (2.28).
Note that these are actually only three independent equations, because $p^\mu\bar{\Pi}_{\mu\nu}U^\nu \equiv 0$, i.e. one

superposition of the four equations $U^\mu \bar{\Pi}_{\mu\nu}$ is zero anyway. This means we can impose one more condition. We observe that the trace of the momentum space infinitesimal transformation (2.30) is

$$\bar{\Pi}'^{\mu}{}_{\mu} = \bar{\Pi}^{\mu}{}_{\mu} + i(p^\mu \epsilon_\mu \cdot 2 - 4p^\lambda \epsilon_\lambda) = \bar{\Pi}^{\mu}{}_{\mu} - 2ip^\lambda \epsilon_\lambda \,. \tag{2.32}$$

This means that we can choose ϵ_μ such that $\bar{\Pi}_{\mu\nu}$ is a traceless tensor. Note that this condition implies that

$$\bar{\Pi}_{\mu\nu} = \Pi_{\mu\nu} \,. \tag{2.33}$$

We can now summarize the conditions on the polarization tensor. These are:

- the polarisation tensor is traceless: $\Pi^\mu{}_\mu = 0$ (one condition),
- the polarisation tensor is transverse to the direction of propagation $p^\mu \Pi_{\mu\nu} = 0$ (four conditions),
- the polarisation tensor is orthogonal to some arbitrary vector: $U^\mu \Pi_{\mu\nu} = 0$ (three conditions).

Of the 10 independent components of the symmetric tensor $\Pi_{\mu\nu}$, two are physical, i.e., the two helicities (spin aligned and anti-aligned with direction of propagation) of a massless spin-two graviton propagate.

Let us briefly generalize to d spacetime dimensions and count more generally. The polarisation tensor $\Pi_{\mu\nu}$ is a symmetric tensor and thus has $d(d+1)/2$ independent components. We have imposed $d + (d-1) + 1$ conditions from transversality, the orthogonality to U^μ and tracelessness, respectively. Hence, for the degrees of freedom of the polarisation, we get

$$\text{dof}(\Pi) = \frac{d(d+1)}{2} - d - (d-1) - 1 = \frac{d^2+d}{2} - 2d = \frac{d(d-3)}{2} \,. \tag{2.34}$$

From this expression, we can see that $d = 4$ is special, because it is the lowest spacetime dimensionality with propagating gravitational waves. We see that for $d \leq 3$, the number of degrees of freedom in the polarisation tensor is zero or negative. The reason is that, as we will see, gravitational waves (unlike electromagnetic waves) oscillate in *two* directions which are perpendicular to the direction of propagation, not just one. In less than four dimensions, there are no two orthogonal spatial dimensions in which the gravitational wave can oscillate in. Hence, we conclude that $d = 4$ dimensions is the minimal spacetime dimensionality in which gravitational waves can propagate. At this stage, one could ask oneself: why is $d = 4$ the spacetime dimensionality of our universe?

Example: Gravitational wave propagating in z-direction

We want to figure out what the two polarisation states look like in $d = 4$. Let us choose propagation in z-direction, so that $p^\mu = (p, 0, 0, p)$. We choose $U^\mu = (1, 0, 0, 0)$. With these choices, we find

$$\Pi_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & e_+ & e_\times & 0 \\ 0 & e_\times & -e_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{2.35}$$

where e_+ and $e_\times$ are the amplitudes of the two polarisation states. For the case of $e_\times = 0$, the infinitesimal line element is given by

$$\begin{aligned} \mathrm{d}s^2 &= \mathrm{d}x^\mu \mathrm{d}x^\nu \left(\eta_{\mu\nu} + h_{\mu\nu}\right) \\ &= -\mathrm{d}t^2 + \mathrm{d}x^2 \left[1 + e_+ \left(e^{ip(z-t)} + e^{-ip(z-t)}\right)\right] + \mathrm{d}y^2 \left[1 - e_+ \left(e^{ip(z-t)} + e^{-ip(z-t)}\right)\right] + \mathrm{d}z^2 \\ &= -\mathrm{d}t^2 + \mathrm{d}x^2 \left[1 + 2e_+ \cos(p(z-t))\right] + \mathrm{d}y^2 \left[1 - 2e_+ \cos(p(z-t))\right] + \mathrm{d}z^2 \,. \end{aligned} \tag{2.36}$$

We can visualize the effect that this $+$-polarized gravitational wave has on a ring of freely falling particles positioned in the $x-y$-plane (see Figure 2.1).

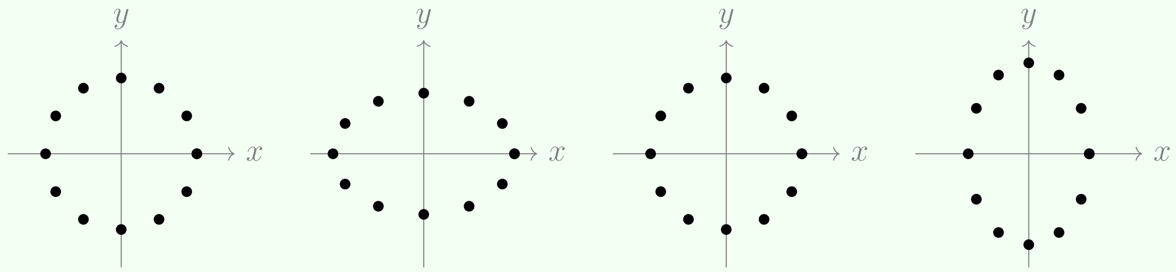


Figure 2.1: *The effect a $+$-polarized gravitational wave propagating in z-direction has on a ring of freely falling particles placed in the $x-y$-plane. Each plot displays a snapshot in time.*

2.2 Quantization of the non-interacting gravitational field

We start from plane-wave solutions and promote coefficients of the expansion to creation/annihilation operators. Hence, we write

$$h_{\mu\nu}(x) = \sum_{\sigma=\mathrm{L,R}} \int \mathrm{d}^3p \left[a_{\vec{p},\sigma} \Pi_{\mu\nu}(\sigma) e^{ip_\alpha x^\alpha} + a^\dagger_{\vec{p},\sigma} \Pi_{\mu\nu}(\sigma) e^{-ip_\alpha x^\alpha}\right], \tag{2.37}$$

where we sum over the two polarisation states

$$e_L = \frac{1}{\sqrt{2}}\left(e_+ - ie_\times\right), \qquad e_R = \frac{1}{\sqrt{2}}\left(e_+ + ie_\times\right). \tag{2.38}$$

We follow the standard rule for quantization and promote $a_{\vec{p},\sigma}$, $a^\dagger_{\vec{p},\sigma}$ to quantum-mechanical operators that satisfy the canonical commutation relations

$$[a_{\vec{p},\sigma}, a_{\vec{p}',\sigma'}] = 0\,, \qquad [a^\dagger_{\vec{p},\sigma}, a^\dagger_{\vec{p}',\sigma'}] = 0\,, \qquad [a_{\vec{p},\sigma}, a^\dagger_{\vec{p}',\sigma'}] = i\delta_{\sigma\sigma'}\delta^{(3)}(\vec{p} - \vec{p})\,. \tag{2.39}$$

Our whole discussion of gravitational waves had the point to find how many independent set of creation operators $a^\dagger$ and annihilation operators a there are, i.e., what set of particles gravitons are. We found that there are two degrees of freedom, the positive and negative helicity of a massless spin-two field. From the canonical commutation relations, we can build a Fock space for non-interacting gravitons in the usual way:

- define the vacuum $|0\rangle$ as the unique state annihilated by all $a_{\vec{p},\sigma}$, i.e. $a_{\vec{p},\sigma}\,|0\rangle = 0\,, \quad \forall \vec{p}, \sigma$,
- define the single-graviton state as $a^\dagger_{\vec{p},\sigma}\,|0\rangle = |\vec{p},\sigma\rangle$,
- etc.

We have now quantized the free theory; the next step is to add interactions.

2.3 Loop corrections in perturbative quantum gravity

To calculate loop corrections, we need to know the **Feynman rules**. We start again from the Einstein-Hilbert Lagrangian and this time expand to higher order in the perturbation field $h_{\mu\nu}$, because we want to access the propagator and the interaction vertices. We also impose de Donder gauge. Doing so, one finds

$$\mathcal{L} = -\frac{1}{4}\partial_\mu h_{\alpha\beta}\partial^\mu h^{\alpha\beta} + \frac{1}{8}(\partial_\beta h)^2 + h_{\mu\nu}U^{\mu\nu\kappa\lambda\rho\sigma}h_{\kappa\lambda}h_{\rho\sigma} + \mathcal{O}(h^4)\,. \tag{2.40}$$

From the first two terms, we can extract the propagator and from the third term, we can extract the graviton three-point vertex. Note that when imposing the gauge condition, we have to do so via the Faddeev-Popov procedure, which produces a ghost term in the Lagrangian. For now we neglect it, because it does not change the power-counting of loop divergences here. We will include it in our treatment later on and it must of course be included if we want to obtain the correct numerical prefactor of a given loop diagram. Because we only care about the divergences in loop diagrams for now, this numerical prefactor will not matter for us.

Now we derive the Feynman rules starting with the propagator. The inverse propagator in momentum space is given by

$$\mathcal{P}^{-1}_{\alpha\beta\mu\nu} = \frac{p^2}{2}\eta_{\alpha\mu}\eta_{\beta\nu} - \frac{p^2}{4}\eta_{\alpha\beta}\eta_{\mu\nu}\,. \tag{2.41}$$

The propagator must satisfy $\mathcal{P}^{-1}\mathcal{P} = \mathbb{I}$, where $\mathbb{I}$ is the unit element in the space of symmetric tensors, i.e.

$$\mathcal{P}^{-1}_{\alpha\beta\mu\nu}\mathcal{P}^{\mu\nu\kappa\lambda} = \frac{1}{2}\left(\delta_\alpha{}^\kappa\delta_\beta{}^\lambda + \delta_\alpha{}^\lambda\delta_\beta{}^\kappa\right)\,. \tag{2.42}$$

Exercise: Derivation of the propagator of the gravitational field

Given the above specifications, one can make the ansatz

$$\mathcal{P}^{\mu\nu\kappa\lambda} = \frac{1}{p^2}\left(c_1\eta^{\mu\kappa}\eta^{\nu\lambda} + c_2\eta^{\mu\lambda}\eta^{\nu\kappa} + c_3\eta^{\mu\nu}\eta^{\kappa\lambda}\right) , \tag{2.43}$$

for three to be determined constant coefficients c_1, c_2 and c_3. This ansatz is motivated by the following consideration: We can guess from Eq. (2.42) that the propagator must go like $1/p^2$. The four indices must be carried by tensors due to Lorentz covariance, and the only tensors we have available are $\eta_{\mu\nu}$ and $\frac{p_\mu p_\nu}{p^2}$. Our ansatz only uses a subset of those, namely only those constructed from the metric. It is a "bonus" of the de Donder gauge condition that the propagator only needs these terms; the graviton propagator in other gauges generically requires more terms, namely the three extra terms that can be constructed from $\frac{p_\mu p_\nu}{p^2}$ and $\eta_{\mu\nu}$, with the indices appropriately symmetrized. Inserting this ansatz and (2.41) into the condition (2.42), we get

$$\begin{aligned}
&\frac{p^2}{4}\left(2\eta_{\alpha\mu}\eta_{\beta\nu} - \eta_{\alpha\beta}\eta_{\mu\nu}\right)\frac{1}{p^2}\left(c_1\eta^{\mu\kappa}\eta^{\nu\lambda} + c_2\eta^{\mu\lambda}\eta^{\nu\kappa} + c_3\eta^{\mu\nu}\eta^{\kappa\lambda}\right) = \frac{1}{2}\left(\delta_\alpha{}^\kappa\delta_\beta{}^\lambda + \delta_\alpha{}^\lambda\delta_\beta{}^\kappa\right)\\
\Longleftrightarrow\ & c_1\left(2\delta_\alpha{}^\kappa\delta_\beta{}^\lambda - \eta_{\alpha\beta}\eta^{\kappa\lambda}\right) + c_2\left(2\delta_\alpha{}^\lambda\delta_\beta{}^\kappa - \eta_{\alpha\beta}\eta^{\lambda\kappa}\right)\\
&+ c_3\left(2\eta_{\alpha\beta}\eta^{\kappa\lambda} - \eta_{\alpha\beta}4\eta^{\kappa\lambda}\right) = 2\left(\delta_\alpha{}^\kappa\delta_\beta{}^\lambda + \delta_\alpha{}^\lambda\delta_\beta{}^\kappa\right) .
\end{aligned} \tag{2.44}$$

We notice that for the relevant terms on the LHS to match with the RHS, we need to set $c_1 = 1 = c_2$. Doing so simplifies the equation for c_3 to the following:

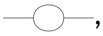
$$\begin{aligned}
&-\eta_{\alpha\beta}\eta^{\kappa\lambda} - \eta_{\alpha\beta}\eta^{\lambda\kappa} + c_3\left(2\eta_{\alpha\beta}\eta^{\kappa\lambda} - 4\eta_{\alpha\beta}\eta^{\kappa\lambda}\right) = 0\\
\Longleftrightarrow\ & -2\eta_{\alpha\beta}\eta^{\kappa\lambda} - 2c_3\eta_{\alpha\beta}\eta^{\kappa\lambda} = 0
\end{aligned} \tag{2.45}$$

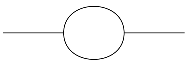
Hence, we can see that we need to set $c_3 = -1$ such that this equation holds.

Feynman propagator of the gravitational field

The Feynman propagator of the gravitational field in de Donder gauge is given by

$$\mathcal{P}_{\mu\nu\kappa\lambda} = \frac{1}{p^2}\left(\eta_{\mu\kappa}\eta_{\nu\lambda} + \eta_{\mu\lambda}\eta_{\nu\kappa} - \eta_{\mu\nu}\eta_{\kappa\lambda}\right) . \tag{2.46}$$

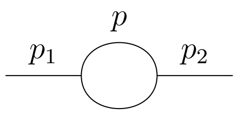
The graviton three-point vertex is lengthy, with 18 terms with six indices each when we use maximally compact notation. Each term is quadratic in momentum; and depends on two out of the three momenta $p_1, p_2, -p_1-p_2$ of the three gravitons. (We already used momentum conservation at the vertex to write that the third graviton has momentum $p_3 = -p_1 - p_2$, if all momenta are defined with respect to ingoing graviton lines.) With these ingredients, we can calculate loop diagrams such as , a one-loop contribution to the graviton propagator. We encounter divergences when performing the loop integral. They organize themselves into terms of the form



$$= (\text{divergent prefactor}) \cdot \left(\partial^2 h \partial^2 h - 2\partial^2 h \partial_\alpha \partial_\beta h^{\alpha\beta} + \partial_\alpha \partial_\beta h^{\alpha\beta} \partial_\mu \partial_\nu h^{\mu\nu}\right) + \text{other terms}\,, \tag{2.47}$$

where we recognize the second term as the expression of R^2 to second order in $h_{\mu\nu}$.
How do these terms arise and what do they imply? To understand their origin, we consider the following two points:

- By plugging in the expressions for the propagator and the vertex explicitly, we arrive at an expression that schematically contains three different terms. These depend on whether the two momenta at each vertex are i) both external momenta, ii) one internal (loop) momentum and one external momentum or iii) both internal (loop) momenta. The above expression arises if all momenta at the vertex are external momenta. We then get an expression of the schematic form,



$$\sim \int \mathrm{d}^4 p \left(\frac{1}{p^2}\right)^2 \cdot f(p_1, p_2)\,, \tag{2.48}$$

  where the first term comes from the two propagators and the second term is a function that is $\mathcal{O}(p_1^2, p_2^2)$, as both vertices are quadratic in momenta.

- The resulting terms must be linearisations of curvature invariants, otherwise coordinate invariance would be broken. At fourth order in derivatives of the perturbation metric, candidate terms are R^2, $\Box R$, $R_{\mu\nu}R^{\mu\nu}$, $R_{\mu\nu\kappa\lambda}R^{\mu\nu\kappa\lambda}$. (Note that $\Box R$ is a total derivative.) In $d = 4$, the Gauss-Bonnet invariant

$$E = R_{\mu\nu\kappa\lambda}R^{\mu\nu\kappa\lambda} - 4R_{\mu\nu}R^{\mu\nu} + R^2\,, \tag{2.49}$$

  is also a total derivative. Hence, only two of the above listed fourth order curvature invariants contribute to the action (in the absence of boundary terms), namely R^2 and $R_{\mu\nu}R^{\mu\nu}$. Therefore, the prefactor of the divergence of the above diagram must be a combination of R^2 and $R_{\mu\nu}R^{\mu\nu}$, linearized to second order in $h_{\mu\nu}$.

To absorb the divergences of the one-loop amplitude, counterterms have to be added to the action. As a consequence of the second point we just made, we know that the counterterms are

$$a \int \mathrm{d}^4 x \sqrt{-g} R^2\,, \qquad \text{and} \qquad b \int \mathrm{d}^4 x \sqrt{-g} R_{\mu\nu}R^{\mu\nu}\,, \tag{2.50}$$

which are **not** of the form of the original action. Hence, two new couplings appear in the action.[ii]

[ii]In the absence of matter, one can perform a field redefinition

$$g_{\mu\nu} \to g_{\mu\nu} + \alpha R g_{\mu\nu} - \beta R_{\mu\nu}\,. \tag{2.51}$$

For an appropriate choice of α and β, the counterterms vanish.

This pattern persists at n^{th}-loop order, where the required counterterms have $2n+2$-derivatives and are thus (curvature)$^{n+1}$-expressions. The consequence is that the perturbative quantization of the Einstein-Hilbert action leads to a non-predictive theory, which has infinitely many couplings (the finite prefactors of the new terms, once divergences have been absorbed). This is the **perturbative non-renormalisability** or simply a breakdown of predictivity.

To support this claim of $2n+2$-derivatives in the counterterms, let us calculate the superficial degree of divergence. This is the expected divergence of a diagram, barring any cancellations (e.g. from a contraction of propagator indices with vertex indices). To count the divergence, we introduce a cut-off in the momentum integral, which we call Λ_{UV}, i.e. the loop integral $\int \mathrm{d}^4p$ becomes $\int \mathrm{d}^4p = 2\pi^2 \int_0^\infty dp^2p^2 \rightarrow 2\pi^2 \int_0^{\Lambda_{\text{UV}}^2} dp^2p^2$. In a diagram with P propagators, L loops and V vertices, we will have

$$D = -2P + dL + 2V\,, \tag{2.52}$$

where D denotes the superficial degree of divergence, i.e. the power of Λ_{UV} when the momentum integral is performed. Upon inserting the topological relation $P = L + V - 1$, we obtain

$$D = (d-2)L + 2. \tag{2.53}$$

From this equation, we note that $d = 2$ is special, as there the superficial degree of divergence is not dependent on the loop order L. This is because gravity is **topological** in two dimensions, i.e. the $d = 2$ analogue of the Einstein-Hilbert action does not give rise to local equations of motion, but instead

$$\chi = \int \mathrm{d}^2x\sqrt{-g}R\,, \tag{2.54}$$

where χ is the Euler character, which is a topological invariant. χ counts the number h of handles that the manifold has, $\chi = 2 - 2h$.

Perturbative non-renormalisability of General Relativity

The superficial degree of divergence D of perturbatively quantized General Relativity is

$$D = (d-2)L + 2\,, \tag{2.55}$$

where L is the number of loops in a Feynman diagram and d the spacetime dimension. In four dimensions, the superficial degree of divergence increases with the loop order. At each loop order, it is always the logarithmically divergent terms (i.e. the lowest order of the divergence), which produces new counterterms. This indicates that General Relativity is perturbatively non-renormalisable.

Despite this, one can still ask: can we use perturbative Quantum Gravity in some way to extract predictions from quantum gravity? The answer is yes. We can treat the theory as an Effective Field Theory (EFT).

2.4 Perturbative Quantum Gravity as an Effective Field Theory

If we add all infinitely many counterterms to our action, the theory is renormalisable. In regimes where the spacetime curvature radius is of the order of the Planck length, the theory is not predictive, because there physical quantities depend on the unknown (and infinitely many) higher-order couplings. However, at low enough curvature (compared to Planckian scales), the extra terms are negligible and we can make predictions. In other words, the rationale for an Effective Field Theory (EFT) is to accept that we do not know physics up to arbitrarily high momentum cut-off Λ_{UV} and write an action with all possible terms compatible with the symmetry of the theory (in our case, diffeomorphism invariance) and organize them by mass-dimensionality. In practice, we write

$$S = \int \mathrm{d}^4x\sqrt{-g}\Big(\underbrace{-\frac{1}{2}m_p^2 R}_{(1)} + \underbrace{\alpha R^2 + \beta R_{\mu\nu}R^{\mu\nu}}_{(2)} + \ldots\Big)\,. \tag{2.56}$$

To see that in some regimes the higher-order terms are negligible, consider scattering of gravitons described by the action (2.56). If gravitons have energy E in this process, the term (1) contributes at $\mathcal{O}(E^2)$ and the terms (2) at $\mathcal{O}(E^4)$. To write the comparison between the two parts in dimensionless form, we pull out a factor m_p^4 from the Lagrangian. Schematically, we find

$$\mathcal{L} \sim m_p^4\left(-\frac{1}{2}\left(\frac{E}{m_p}\right)^2 + \alpha\left(\frac{E}{m_p}\right)^4 + \beta\left(\frac{E}{m_p}\right)^4 + \ldots\right)\,, \tag{2.57}$$

and thus conclude that, if $\alpha \sim \mathcal{O}(1)$ and $\beta \sim \mathcal{O}(1)$, the contribution from the higher-order part of (2.56) is negligible for $E \ll m_p^2$. If α and β are large, this enhances the contribution from the terms (2) and then the cut-off scale, at which the term (1) on its own is no longer a good approximation, is lowered.

Example: Classical Newtonian potential from Perturbative Quantum Gravity

We start from the tree-level diagram for the interaction between two massive non-relativistic scalar particles, shown in Fig. 2.2.

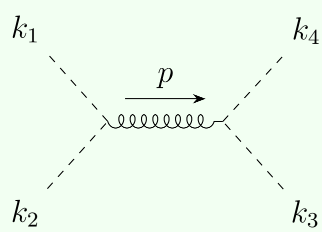


Figure 2.2: *Tree-level Feynman diagram for two massive scalar particles interacting via graviton exchange.*

The graviton propagator for a graviton of momentum q^μ in de Donder gauge is

$$\mathcal{D}_{\mu\nu\kappa\lambda}(-q^2) = \frac{1}{-q^2}\left(\eta_{\mu\kappa}\eta_{\nu\lambda} + \eta_{\mu\lambda}\eta_{\nu\kappa} - \eta_{\mu\nu}\eta_{\kappa\lambda}\right) , \tag{2.58}$$

as derived above and defined in (2.42). The vertex comes from the scalar action

$$S_\phi[g,\phi] = \int \mathrm{d}^4x\sqrt{-g}\left(-\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2\right) . \tag{2.59}$$

From there, we can derive the vertex involving one graviton and two scalar particles of momenta k_1^μ and k_2^μ,

$$V^{\mu\nu}(k_1,k_2) = 2\sqrt{8\pi G}\left[k_1^{(\mu}k_2^{\nu)} - \frac{1}{2}\eta^{\mu\nu}\left(k_1\cdot k_2 + m^2\right)\right] . \tag{2.60}$$

This is derived by expanding the scalar action around flat spacetime $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ and only considering the terms linear in $h_{\mu\nu}$, as we are only interested in the interaction between a single graviton and the scalar particles. Using that to linear order in $h_{\mu\nu}$, it holds that $g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}$ and $\sqrt{-g} = 1 + h/2$, we find

$$S[\eta+h,\phi] - S[\eta,\phi] = -\frac{1}{2}\int \mathrm{d}^4x\; h^{\mu\nu}\left[\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}\eta_{\mu\nu}\left(m^2\phi^2 + \frac{1}{2}\partial_\alpha\phi\partial^\alpha\phi\right)\right] , \tag{2.61}$$

where the term in square parentheses is proportional to the energy-momentum tensor of the scalar field. Transforming the term in parentheses to momentum space yields the vertex expression in Eq. (2.60), where we take care to consider that each scalar can have momentum k_1 or k_2, hence the appearance of the symmetrisation

$$k_1^{(\mu}k_2^{\nu)} = \frac{1}{2}\left(k_1^\mu k_2^\nu + k_1^\nu k_2^\mu\right) . \tag{2.62}$$

Thus, the amplitude of the Feynman diagram 2.2 is

$$\mathcal{M} = V^{\mu\nu}(k_1,k_2)D_{\mu\nu\rho\sigma}V^{\rho\sigma}(k_3,k_4) \approx -\frac{16\pi G m_1^2 m_2^2}{(k_2-k_1)^2} = -\frac{16\pi G m_1^2 m_2^2}{\vec{p}^{\,2}} , \tag{2.63}$$

where we are only considering the non-relativistic limit, i.e. the condition $|\vec{p}| \ll m_{1/2}$ for $\vec{p}$ the outgoing momentum. Using a well-known result from QFT relating the Fourier-transformed non-relativistic amplitude $\tilde{\mathcal{M}}_{\text{non-rel}}$ and the scattering potential, we find

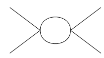
$$V(r) = \frac{\tilde{\mathcal{M}}_{\text{non-rel}}}{2m_1 2m_2} = \int \frac{\mathrm{d}^3p}{(2\pi)^3}e^{i\vec{p}\cdot\vec{r}}\left(-\frac{16\pi G}{\vec{p}^{\,2}}m_1^2m_2^2\right)\frac{1}{4m_1m_2} = -\frac{Gm_1m_2}{r} , \tag{2.64}$$

which corresponds to the classical Newtonian potential.

Now that we have re-derived the Newtonian potential, we can use perturbative Quantum Gravity to calculate how loop corrections modify it. Following dimensional analysis and the assumptions

$r \gg l_{\mathsf{Pl}}$ and $r \gg r_s = 2GM/c^2$, where l_{Pl} is the Planck length and r_s the Schwarzschild radius of an object of mass $m_{1/2}$, we can make the ansatz

$$V(r) = -\frac{Gm_1m_2}{r}\left(1 + a\frac{G(m_1+m_2)}{rc^2} + b\frac{G\hbar}{r^2c^3} + ...\right) . \tag{2.65}$$

The first correction is actually a classical term, but can be obtained from the loop expansion. It is one example of the fact that the loop expansion is **not** equal to an expansion in $\hbar$. The fact that one can calculate such post-Newtonian corrections from loop diagrams has given rise to a new research direction in the past years, where QFT methods are used to improve gravitational-wave calculations which are necessary to interpret the LIGO data. Reviews on this topic are, e.g., Levi 2020. The term proportional to b is a quantum gravity contribution from loop diagrams such as  and several others. These diagrams are all divergent and one may worry that counterterms (and their unknown prefactors) enter the expression for b. However, an r^{-3} term, as we are expecting in the potential, transforms in the following way to Fourier space:

$$\frac{1}{r^3} = -2\pi^2 \int \frac{\mathrm{d}^3p}{(2\pi)^3} \ln\left(\frac{\vec{p}^{\,2}}{\mu^2}\right) e^{i\vec{p}\cdot\vec{r}} . \tag{2.66}$$

This contribution to the amplitude is distinct from contributions of local counterterms. For instance, R^2 and $R_{\mu\nu}R^{\mu\nu}$ counterterms produce a q^4-vertex, and

$$\frac{1}{-q^2} q^4 \frac{1}{-q^2} \approx 1 , \tag{2.67}$$

is the corresponding amplitude, which is not of logarithmic form. Hence, it is possible to disentangle the infrared (IR) effect (this is the leading order correction to the potential in l_{Pl}/r) from the ultraviolet (UV) effect (this is the contribution of local counterterms). It was shown in 1994 by John Donoghue that $b = 41/10\pi$ (Donoghue 1994).

It is sometimes stated that we do not know how to reconcile quantum physics with gravity. The above example exemplifies that this statement is not true. In the EFT approach, one can extract the leading-order quantum-gravity contributions in the l_{Pl}/r – or E/m_p – expansion. In other words, there is a predictive quantum theory of gravity.

The remaining problem is that this theory does not extend to length-scales below the Planck length (energy scales above the Planck mass). Thus, we still need a UV-completion.

There are various possible strategies to try to quantize gravity in an ultraviolet complete way. A non-exhaustive list of examples is the following:

- We interpret the appearance of R^2 and $R_{\mu\nu}R^{\mu\nu}$ at one-loop order as a sign that we need to add these terms to the classical theory. This is Quadratic Gravity.

- We note that infinitely many coupling are not a priori a problem, as long as their values are known. What happens if we relate couplings to each other through a symmetry principle? This is Asymptotically Safe Gravity.

- We conclude that perturbation theory has failed for gravity and we need to go beyond treating perturbations about a flat background $\eta_{\mu\nu}$ and instead consider all possible metric configurations $g_{\mu\nu}$ on an equal footing, implementing background independence in quantum gravity. This leads, e.g., to Loop Quantum Gravity.

- We conclude that QFT has failed for gravity and a new formalism is required. In particular, because the predictivity problem arises, once we try to take the QFT-approach to sub-Planckian length scales, we might conclude that the new formalism has to account for a non-minimal length scale. This leads for example to String Theory, where the string scale constitutes a scale where a purely local framework breaks down, or to Causal Sets, in which continuous spacetime is treated as an approximation to a fundamentally discrete spacetime. Regarding the latter, a brief introduction to Causal Set Theory can be found in Appendix A

Chapter 3

Quadratic Gravity

We consider the action

$$S = \frac{1}{16\pi G} \int \mathrm{d}^4x \sqrt{-g} R + \int \mathrm{d}^4x \sqrt{-g} \left(\alpha R^2 + \beta R_{\mu\nu} R^{\mu\nu} \right) . \tag{3.1}$$

We will attempt to answer the following questions:

1. Is it compatible with observations to add the higher-order terms?
2. What are physical effects of the higher-order terms?
3. What is the status of renormalisation of this theory?
4. Why is this theory (presumably) unstable?

3.1 Is Quadratic Gravity compatible with observations?

In experimentally accessible situations, curvature is small, i.e. $|\alpha R^2| , |\beta R_{\mu\nu} R^{\mu\nu}| \ll m_p^2 R$. This is the case even for observations of black holes, which probe some of the largest values of curvature accessible to us. The currently strongest bounds on α and β come from constraints on modifications of Newton's law. In quadratic gravity, the Newtonian potential takes the form

$$\phi(r) = -\frac{Gm}{r} \left[1 + \frac{1}{3} \exp\left(-\sqrt{32\pi G(3\alpha - \beta)} r \right) - \frac{4}{3} \exp\left(-\sqrt{16\pi G \beta r} \right) \right] . \tag{3.2}$$

Experimental tests yield the constraints $|\alpha| , |\beta| \ll 10^{60}$. These are not very strong, because the observationally accessible curvature scales are much lower than Planckian curvature radii. Note that from this expression, we already see the presence of two massive modes which create the Yukawa term e^{-mr}. These two will become important for the fourth question, namely the stability of the theory.

3.2 Classical phenomenology of Quadratic Gravity

Modifications of General Relativity at the classical level can be expected to be important in (at least) two settings, namely the early universe and black holes.

In the early universe, an R^2 term in the action constitutes a simple model of inflation, known as Starobinsky inflation. For an appropriate choice of couplings ($\beta = 0$, $\alpha \sim 10^5$), it is in good agreement with measurements of the Cosmic Microwave Background. This is an attractive scenario for inflation, because it does not rely on the explicit addition of a new scalar field with an ad-hoc postulated potential. Instead, it is actually the scalar degree of freedom in the metric that becomes dynamical through the addition of the R^2 coupling and has a potential that is suitable for a slow-roll regime.
While inflation is just one, and neither the only, nor an experimentally unequivocally confirmed, possibility for the physics of the early universe, Starobinsky inflation may nevertheless be taken as an indication that quadratic-gravity terms can be physically interesting.

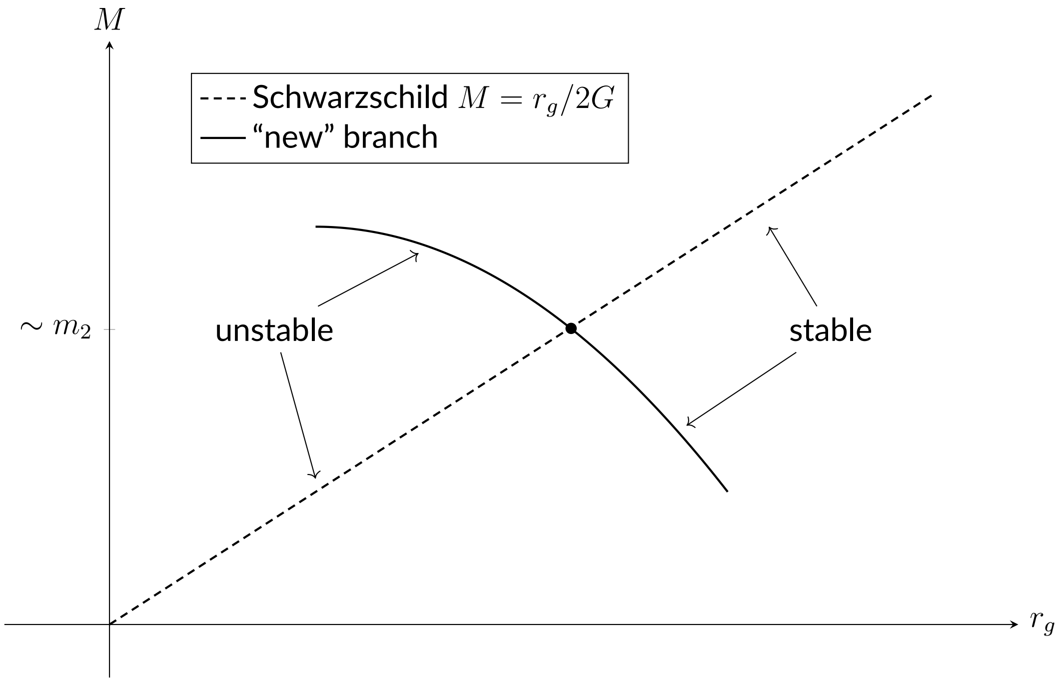


Figure 3.1: *Scaling of the event horizon r_g with the mass M for Quadratic Gravity black hole solutions and for the Schwarzschild solution of General Relativity*

For black holes, our first consideration is that any vacuum solution of General Relativity is also a vacuum solution of Quadratic Gravity, because the vacuum equations of motion are

$$\frac{1}{8\pi G}G_{\mu\nu} + \frac{1}{2}g_{\mu\nu}\left(\alpha R^2 + \beta R_{\mu\nu}R^{\mu\nu}\right) + 2\alpha R\frac{\delta R}{\delta g_{\mu\nu}} + 2\beta R^{\kappa\lambda}\frac{\delta R_{\kappa\lambda}}{\delta g_{\mu\nu}} = 0\,, \tag{3.3}$$

and thus $R = 0 = R_{\mu\nu}$, as is the case for every General Relativity vacuum solution, sets the extra terms to zero. Hence, Kerr black holes are solutions of Quadratic Gravity. However, uniqueness theorems from General Relativity no longer hold and there may therefore be **extra** black-hole solutions (even in $d = 4$).[i] Indeed, new solutions can be found, see Lu et al. 2015. These are only known numerically, not analytically. Numerically, one has the result (at finite m_2) displayed in Fig. 3.1, obtained in Held and Zhang 2023. In addition to the Kerr solution (or, as shown in the plot, its Schwarzschild limit), there is a new branch of solutions that is degenerate with the Kerr solution at one particular value of the mass. For lower values of the mass, the new solution is dynamically stable (and the Schwarzschild solution is unstable), whereas for larger values of the mass, the opposite is true.

3.3 What is the status of renormalisability of Quadratic Gravity?

As in General Relativity, we can compute the superficial degree of divergence in the perturbative quantization of Quadratic Gravity. First, we determine the momentum-scaling of the propagator. Following the observations we made for General Relativity, where we found that in momentum space $R \sim p^2$, we can conclude $R^2 \sim p^4$ and $R_{\mu\nu}R^{\mu\nu} \sim p^4$. And hence, at large momentum $p \gg m_{\mathrm{mPl}}$, we obtain the scaling

$$\mathcal{P} \sim \frac{1}{p^4}\,, \tag{3.4}$$

of the propagator $\mathcal{P}$. For the vertices V, we similarly conclude

$$V \sim p^4\,, \tag{3.5}$$

for $p \gg m_{\mathrm{Pl}}$. The superficial degree of divergence D in d-dimensions is then given by

$$D = dL - 4P + 4V = dL - 4(L + V - 1) + 4V = (d-4)L + 4\,, \tag{3.6}$$

where we inserted the topological relation $L = P - V + 1$. In $d = 4$, we hence find

$$D = (d-4)L + 4 \xrightarrow{d\to 4} 4\,, \tag{3.7}$$

a finite value at all loop orders. This indicates that only a finite number of counterterms are needed to renormalise the theory. Hence, Quadratic Gravity is perturbatively renormalisable.

But is renormalisability enough to ensure that a theory is fundamental, i.e. valid (at least theoretically) at arbitrary small distances? The answer is **no**. The reason is that renormalisability only ensures that UV divergences can be absorbed in a finite number of free parameters of the theory, namely the couplings in front of the counterterms. However, it has no implications for how these couplings change under the Renormalisation Group (RG) flow, i.e. there can still be divergences in the scale-dependence of the finitely many couplings (so-called Landau poles). We will introduce the

[i] In General Relativity, there are additional solutions (e.g., black strings) in $d \gtrsim 5$, but the Kerr-Newman family is the unique vacuum black-hole family.

RG flow more thoroughly below, when we introduce the functional RG for quantum gravity. For now, it is sufficient to say that when we probe the theory at different scales (e.g. experimentally by performing scattering, or theoretically, by introducing a cut-off that we shift[ii]), the couplings change. We will call the RG scale μ for now and introduce $\beta_g := \mu\partial_\mu g(\mu)$ for a coupling g.

Example: β-function in $\lambda\phi^4$-theory

A simple example that perturbative renormalisability is not sufficient to make a theory UV complete, i.e., to guarantee that the couplings are finite at all scales μ, is $\lambda\phi^4$-theory (in $d = 4$). It is renormalisable but

$$\beta_\lambda = \frac{3}{16\pi^2}\lambda^2 + \mathcal{O}(\lambda^3)\,, \tag{3.8}$$

which, when applying the definition of the β-function $\beta_\lambda = \mu\partial_\mu\lambda$, translates to

$$\lambda(\mu) = \frac{\lambda(\mu_0)}{1 + \frac{3}{16\pi^2}\lambda(\mu_0)\ln(\mu_0/\mu)}\,, \tag{3.9}$$

where μ_0 is a (low-energy) reference scale at which we provide an initial condition for the RG flow. We see that $\lambda(\mu) \to \infty$ for $\mu \to \mu_c < \infty$, unless $\lambda(\mu_0) = 0$. Thus, there is a divergence in the coupling at a finite scale, unless $\lambda(\mu_0) = 0$, which makes the theory non-interacting (i.e. trivial).

Note that as $\lambda(\mu)$ increases, higher-order terms in the loop expansion become important and the perturbative expansion we used to make this argument breaks down. But non-perturbative lattice studies actually confirm the result of our above argument, i.e., ϕ^4 theory is trivial in four dimensions.

Thus, we can ask ourselves whether Quadratic Gravity is UV complete? We will consider its β-functions in the following parametrisation of the action of Quadratic Gravity:

$$S = \int \mathrm{d}^4x\sqrt{-g}\left(\frac{m_p^2}{2}(R - 2\Lambda) - \frac{1}{2\lambda}C^2 - \frac{1}{\xi}R^2\right)\,, \tag{3.10}$$

where

$$C_{\beta\nu\kappa\lambda} = R_{\beta\nu\kappa\lambda} - \frac{1}{2}\left(R_{\beta\lambda}g_{\nu\kappa} - R_{\beta\kappa}g_{\nu\lambda} + R_{\nu\kappa}g_{\beta\lambda} - R_{\nu\lambda}g_{\beta\kappa}\right) + \frac{1}{6}R(g_{\beta\kappa}g_{\nu\lambda} - g_{\beta\lambda}g_{\nu\kappa})\,, \tag{3.11}$$

is the Weyl tensor. We consider the β-functions by Buccio et al. 2024, where it was found that

$$\beta_\lambda = -\frac{1}{16\pi^2}\left(\frac{1617\lambda - 20\xi}{90}\right)\lambda\,, \tag{3.12}$$

$$\beta_\xi = -\frac{1}{16\pi^2}\left(\frac{\xi^2 - 36\lambda\xi - 2520\lambda^2}{36}\right)\,. \tag{3.13}$$

[ii]These do not necessarily lead to the same response of the theory; only some aspects of such scale dependencies are universal – more later.

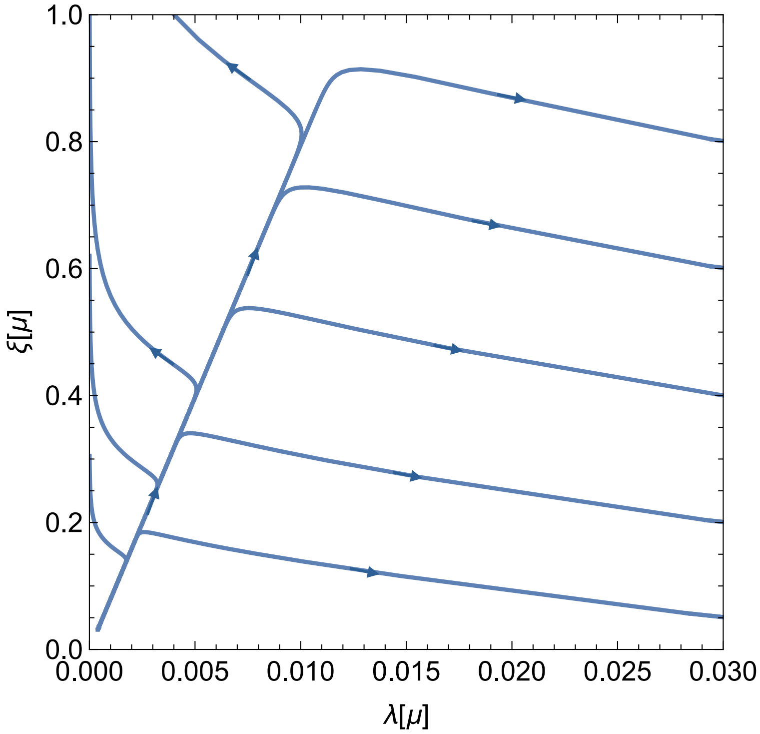


Figure 3.2: *We show numerically integrated RG trajectories, where the arrows show how the couplings change as μ is lowered. All trajectories start out in the close vicinity to the free fixed point $(\lambda = 0, \xi = 0)$ and lead to nonzero values of both couplings.*

This is a coupled set of equations, so for simplicity we first consider the two beta functions on the line where the respective other coupling is set to zero:

$$\beta_\lambda\Big|_{\xi=0} = -\frac{1617}{90}\frac{\lambda^2}{16\pi^2}\,, \tag{3.14}$$

$$\beta_\xi\Big|_{\lambda=0} = -\frac{1}{36}\frac{\xi^2}{16\pi^2}\,. \tag{3.15}$$

The negative signs are indicative of **asymptotic freedom**, like in Quantum Chromodynamics (QCD), but to actually show asymptotic freedom, we must consider the coupled set of beta functions. We can do so in a plot that shows numerically integrated RG trajectories in the λ- ξ plane, see Fig. 3.2. The figure shows that asymptotic freedom is indeed achieved, because nonzero values of the couplings can be traced back to values which are essentially zero at large values of μ.

Given that the curvature-squared couplings are not strongly constrained by observations (i.e. observations allow a large range of possible values), the result that adding the couplings results in interesting phenomenology (e.g., inflation without an explicitly added scalar field) and the fact that the theory is perturbatively renormalisable and even asymptotically free, why is Quadratic Gravity then not the accepted theory of quantum gravity? The problem is **stability**.

3.4 Why is Quadratic Gravity (presumably) unstable?

Quadratic Gravity is widely expected to be unstable in the sense of showing runaway behavior in its classical evolution and being non-unitarity as a quantum theory. More recently, both expectations have been questioned Donoghue and Menezes 2019; Held and Lim 2023, but here we present the argument on which this expectation relies.

This argument is tied to the fourth-order time derivatives in R^2 and $R_{\mu\nu}R^{\mu\nu}$, and it applies generically to many (not all) systems with higher than second order time derivatives. It is called the **Ostrogradsky instability**. We will review it in the form in which Ostrogradsky originally showed it, namely for the classical mechanics of point particles. Consider then a Lagrangian function $L = L(x, \dot{x}, \ddot{x})$. This generalises the standard situation in classical mechanics, where $L = L(x, \dot{x})$, to a setting with a higher-order time derivative, here a second derivative. It will introduce an extra initial condition that we need to solve the equations of motion, and this extra degree of freedom will make the Hamiltonian unbounded from below. This is generically expected to lead to runaway-instabilities in the time-evolution of the system.
The Euler-Lagrange equations of the system are

$$\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial L}{\partial \ddot{x}} = 0\,. \tag{3.16}$$

There are four initial conditions required and there are therefore four canonical variables (two generalized positions and two momenta). Ostrogradsky chose

$$x_1 = x\,, \tag{3.17}$$

$$x_2 = \dot{x}\,, \tag{3.18}$$

$$p_1 = \frac{\partial L}{\partial \dot{x}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \ddot{x}}\,, \tag{3.19}$$

$$p_2 = \frac{\partial L}{\partial \ddot{x}}\,. \tag{3.20}$$

The Hamiltonian is obtained by Legendre transformation

$$H = \sum_{i=1}^{2} p_i \dot{x}^i - L \tag{3.21}$$

$$= p_1\dot{x} + p_2\ddot{x} - L\,. \tag{3.22}$$

Upon inserting the Lagrangian, one is left with a Hamiltonian that is linear in one of the momenta and thus unbounded from below. This is generically expected to lead to runaway instabilities. (Already at the level of classical mechanics, there are exceptions to this, e.g., if the Hamiltonian is degenerate (such that there are actually fewer degrees of freedom) or if there are additional constants of motion that prevent the system from exploring the unstable regime.)

Example: Ostrogradsky instability

Consider the Lagrangian

$$L = -\frac{\epsilon m}{2\omega^2}\ddot{x}^2 + \frac{m}{2}\dot{x}^2 - \frac{m\omega}{2}x^2\,, \tag{3.23}$$

where ϵ quantifies the deviation from the Lagrangian of the harmonic oscillator. The canonical momenta are

$$p_1 = m\dot{x} - \frac{\mathrm{d}}{\mathrm{d}t}\left(-\frac{\epsilon m}{\omega^2}\ddot{x}\right)\,, \tag{3.24}$$

$$p_2 = -\frac{\epsilon m}{\omega^2}\ddot{x} \tag{3.25}$$

Inserting these into the expression for the Hamiltonian (3.21), we find

$$H = \left(m\dot{x} + \frac{\epsilon m}{\omega^2}\dddot{x}\right)\dot{x} - \frac{\epsilon m}{\omega^2}\ddot{x}\ddot{x} + \frac{\epsilon m}{2\omega^2}\ddot{x}^2 - \frac{m}{2}\dot{x}^2 + \frac{m\omega^2}{2}x^2 \tag{3.26}$$

$$= \frac{\epsilon m}{\omega^2}\dddot{x}\dot{x} - \frac{\epsilon m}{2\omega^2}\ddot{x}^2 + \frac{m}{2}\dot{x}^2 + \frac{m\omega^2}{2}x^2\,. \tag{3.27}$$

We notice that the first term is linear $\dddot{x}$, meaning that it can be made arbitrarily negative. (Note that this particular example is actually just two decoupled harmonic oscillators, which one can see upon diagonalizing the Hamiltonian. This is therefore an example which also highlights that a Hamiltonian that is unbounded from below does not automatically give rise to an instability.)

In a quantum theory, the Hamiltonian can be made bounded from below if one flips the role of creation and annihilation operators, which amounts to introducing a negative sign for the norm in Hilbert space – i.e., a violation of unitarity. Therefore, one generically expects that theories with higher than second order time derivatives violate unitarity upon quantization. This expectation is subject to the same caveats as discussed above.

To see that quadratic gravity has the same problem, we only need to note that there are four derivatives in R^2 and $R_{\mu\nu}R^{\mu\nu}$ and because of local Lorentz invariance, they cannot just be spatial derivatives but must also contain time derivatives.[iii] We will take a closer look, because we also want to understand which degrees of freedom actually cause the (presumed) instability.

To that end, we consider the propagator in the theory. For that step, it is useful to introduce spin projectors. These allow us to separate the gauge degrees of freedom from the physical degrees of freedom. These projectors split a general symmetric rank-2 tensor field $h_{\mu\nu}$ into its components, which transform in irreducible representations of the Lorentz group. This is similar to how a vector V^μ can be decomposed into a transverse part $\left(V^T\right)^\mu$ that transforms in the spin-1 representation

[iii]The idea that a quantum gravity theory which breaks Lorentz invariance, such that only spatial derivatives occur at higher order has a better behavior under renormalization than the perturbative quantization of GR has been explored in Hořava-Lifshitz gravity.

and a longitudinal part $\left(V^L\right)^\mu$ that transforms in the spin-0 representation. In Fourier space, this is achieved by using the projectors

$$T^\mu{}_\nu = \delta^\mu{}_\nu - \frac{p^\mu p_\nu}{p^2}\,, \tag{3.28}$$

$$L^\mu{}_\nu = \frac{p^\mu p_\nu}{p^2}\,, \tag{3.29}$$

which project out the component V^T that is transverse to the momentum p^μ and V^L, which picks the longitudinal component, i.e. the one aligned with p^μ.

Properties of the rank-1 spin projectors

The goal of this exercise is to check that the spin projectors defined above are orthogonal projectors, i.e. we need to check that

$$T^\mu{}_\nu T^\nu{}_\kappa = T^\mu{}_\kappa\,, \tag{3.30}$$

$$T^\mu{}_\nu L^\nu{}_\kappa = 0\,, \tag{3.31}$$

$$L^\mu{}_\nu L^\nu{}_\kappa = L^\mu{}_\kappa\,. \tag{3.32}$$

Rewriting the spin projectors in matrix notation rather than in index notation allows us to quickly verify these equations. Doing so, we find

$$T = \mathbf{1} - \frac{p\otimes p}{p^2}\,, \tag{3.33}$$

$$L = \frac{p\otimes p}{p^2}\,. \tag{3.34}$$

With this, we can compute

$$\begin{aligned}
T^2 &= \left(\mathbf{1} - \frac{p\otimes p}{p^2}\right)\left(\mathbf{1} - \frac{p\otimes p}{p^2}\right) && (3.35)\\
&= \mathbf{1} - \frac{2\,(p\otimes p)}{p^2} + \frac{1}{p^4}(p\otimes p)(p\otimes p) && (3.36)\\
&= \mathbf{1} - \frac{2\,(p\otimes p)}{p^2} + \frac{1}{p^4}p^2(p\otimes p) && (3.37)\\
&= T\,, && (3.38)
\end{aligned}$$

$$\begin{aligned}
TL &= \left(\mathbf{1} - \frac{p\otimes p}{p^2}\right)\frac{1}{p^2}p\otimes p && (3.39)\\
&= \frac{p\otimes p}{p^2} - \frac{1}{p^4}(p\otimes p)(p\otimes p) && (3.40)\\
&= \frac{p\otimes p}{p^2} - \frac{1}{p^4}p^2(p\otimes p) && (3.41)\\
&= 0\,, && (3.42)
\end{aligned}$$

$$L^2 = \frac{p\otimes p}{p^2}\frac{p\otimes p}{p^2} \tag{3.43}$$
$$= \frac{1}{p^4}p^2(p\otimes p) \tag{3.44}$$
$$= L\,. \tag{3.45}$$

In this calculation, we repeatedly used the identity

$$(p\otimes p)(p\otimes p) = p^2(p\otimes p)\,, \tag{3.46}$$

which follows by the rules of matrix multiplication.

Similarly, a symmetric rank-2 tensor $h_{\mu\nu}$ (in four dimensions) can be decomposed into

- a spin-2 representation by acting with

$$\mathcal{P}^{(2)}{}_{\mu\nu}{}^{\rho\sigma} = \frac{1}{2}\left(T^\rho{}_\mu T^\sigma{}_\nu + T^\sigma{}_\mu T^\rho{}_\nu\right) - \frac{1}{3}T_{\mu\nu}T^{\rho\sigma}\,, \tag{3.47}$$

- a spin-1 representation

$$\mathcal{P}^{(1)}{}_{\mu\nu}{}^{\rho\sigma} = \frac{1}{2}\left(T_\mu{}^\rho L^\sigma{}_\nu + T_\mu{}^\sigma L^\rho{}_\nu + T_\nu{}^\rho L^\sigma{}_\mu + T_\nu{}^\sigma L^\rho{}_\mu\right) \tag{3.48}$$

- two spin-0 representations

$$\mathcal{P}^{(0,\text{tr})}{}_{\mu\nu}{}^{\rho\sigma} = \frac{1}{3}T_{\mu\nu}T^{\rho\sigma}\,, \tag{3.49}$$
$$\mathcal{P}^{(0,l)}{}_{\mu\nu}{}^{\rho\sigma} = L_{\mu\nu}L^{\rho\sigma}\,. \tag{3.50}$$

After gauge-fixing, only the spin-2 representation and one scalar representation contribute to the propagating degrees of freedom in Quadratic Gravity. If we write the action as

$$S = m_p^2\int \mathrm{d}^4x\sqrt{-g}R + \int \mathrm{d}^4x\sqrt{-g}\left(\frac{1}{2\lambda}C^2 + \frac{1}{\xi}R^2\right) \tag{3.51}$$

we get the propagators which are products of the projects with momentum- and coupling-dependent prefactors. These prefactors are

$$\frac{4\lambda}{p^4 - \frac{1}{2}\lambda m_p^2 p^2} \qquad \text{for the spin-2 part}\,, \tag{3.52}$$
$$\frac{\xi/3}{p^4 + \frac{\xi}{12}m_p^2 p^2} \qquad \text{for the scalar part}\,. \tag{3.53}$$

These can be decomposed by using a partial-fractions decomposition. For the spin-2 part, we find

$$\frac{4\lambda}{p^4 - \frac{1}{2}\lambda m_p^2 p^2} = \frac{8}{m_p^2}\left(\frac{1}{-p^2} - \frac{1}{-p^2 + \frac{1}{2}\lambda m_p^2}\right)\,. \tag{3.54}$$

The first term comes from the contribution of a massless spin-2 particle – this is the graviton. The second term is the contribution of a spin-2 graviton with the **opposite** sign of the propagator. This can be traced back to the kinetic term having the opposite sign. This is the (presumed) instability (“ghost”). For the scalar part, we find

$$\frac{\xi/3}{p^4 + \frac{\xi}{12} m_p^2 p^2} = \frac{4}{m_p^2} \left(-\frac{1}{-p^2} + \frac{1}{-p^2 - \frac{\xi}{12} m_p^2} \right) . \tag{3.55}$$

The first term is the contribution of a massless scalar ghost and the second term is a massive scalar. It is a tachyon if $\xi < 0$.

Chapter 4

Asymptotic Safety

We have seen that perturbative non-renormalisability represents a breakdown of predictivity. If we wish to find a theory of Quantum Gravity within the framework of QFT, we will need to answer the following questions:

1. Can we find a mechanism/principle that relates couplings to each other, so that even if there are infinitely many couplings, there are just finitely many free parameters?

2. Can such a theory avoid Ostrogradski's theorem and be unitary?

4.1 Running couplings and β-functions

We first focus on the first question. There is one piece of information we have not used so far, when assessing the predictivity of a theory and that is how couplings depend on the scale. The scale can be thought of as a momentum/wavenumber, or its inverse, a wavelength. We caution, though, that this need not be a physical wavenumber (in the sense of the wavenumber of an external state in a scattering experiment), but it can also be a quantity with units of momentum that is introduced as an auxiliary parameter into the calculation. For now, we keep the notion of this scale general, later we will specialize to the functional Renormalisation Group and its associated RG scale k. Roughly, we can think of the scale as one that determines in which the range of energies/momenta quantum fluctuations have been accounted for.

Given a set of interactions parameterized by the couplings $\bar{g}_i$, we can always calculate

$$\beta_{\bar{g}_i} = k\partial_k \bar{g}_i(k)\,, \tag{4.1}$$

which encodes the scale dependence. We can (in principle) do this for the infinitely many couplings in gravity.

What we need in order to make the theory predictive at all scales is some principle that gives us an "initial condition" for the couplings, i.e. a principle that determines the values of all (or all but

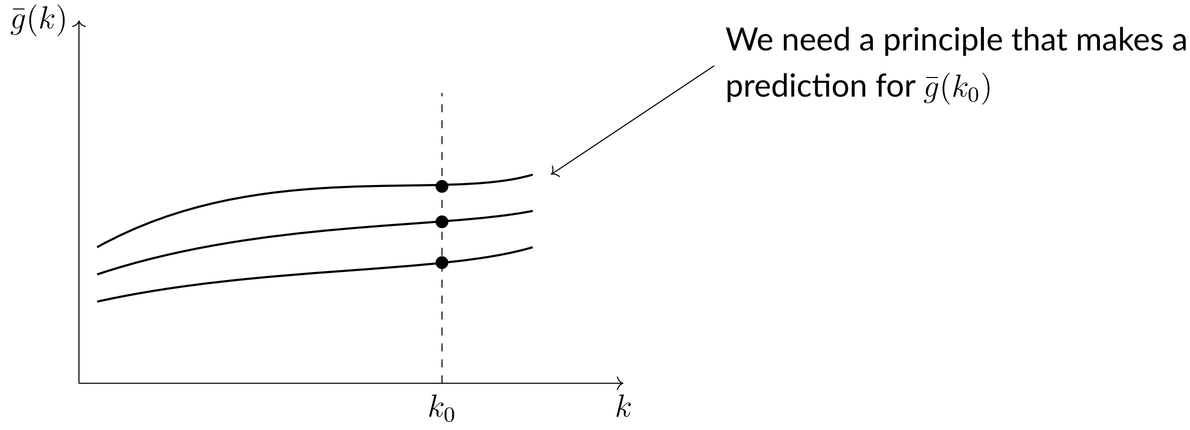


Figure 4.1: *The scale dependence of the coupling constants $\bar{g}_i(k)$ as a function the scale k; the scale dependence is determined by the Renormalization Group (RG) flow.*

finitely many) couplings at some scale k_0, so that we do not have to perform an infinite number of experiments to make the theory predictive (see Fig. 4.1). One such principle is to demand a Renormalisation Group (RG) fixed point, i.e. we demand **scale symmetry**.
This is not defined by $\beta_{\bar{g}_i} = 0$, because some couplings have non-zero dimensionality, i.e. they are not a pure number, but define a scale. In a theory that is classically scale symmetric, such couplings have to be set to zero. We are interested in a slightly different notion of scale symmetry, namely quantum scale symmetry. This is a version of scale symmetry that arises in the presence of quantum fluctuations, as we will see below. To detect a quantum scale symmetric regime in the beta functions, we work with dimensionless quantities. Thus, we define

$$g_i = \bar{g}_i \cdot k^{-d\bar{g}_i} \,, \tag{4.2}$$

where $-d\bar{g}_i$ is the mass-dimension of $\bar{g}_i$. By this definition, the new couplings g_i are dimensionless. Now, we know that scale symmetry is achieved, if $\beta_{g_i} = 0$ for all couplings g_i.

Why does this restore predictivity? The equation $\beta_{g_i} = 0$ provides one constraint for each coupling; the system of beta functions is in general a coupled system of polynomials (in perturbation theory) or even non-polynomial expressions (e.g. algebraic or exponential expressions beyond perturbation theory). Generically, there are at most a few (and typically not more than one non-trivial) real zeroes of this system.

A theory that is fully scale-symmetric is therefore expected to yield predictions for the values of all couplings. However, Nature is not scale symmetric, there are distinct scales in nature (e.g. masses of elementary particles). This means that, at best, scale-symmetry can only be realized asymptotically, at large k (i.e. microscopically), but there must be a transition scale $k_{\rm tr}$ at which a transition away from scale–symmetry can occur (see Fig. 4.2). At first, one might think that for $k < k_{\rm tr}$, the theory is no longer predictive, because the couplings are not constrained by the requirement β_{g_i} any more. However, as we will discuss in detail below, predictivity persists. Losely speaking, a

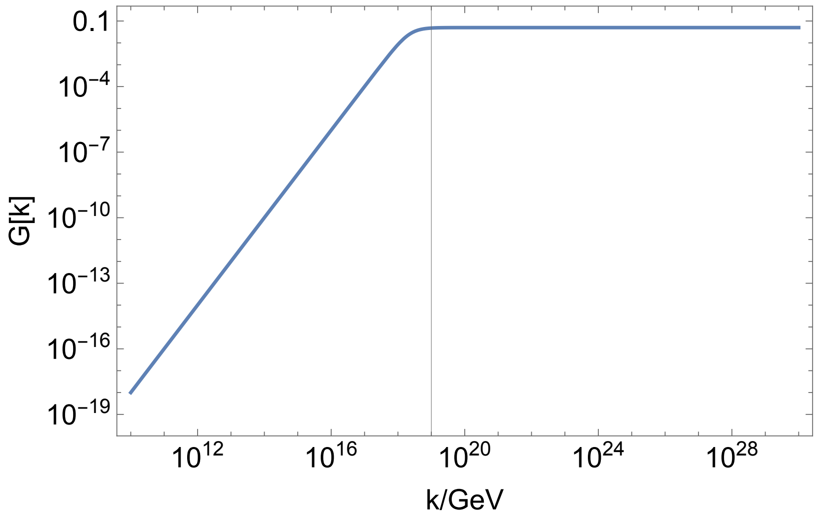


Figure 4.2: *The dependence of the dimensionless Newton constant G on the squared scale k^2 (with $\# = 40$ for purposes of illustration). Below $k_{\rm tr} \approx 1019\,\mathrm{GeV}$, there is no scale symmetry; above, there is approximate scale symmetry.*

theory that is (approximately) scale symmetry in the microscopic regime cannot realize arbitrary values of all of its couplings at macroscopic scales.

Example: β–function for the Newton coupling

We will later see how to derive the β–function for the Newton coupling. We quote the result here:

$$\beta_G = 2G - \#G^2\,, \tag{4.3}$$

where $\# > 0$ is some number. The first term comes about, because $[\bar{G}] = -2$ in $d = 4$ and hence

$$G = \bar{G}k^2\,. \tag{4.4}$$

Inserting this into (4.1) leads to

$$\beta_G = k\partial_k(\bar{G}k^2) = 2k^2\bar{G} + k^2 k\partial_k\bar{G} = 2G - \#G^2\,. \tag{4.5}$$

The computation of the second term in the last step requires new techniques we will develop below (Functional RG). Setting the β–function equal to zero and solving for G, we find the solutions

$$\beta_G = 0 \iff G_* = \left\{0, \frac{2}{\#}\right\}\,. \tag{4.6}$$

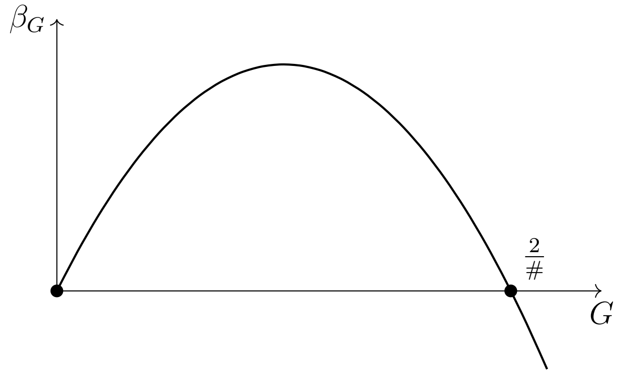


Figure 4.3: *The dependence of the β_G–function on the Newton constant G. We had found the explicit dependence $\beta_G = 2G - \#G^2$ above.*

4.2 How does asymptotic safety generate predictivity at all scales?

To understand how asymptotic safety generates predictivity at $k < k_{\mathrm{tr}}$, where $\beta_{g_i} = 0$ no longer holds, we will introduce the relevant and irrelevant couplings of the RG flow (or its sources and sinks). The relevant couplings are the only ones that introduce free parameters and the irrelevant couplings can be calculated in terms of the relevant ones *without* additional free parameters.

We first consider a coupling that is relevant, namely the Newton coupling. In Fig. 4.3, the dependence of the β_G–function on G^2 has been displayed. The beta function encodes whether the coupling increases or decreases, because it is the (dimensionless) scale derivative. There is one extra sign we have to consider, because the beta function is defined as the dimensionless scale derivative with respect to a momentum scale, i.e., it tells us whether the coupling grows or shrinks as we "zoom in" into the microscopic regime. However, the *physical* direction of the RG flow is actually from large to small k, i.e., from microphysics to macrophysics. We are therefore interested in determining whether a coupling grows or shrinks as k decreases. We thus have that:

- $\beta_G < 0 \implies k\partial_k G < 0$, this implies that $G(k)$ **grows** towards low k. We say that the coupling is anti-screened.
- $\beta_G > 0 \implies k\partial_k G > 0$, this implies that $G(k)$ **decreases** towards low k. The coupling is screened.

Hence, we see that the fixed point at $G = 2/\#$ is unstable, i.e. a tiny perturbation away from the scale-symmetric point

$$\beta_G(G_* + \delta G) = 2(G_* + \delta G) - \#(G_* + \delta G)^2 = -\#\delta G^2 < 0\,, \tag{4.7}$$

is negative and by our characterization above, means that coupling grows towards low k. We call the coupling relevant, because the perturbation away from scale-symmetry is generically large at low energies and thus relevant for the dynamics of the theory. A relevant coupling enables us to move away from a scale-symmetric regime and reconcile the idea of scale-symmetry at microscopic scales with the existence of scales in Nature.

Relevant couplings realize the idea that scale symmetry at $k > k_{\rm tr}$ does not impose restrictions at $k < k_{\rm tr}$, because we can reach any value of G at $k < k_{\rm tr}$.[i]

The case of irrelevant couplings is much more non-trivial, because irrelevant couplings are constrained at all k, even $k < k_{\rm tr}$. To see this on a second example, let us introduce a second coupling, g, with

$$\beta_g = -Gg + \beta_1 g^3\,, \qquad \beta_1 < 0\,,$$

and attempt to determine the RG flow of the two couplings g and G.[ii]

Exercise: Finding the fixed points and relevant directions in the $G-g$-system.

The goal of this exercise is to find the fixed points and characterize the RG flow of the system

$$\begin{cases} \beta_G(G) = 2G - \#_1 G^2 \\ \beta_g(g) = -Gg + \#_2 g^3 \end{cases} \tag{4.8}$$

As we now know, the fixed points are found at the points (G_*, g_*) in couplings-space where both β–function are simultaneously zero. We already know the points G_*, where the β_G–function vanishes (c.f. (4.6)). And because β_g has no constant term, we know of two fixed points (FP) already: FP A at $(\frac{2}{\#}, 0)$ and FP B at $(0, 0)$. To see if there are additional fixed point, we set the β_g–function to zero and solve for g. Doing so, we find

$$\beta_g = 0 \iff g_* = \left\{0, \pm\sqrt{\frac{G}{\#_2}}\right\}\,. \tag{4.9}$$

Hence, there is a third fixed point in the region, where the couplings are both positive, namely FP C at $(\frac{2}{\#_1}, \sqrt{\frac{2}{\#_1\#_2}})$. We can characterize the three fixed points as follows:

- FP A has two relevant directions, i.e. two free parameters. How do we know this? First, we saw in (4.7) by considering a perturbation about the FP that in the G-direction, the FP at $G_A = \frac{2}{\#}$ is unstable, hence we denote this by arrows moving away from the fixed point. To get the behavior in g-direction, we can similarly consider a perturbation of the β_g–function around the FP. We find

$$\beta_g(g)\Big|_{G=G_*} = -G_* g + \#_2 g^3 \approx -G_* g < 0\,, \tag{4.10}$$

  where we are considering small deviations from the value of g at the FP, i.e. small deviations from $g = 0$. The perturbed β_g function is negative and hence, the coupling grows towards low k. This is denoted similarly with an arrow moving away from the FP.

[i] Different choices of $G(k = k_{\rm IR})$, where $k_{\rm IR}$ is some fixed infrared scale, correspond to different $k_{\rm tr}$.

[ii] We will later see that the Abelian gauge coupling in the Standard Model with gravity has a beta function that corresponds to that of g.

- FP B has two irrelevant directions. This can be seen by once again perturbing the two β–functions around the FP and observing that these perturbations are positive, signifying a stable FP. We denote this by arrows pointing towards the FP from both G- and g-direction.

- FP C has one relevant direction and one irrelevant direction. The beta function for G does not depend on g, therefore, G is a relevant direction of any fixed point with $G = G_* = 2/\#_1$. Let us now turn our attention to the g-direction. Again considering a small perturbation, this time of the β_g–function, we find

$$\beta_g(g_C + \delta g) = 2\#_2 g_C \delta g^2 + \#_2 \delta g^3 > 0\,, \tag{4.11}$$

and hence an irrelevant direction. This results in the arrows pointing towards C in the g-direction and away from it in the G-direction (see Fig. 4.4)[a].

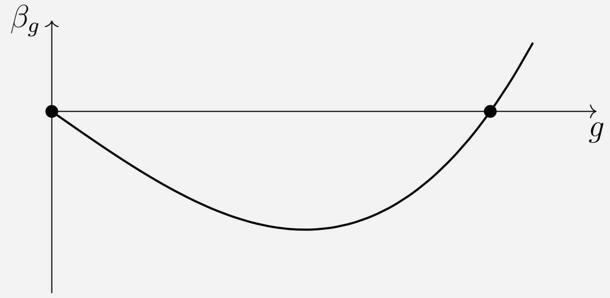


Figure 4.4: *The dependence of the β_g–function on the Newton constant g. We see that the β_g–function is positive for slight positive perturbations of the coupling, signifying that the coupling decreases towards low k. This is the marker of an irrelevant direction.*

With these specifications, we can draw the flow of these couplings (see Figure 4.5).

[a]More in detail, we see that the relevant direction is not exactly parallel to the G-axis. This can be understood from the stability matrix and its eigenvalues that we consider below.

Through this exercise, we have found the fixed points of a coupled system of β–functions. This was done by first finding the points where the β–functions both vanish and analyzing their behavior around those points. By doing so, we determined whether each fixed point is stable or unstable (or both, depending on which direction in couplings-space we are considering) and thus, whether the coupling is relevant or irrelevant. The fixed point we called C is an example for how a FP has predictive power even if the RG flow moves away from it. Indeed, it allows $g(k)$ to be determined at all scales as a function of $G(k)$. In other words, by knowing $g(k)$ at some scale k_0, we can follow the flow and connect it to a known value of $G(k)$.

In the space of all couplings g_i, if there is a fixed point, its **critical surface** is spanned by all relevant directions, i.e. any RG flow that starts in (or very close to) the scale-symmetric regime at small scales lies on (or very close to) the critical surface. The dimension of the critical surface is the number of

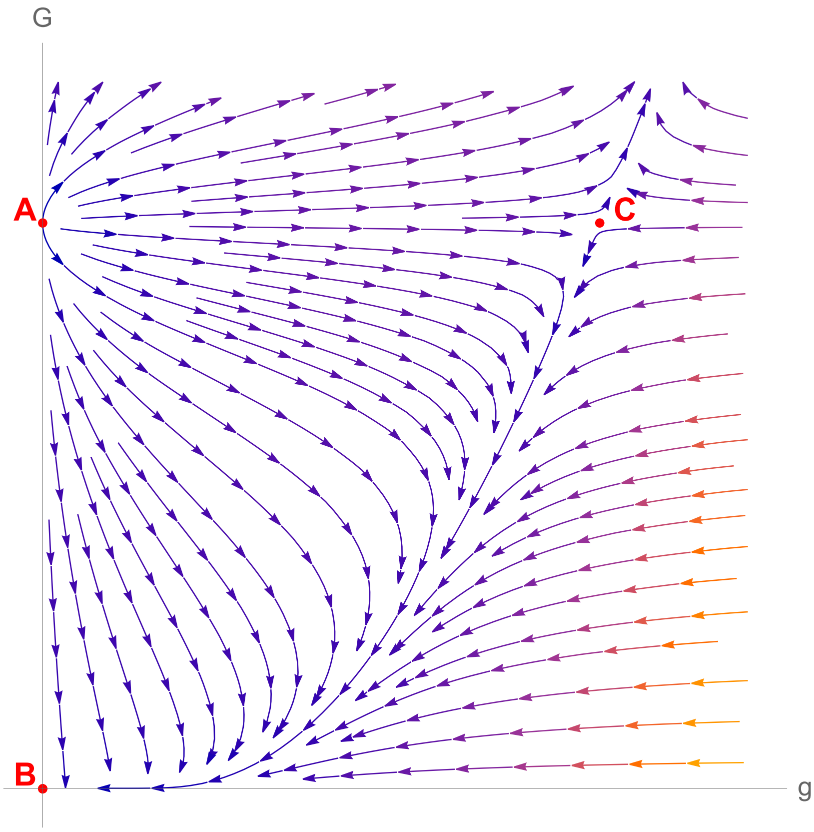


Figure 4.5: *A view of the coupling space G–g with the arrows displaying the flow of the couplings with respect to the scale (as discussed above, the arrows point towards the IR). The fixed points are displayed in red.*

free parameters of the theory; i.e., in general it is a hypersurface in the infinite-dimensional space of all couplings. This is because one experiment is needed to fix the value of a relevant coupling at some scale.

We used the above example to gain some intuition of predictive power associated to RG fixed points. We now want to formalize and generalize the above considerations. To calculate the number of relevant directions, we start from the stability matrix.

Stability matrix

Given the β–functions β_{g_i}, we define the stability matrix at the fixed point g_{i*} as

$$M_{ij} = \frac{\partial \beta_{g_i}}{\partial g_j}\Big|_{g_i = g_{i*}} . \tag{4.12}$$

We multiply its eigenvalues $\mathrm{eig}\left(M_{ij}\right)$ by an additional negative sign and call these the critical exponents θ_I, i.e.,

$$\theta_I = -\mathrm{eig}\left(M_{ij}\right) . \tag{4.13}$$

The stability matrix is important, because it encodes the RG flow close to a fixed point, i.e., the linearized RG flow,

$$\begin{aligned}\beta_{g_i} &= \beta_{g_i}\Big|_{g=g_*} + \sum_j \frac{\partial \beta_{g_i}}{\partial g_j}(g_j - g_{j*}) + \mathcal{O}\left((g_j - g_{j*})^2\right) \\ &= \sum_j \frac{\partial \beta_{g_i}}{\partial g_j}(g_j - g_{j*}) + \mathcal{O}\left((g_j - g_{j*})^2\right) \approx \sum_j M_{ij}(g_j - g_{j*})\,, \end{aligned} \tag{4.14}$$

where we used that $\beta_{g_i}\Big|_{g=g_*} = 0$ in going from the first to the second line. The linearized RG flow can be solved analytically to give

$$g_i(k) = g_{i*} + \sum_J c_J V_i^{(J)} \left(\frac{k}{k_0}\right)^{-\theta_J}\,, \tag{4.15}$$

where c_J are constants of integration, V^J are the (right) eigenvectors of the stability matrix and k_0 is a reference scale. One may check that this expression solves the differential equation 4.14 by plugging it in. The solution contains a sum over all eigenvectors, because (as we already saw close to fixed point C in Fig. 4.5, the (ir)relevant directions of the flow are not always parallel to the coupling-axis that we have chosen, instead, (ir)relevant directions may be superpositions of couplings. Below, we derive this solution.

To solve the differential equation 4.14, we will change the basis in the space of couplings, such that the stability matrix is diagonal in our new basis. We first shift the couplings by defining $h_j := g_j - g_{j*}$. Then, the equation becomes in matrix notation

$$k\frac{\partial h}{\partial k} = Mh\,. \tag{4.16}$$

The stability matrix can now be diagonalized $M = PDP^{-1}$, where P is the matrix built up of the eigenvectors V of M as its columns and D is of the form diag$(\theta_1, ..., \theta_N)$, where θ_I are the eigenvalues of the stability matrix M. Hence, rotating the couplings to new couplings $u := P^{-1}h$, we may write

$$k\frac{\partial u}{\partial k} = Du\,, \tag{4.17}$$

which decouples the equations. All the linear ODE's can now be solved separately to give

$$u_I(k) = c_I \left(\frac{k}{k_0}\right)^{-\theta_I}\,, \tag{4.18}$$

where c_I is an integration constant and k_0 is some constant scale. Transforming back to our original variables, we obtain

$$h_i(k) = g_i(k) - g_{i*} = \sum_J V_i^{(J)} c_J \left(\frac{k}{k_0}\right)^{-\theta_J} \implies g_i(k) = g_{i*} + \sum_J c_J V_i^{(J)} \left(\frac{k}{k_0}\right)^{-\theta_J}\,. \tag{4.19}$$

We can distinguish two cases:

- Case 1:

$$\theta_I < 0 \implies \left(\frac{k}{k_0}\right)^{-\theta_I} \to 0\,, \qquad \text{for } k \ll k_0 \tag{4.20}$$

Looking at the solution of our differential equation (4.19), this implies that the corresponding constant of integration c_I (which is a free parameter) is suppressed in the expression for g_i at scales much smaller than the reference scale k_0. Hence, $\theta_I < 0$ indicates an irrelevant direction that is not connected to a free parameter.

We can connect this to our previous one-dimensional example. There, the stability matrix is a 1×1-matrix and the critical exponent is simply

$$\theta = -\frac{\partial \beta_G}{\partial g}\Big|_{G=G_*}, \tag{4.21}$$

i.e. the local slope at the fixed point. The case $\theta < 0$ signifies a positive slope and hence, as we saw above (e.g. in Figure 4.4), an irrelevant direction.

- Case 2:

$$\theta_I > 0 \implies \left(\frac{k}{k_0}\right)^{-\theta_I} \gg 0\,, \qquad \text{for } k \ll k_0 \tag{4.22}$$

Again looking at the solution of our differential equation (4.19), this implies that the corresponding c_I contributes to the expression for g_i at scales much smaller than the reference scale k_0. Hence, $\theta_I > 0$ indicates a relevant direction that is connected to a free parameter.

Example: Stability matrix of the g-G-system

As a direct example of a stability matrix, we can compute that of the system of β-functions (4.8) we considered above. We find without yet specifying the fixed point

$$M(G,g) = \begin{pmatrix} 2 - 2\#_1 G & 0 \\ -g & -G + 3\#_2 g^2 \end{pmatrix}. \tag{4.23}$$

Now, if we consider the FP C, where we recall $(G,g) = (\frac{2}{\#_1}, \sqrt{\frac{2}{\#_1\#_2}})$ we find the stability matrix

$$M\left(\frac{2}{\#_1}, \sqrt{\frac{2}{\#_1\beta_1}}\right) = \begin{pmatrix} -2 & 0 \\ -\sqrt{\frac{2}{\#_1\#_2}} & \frac{4}{\#_1} \end{pmatrix}. \tag{4.24}$$

We notice that the determinant of this matrix is negative. This already implies that one of the eigenvalues of this stability matrix is negative. This confirms our earlier finding that the FP C has one relevant direction (corresponding to the negative eigenvalue) and one irrelevant direction (corresponding to the positive eigenvalue). The two critical exponents (note the extra sign with respect to the eigenvalues) are $\theta_1 = 2$ and $\theta_2 = \frac{-4}{\#_1}$.

4.3 Why should we expect finitely many relevant couplings? – an argument using dimensional analysis

We now ask ourselves: Why can we expect only finitely many relevant couplings at an asymptotically safe fixed point? In practise, at an asymptotically safe fixed point, the number of relevant couplings is not determined a priori. There is, however, an argument based on dimensional analysis, which supports that there should only be a finite and relatively small number of free parameters. To understand the argument, we consider the structure of the stability matrix. The important terms for the argument are the dimensional terms in the beta functions, i.e.,

$$\beta_{g_i} = -d_{\bar{g}_i} g_i + \mathcal{O}\left(\left\{g_i^2\right\}\right) . \tag{4.25}$$

These result in entries on the diagonal of M,

$$M_{ii} = -d_{\bar{g}_i} + \mathcal{O}\left(\{g_{i*}\}\right) . \tag{4.26}$$

In contrast, the off-diagonal are of the same order as the terms we have neglected on the diagonal,

$$M_{ij,i\neq j} = \mathcal{O}\left(\{\partial_j g_{i*}\}\right) . \tag{4.27}$$

We find schematically

$$M = \begin{pmatrix} -d_{g_1} + ... & \dots & \dots \\ \vdots & -d_{g_2} + ... & \dots \\ \vdots & \dots & \ddots \end{pmatrix} . \tag{4.28}$$

We are free to order our vector of β-functions as we wish. Using this freedom, we can sort couplings by mass dimensions, such that they grow along the diagonal, i.e.

$$d_{g_i} \geq d_{g_{i+1}} \implies -d_{g_i} \leq -d_{g_{i+1}} , \qquad \forall i . \tag{4.29}$$

Now, in a general field theory, it holds that the higher the power of fields and derivatives in an interaction term, the more negative is the mass dimension of the coupling, i.e. the more negative is the value of d_{g_i}. To see this, let us consider an example.

Example: $\mathbb{Z}_2$-symmetric 4d scalar field theory

Let us consider the following action of a $\mathbb{Z}_2$-symmetric 4d scalar field theory:

$$S = \int \mathrm{d}^4x \left(-\frac{1}{2}\partial_\alpha\phi\partial^\alpha\phi + \sum_{i=1}^{\infty} G_{2i}\phi^{2i} + \sum_{j=1}^{\infty} H_{2j}\left(\partial_\beta\phi\partial^\beta\phi\right)\phi^{2j} + \cdots\right) . \tag{4.30}$$

From our unit conventions, we know the action is dimensionless and the integral measure has mass dimension $[\mathrm{d}^4x] = -4$. This implies that the derivative has mass dimension $[\partial] = 1$

and the scalar field $[\phi] = 1$. Using this, we find

$$[G_{2i}] = 4 - 2i\,, \qquad [H_{2j}] = -2j\,. \tag{4.31}$$

We thus see our claim confirmed in this simple example. For i and j increasing, the mass dimension of the coupling constants will become more negative.

Accordingly, the entries of M_{ij} follow the pattern

$$M_{ii} \to -d_{\bar{g}_i}\,, \qquad M_{ij,i\neq j} \to 0\,, \qquad i \gg 1\,, \tag{4.32}$$

if we assume that $\mathcal{O}\left(g_{i*}\right)$ is bounded. The observation we made regarding how the mass dimension of the coupling changes thus implies that the diagonal elements of the stability matrix will grow as we go along the diagonal. This implies for the critical exponents

$$\theta_i \to d_{g_i} < 0\,. \tag{4.33}$$

We hence expect finitely many relevant directions, unless g_{i*} grows without bound for higher-order g_i. Explicit examples of asymptotically safe fixed points confirm this argument.[iii]

4.4 Asymptotic Freedom: a special case of an RG fixed point

We now consider a special case of a fixed point: the Gaussian fixed point (GFP). It is the fixed point at $g_{i*} = 0$. This fixed point is always guaranteed, because a theory without interactions always stays a free, non-interacting theory. But, for a fixed point to be a viable at high k (i.e. an "ultraviolet" fixed point), the couplings that we want to be non-zero at low energies have to relevant at this fixed point. In other words, the mere existence of the Gaussian fixed point is insufficient to make a theory asymptotically free. To illustrate this, we contrast non-Abelian gauge theories with few and with many matter fields. Both have a Gaussian fixed point at vanishing gauge coupling, but only the former are asymptotically free.

We consider an SU(3) Yang-Mills theory with gauge coupling g and with N_f fermions in the fundamental representation. Because non-Abelian gauge bosons also couple to each other, there is a contribution from the gauge bosons and one from the fermions in the beta function

$$\beta_g = -\left(11 - \frac{2N_f}{3}\right)\frac{g^3}{16\pi^2} + \dots\,, \tag{4.34}$$

which boasts a GFP as expected. For $N_f < 16.5$, the first term dominates and thus the beta function is negative and the coupling grows from UV to IR. Thus, in this regime, the theory is asymptotically

[iii]This is of course just an argument for finitely many relevant directions, not a proof. There is always the possibility of a fixed point that is very strongly coupled, so that the $\mathcal{O}(g_{i\,*})$-terms are never negligible compared to the canonical dimension.

free. In contrast, for $N_f > 16.5$, the second term dominates and the beta function is positive and the coupling decreases from UV to IR. Therefore, the theory cannot be asymptotically free in this regime. This is similar to an Abelian gauge theory, e.g., for QED with gauge coupling e we have

$$\beta_e = \frac{e^3}{12\pi^2} + \dots \,, \tag{4.35}$$

which also boasts a GFP. We note that the leading order coefficient is **not** negative and hence that the coupling is marginally irrelevant. If $e = e_* = 0$ in the UV, then $e(k) \equiv 0$ necessarily. The scale dependence of the coupling is

$$e^2(\mu) = \frac{1}{\frac{1}{e^2(\mu_0)} - \frac{c}{2}\ln\left(\frac{\mu}{\mu_0}\right)} \,. \tag{4.36}$$

Thus $e^2(\mu) \nearrow 0$ as $\mu \to \infty$. This implies that QED is not asymptotically free.

4.5 Mechanisms to generate Asymptotic Safety

Asymptotic safety requires a zero of the beta function. Such a zero can be generated in different ways. The generic form of a β-function is

$$\beta_{g_i} = -d_{\bar{g}_i} g_i + \mathcal{O}\left(\left\{g_i^2, g_i\, g_j,\, g_j^2\right\}\right) \,. \tag{4.37}$$

The first term arises from the canonical dimension of the coupling (and vanishes for canonically marginal couplings); the second term denotes contributions from quantum fluctuations. These can arise from different degrees of freedom and can therefore depend on the coupling in question as well as the other couplings in the theory.

We see from the above form that we need to balance different terms in β-functions against each other to achieve asymptotic safety. There are different mechanisms to do so and we will look at them through simple examples of one or two couplings. The mechanisms are

1. one-loop vs. two-loop (for dimensionless couplings),
2. canonical scaling vs. quantum scaling (for couplings with $d_{\bar{g}_i} \neq 0$,
3. screening vs. antiscreening from different degrees of freedom.

Note that there is also some overlap between the mechanisms, e.g. 1. can be seen as an example of 3.

4.5.1 One-loop vs two-loop

Consider the β-function of the gauge coupling g in $\mathrm{SU}(N_c)$ Yang-Mills theory with N_f fermions in the fundamental representation. The β-function is

$$\beta_g = -\left(\frac{11}{3}N_c - \frac{2}{3}N_f\right)\frac{g^3}{16\pi^2} + \mathcal{O}(g^5)\,. \tag{4.38}$$

For $N_f < 11N_c/2$, the one-loop coefficient is negative, indicating that the coupling is **antiscreened** (c.f. discussion at the beginning of Section 4.2). For $N_f > 11N_c/2$, asymptotic freedom is lost. In this regime, it is possible to obtain asymptotic safety instead, if we add the two-loop term. Then, the β-function has the form

$$\beta_g = \left(B + C\frac{g^2}{16\pi^2}\right)\frac{g^2}{16\pi^2}\,, \qquad B > 0\,, \tag{4.39}$$

where B denotes the one-loop coefficient, as in Eq. (4.38) for N_f fermions in the fundamental representation, and C is the two-loop coefficient. Both B and C depend on the matter content (i.e., the number of fields as well as the representation of $SU(N_c)$ that they transform in). We write the beta function in this more general form, because it can be shown on general grounds that $C > 0$, when only fermions are present. This means that no non-trivial zero can be achieved with fermions only

· There is one way of achieving $C < 0$, and that is to also add scalars, so that there is also a Yukawa interaction in the theory, parameterized by the Yukawa coupling y. We analyse the theory in the Veneziano limit $N_f \to \infty$, $N_c \to \infty$ with

$$\epsilon = \frac{N_f}{N_c} - \frac{11}{2} \ll 1\,, \tag{4.40}$$

fixed. The reason to consider this limit is that the fixed point will be controlled by the parameter ϵ and therefore, despite being an interacting fixed point, can be made arbitrarily perturbative (i.e., close to the free fixed point). Then, it becomes an excellent approximation to neglect terms higher than 2 loop in our considerations. In the Veneziano limit, we have to rescale the couplings in the theory, because the beta functions for the original couplings feature explicit factors of N_c and N_f and therefore diverge in this limit. We define

$$\hat{\alpha}_y = \frac{y^2 N_c}{16\pi^2}\,, \qquad \hat{\alpha}_g = \frac{g^2 N_c}{16\pi^2}\,. \tag{4.41}$$

We work in terms of the squares of the couplings instead of the couplings purely for convenience, because the physics does not change under a change in sign of these two couplings. The scaling with N_c is chosen such that the right-hand sides of the beta functions are finite and nonzero for $N_c \to 0$. This requirement fixes the power of N_c with which we have to rescale. After rescaling, we find the β-functions

$$\beta_{\hat{\alpha}_g} = \hat{\alpha}_g^2\left(\frac{4}{3}\epsilon + \left(25 + \frac{26}{3}\epsilon\right)\hat{\alpha}_g - 2\left(\frac{11}{2} + \epsilon\right)^2\hat{\alpha}_y\right)\,, \tag{4.42}$$

where the negative sign allows a balance between the one-loop term and the combined two-loop terms, and

$$\beta_{\hat{\alpha}_y} = \hat{\alpha}_y\left((13 + 2\epsilon)\,\hat{\alpha}_y - 6\hat{\alpha}_g\right)\,, \tag{4.43}$$

where the negative sign comes from the screening vs. anti-screening mechanism.

Exercise: Computing the non-trivial fixed points and the critical exponents

As we are only considering non-trivial fixed points, we consider $\hat{\alpha}_{y*,g*} \neq 0$. Then solving $\beta_{\hat{\alpha}_y} = 0$, we find

$$\hat{\alpha}_{y*} = \frac{6}{13}\hat{\alpha}_{g*} + \mathcal{O}(\epsilon) \ . \tag{4.44}$$

Setting $\beta_{\hat{\alpha}_g} = 0$, we find

$$0 = \frac{4}{3}\epsilon + 25\hat{\alpha}_{g*} - \frac{121}{2}\hat{\alpha}_{y*} + \mathcal{O}\left(\epsilon^2\right) \ . \tag{4.45}$$

Inserting (4.44) into (4.45), we find

$$\hat{\alpha}_{g*} = \frac{26}{57}\epsilon \, , \tag{4.46}$$

which implies

$$\hat{\alpha}_{y*} = \frac{4}{19}\epsilon \, . \tag{4.47}$$

Through these equations, we see that ϵ (i.e. the difference in N_f from the point $N_f = 11N_c/2$, where B switches sign) is a control parameter that makes $\hat{\alpha}_{g*}$ and $\hat{\alpha}_{y*}$ arbitrarily small. Let us now compute the stability matrix at the fixed point $(\hat{\alpha}_{g*}, \hat{\alpha}_{y*}) = (\frac{26}{57}\epsilon, \frac{4}{19}\epsilon)$. To order $\mathcal{O}(\epsilon^2)$, we find

$$\left.\frac{\partial\beta_{\hat{\alpha}_g}}{\partial\hat{\alpha}_g}\right|_* \approx 25\hat{\alpha}_{g*}^2 \, , \tag{4.48}$$

$$\left.\frac{\partial\beta_{\hat{\alpha}_g}}{\partial\hat{\alpha}_y}\right|_* \approx -\frac{121}{2}\hat{\alpha}_{g*}^2 \, , \tag{4.49}$$

$$\left.\frac{\partial\beta_{\hat{\alpha}_y}}{\partial\hat{\alpha}_y}\right|_* \approx 13\hat{\alpha}_{y*} = 6\hat{\alpha}_{g*} \, , \tag{4.50}$$

$$\left.\frac{\partial\beta_{\hat{\alpha}_y}}{\partial\hat{\alpha}_g}\right|_* \approx -6\hat{\alpha}_{y*} = -\frac{36}{13}\hat{\alpha}_{g*} \, , \tag{4.51}$$

where we made sure to only include terms of order $\mathcal{O}(\epsilon^2)$ at most. Hence, to order $\mathcal{O}(\epsilon^2)$ we find

$$M \approx \begin{pmatrix} 25\hat{\alpha}_{g*}^2 & -\frac{121}{2}\hat{\alpha}_{g*}^2 \\ -\frac{36}{13}\hat{\alpha}_{g*} & 6\hat{\alpha}_{g*} \end{pmatrix} . \tag{4.52}$$

The eigenvalues of this matrix are related to its determinant and its trace, for these we find

$$\det M = \theta_1 \cdot \theta_2 \approx -\frac{228}{13}\hat{\alpha}_{g*}^3 \, , \qquad \operatorname{tr} M = -(\theta_1 + \theta_2) = 6\hat{\alpha}_{g*} + 25\hat{\alpha}_{g*}^2 \, . \tag{4.53}$$

Solving this system of equations for the critical exponents leads to

$$\theta_1 = \frac{104}{171}\epsilon^2 \, , \qquad \theta_2 = -\frac{52}{19}\epsilon \, , \tag{4.54}$$

which means that $\hat{\alpha}_y$ is predicted in terms of $\hat{\alpha}_g$.

For further reading on the model, see Litim and Sannino 2014. Also for a review on Asymptotic Safety and its mechanisms, see Eichhorn 2019.

4.5.2 Canonical scaling vs. quantum scaling

In this mechanism, the leading-order, linear term in the beta function for a coupling with $d_{\bar{g}} \neq 0$ cancels with all terms from quantum fluctuations.

As an example, consider a β-function that encodes asymptotic freedom, e.g. Yang-Mills theory in $d = 4$:

$$\beta_g = -\frac{11}{3}\frac{N_c}{16\pi^2} g^3 + \dots \,. \tag{4.55}$$

Now consider the theory in $d > 4$ by setting $d = 4 + \epsilon$ for $\epsilon > 0$. The mass dimension of the gauge coupling is

$$[\bar{g}] = \frac{4-d}{2}\,, \tag{4.56}$$

which is derived the same way we found the the mass dimension of the couplings G_{2i} and H_{2i} above. Here, we use that for the gauge field, one finds $[A] = (d-2)/2$. In $d > 4$, the coupling therefore has negative mass dimension (which implies that the theory is no longer perturbatively renormalizable). We introduce the dimensionless coupling

$$g = \bar{g}\, k^{\frac{d-4}{2}}\,. \tag{4.57}$$

Inserting this into the expression for the β-function, we find

$$\beta_g = k\partial_k\left(\bar{g}k^{\frac{d-4}{2}}\right) = \frac{d-4}{2}g + k^{\frac{d-4}{2}}k\partial_k\bar{g} = \frac{\epsilon}{2}g + k^{\frac{\epsilon}{2}}\beta_{\bar{g}} \approx \frac{\epsilon}{2}g - \frac{11}{3}\frac{N_c}{16\pi^2}g^3\,, \tag{4.58}$$

where in the last step we used that for $\epsilon \ll 1$, the one-loop term in the $\beta_{\bar{g}}$-function stays the same as in $d = 4$ to leading order in ϵ. The resulting β-function contains a canonical scaling dimension, which leads to an increasing term (first term) and quantum fluctuations that antiscreen the coupling. We find the fixed point

$$g_* = \sqrt{\frac{16\pi^2}{11N_c}\frac{3\epsilon}{2}}\,. \tag{4.59}$$

In Yang-Mills theory, it is an open question, how large ϵ can be taken, see Eichhorn 2019 for further references to studies in $d = 5$.

This mechanism is also relevant for gravity, because in d dimensions we have $[G_N] = d - 2$ and thus

$$\beta_G = (d-2)\,G + \mathcal{O}(G^2)\,. \tag{4.60}$$

Thus, if there is asymptotic safety, it must be related to an antiscreening term from quantum fluctuations that can compensate the linear term from the canonical scaling. In particular, in $d = 2 + \epsilon$, the situation is very similar to that in Yang-Mills theory, and studies to $\mathcal{O}(\epsilon^3)$ show an asymptotically safe fixed point, see Eichhorn 2019 for further references.

4.6 Computing β-functions with the Functional Renormalization Group

So far we have considered the β-functions of multiple quantum field theories without explaining how they are computed. This we will do now within the effective action framework of QFT.
The effective action $\Gamma[\Phi]$ is defined as the quantum analogue of the classical action $S[\Phi]$, in the sense that the variation of $\Gamma[\Phi]$ with respect to the field produces the field equations. In the presence of quantum fluctuations, equations of motion for a "classical" field configuration ϕ do not have much physical meaning[iv]. Instead, we are looking for an object that provides equations of motion for the expectation value $\Phi := \langle\phi\rangle$, i.e. we would like

$$\frac{\delta\Gamma[\Phi]}{\delta\Phi} = J\,, \tag{4.61}$$

for some source field J. This motivates a definition along the lines of

$$\exp\left(i\Gamma[\Phi] + i\int \mathrm{d}^4x J(x)\Phi(x)\right) = \int \mathcal{D}\phi \exp\left(iS[\phi] + i\int \mathrm{d}^4x J(x)\phi(x)\right)\,, \tag{4.62}$$

as going from the classical action S on the RHS to the effective action Γ on the LHS entails integrating over all field configurations. More precisely, the effective action $\Gamma[\Phi]$ and the generating functional $Z[J]$ are related by a Legendre transform:

$$\Gamma[\Phi] = \sup_J\left(\int \mathrm{d}^4x J(x)\Phi(x) - \ln Z[J]\right) = \int \mathrm{d}^4x J_{\mathsf{sup}}(x)\Phi(x) - \ln Z[J_{\mathsf{sup}}]\,, \tag{4.63}$$

where in the second term J_{sup} is understood as a field-dependent source $J_{\mathsf{sup}}[\Phi]$ and $Z[J]$ denotes the generating functional, given by

$$Z[J] = \int \mathcal{D}\phi \exp\left(iS[\phi] + i\int \mathrm{d}^4x J(x)\phi(x)\right)\,. \tag{4.64}$$

By defining $W[J] := \ln Z[J]$, we conclude from (4.63) that $\Gamma[\Phi]$ is a Legendre transform of $W[J]$. This is similar to relations between partition function and free energy in statistical physics. Note that the supremum in the definition of $\Gamma[\Phi]$ guarantees that Γ is convex and has a well-defined Legendre transform. Let us now see if this definition corresponds to our expectation from (4.61):

$$\begin{aligned}
\frac{\delta\Gamma[\Phi]}{\delta\Phi(y)} &= \int \mathrm{d}^4x \left(\frac{\delta J(x)}{\delta\Phi(y)}\Phi(x) + J(x)\frac{\delta\Phi(x)}{\delta\Phi(y)}\right) - \frac{\delta \ln Z[J]}{\delta\Phi(y)} && (4.65)\\
&= \int \mathrm{d}^4x \left(\frac{\delta J(x)}{\delta\Phi(y)}\Phi(x) + J(x)\delta^{(4)}(x-y)\right) - \int \mathrm{d}^4x \frac{\delta \ln Z[J]}{\delta J(x)}\frac{\delta J(x)}{\delta\Phi(y)} && (4.66)\\
&= \int \mathrm{d}^4x \frac{\delta J}{\delta\Phi(y)}\Phi(x) + J(y) - \int \mathrm{d}^4x\; \Phi(x)\frac{\delta J(x)}{\delta\Phi(y)} && (4.67)\\
&= J(y)\,. && (4.68)
\end{aligned}$$

where crucially, in the first step, the supremum in the definition of the effective action forces us to also take a functional derivative of the source field. We thus see that $\Gamma[\Phi]$ indeed does provide equations of motion for Φ in analogy to $S[\phi]$ for ϕ.

[iv]This can be seen by considering that the path integral $\int \mathcal{D}\phi\, e^{iS[\phi]}$ does not pick out a single field configuration ϕ, but sums over all of them.

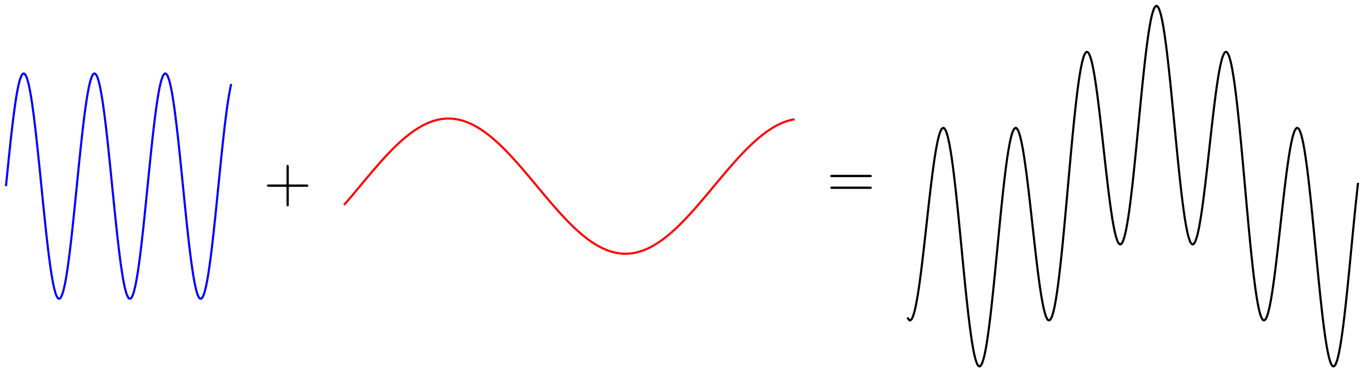


Figure 4.6: *From Fourier analysis, we know that one can decompose a wave into a superposition of constituent waves. This is the intuition behind the infrared cut-off $\Delta S_k[\phi]$ we are introducing into the generating functional Z_k. It is meant to suppress the low momentum modes (red) such that the high momentum modes (blue) are integrated out.*

In general, the effective action contains all possible monomials of the field and its derivatives that are allowed by the symmetries of the theory. The underlying reason is that even if one starts with just a subset of interactions at a microscopic scale, quantum fluctuations generate new, effective interactions, i.e., upon integrating out microscopic field configurations, all additional interactions that are compatible with the symmetries are generated. If we write the interaction monomials (which depend on the field and its derivatives) as $\mathcal{O}_i$, then we can write

$$\Gamma[\Phi] = \sum_i \int \mathrm{d}^4x\, \bar{g}_i \mathcal{O}^i \,, \tag{4.69}$$

where the $\bar{g}_i$ are couplings containing the effects of all quantum fluctuations. Microscopically, many of the $\bar{g}_i$ may be zero, but then deviate from zero at lower RG scales. We are interested in how the $\bar{g}_i$ change, when we account for quantum fluctuations not all at once, but step by step, starting with very large momenta first (see Fig. 4.6). To do so, we introduce an infrared cutoff into the path integral, such that field configurations with small-scale structure are integrated over and field configurations that only vary slowly (i.e., are only composed out of low-wavelength Fourier modes) are not.

This cannot be done in Lorentzian signature. If we were to only consider scales $p^2 > k^2$ in our path integral, then this condition translates to $p^2 = E^2 - \vec{p}^{\,2} > k^2$, which even for $p^2 \approx 0$ does not imply that the components are small, E^2 and $\vec{p}^{\,2}$ could still be very large. To fix this, one could think to cutoff E^2 or $\vec{p}^{\,2}$ separately but this breaks Lorentz invariance.

We therefore work with a Euclidean generating functional. We also add an ultraviolet cutoff Λ to make the latter well-defined. Hence, we define

$$Z_k = \int_\Lambda \mathcal{D}\phi \exp\left(-S[\phi] - \Delta S_k[\phi] + \int \mathrm{d}^4x J(x)\phi(x)\right) . \tag{4.70}$$

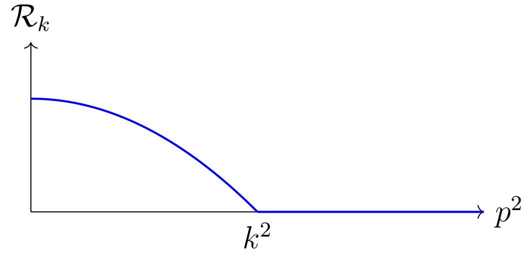


Figure 4.7: *Schematically, the regulator $\mathcal{R}_k$ takes the above form to achieve our goal of integrating out only modes above the infrared scale k^2. The exact form for the regulator for $p^2 < k^2$ is not fixed and there is freedom in the actual choice of function, but we impose $\mathcal{R}_k > 0$ for $p^2 < k^2$ and $\mathcal{R}_k = 0$ for $p^2 > k^2$.*

Here, we introduced

$$\Delta S_k[\phi] = \frac{1}{2}\int \frac{\mathrm{d}^4 p}{(2\pi)^4}\phi(p)\mathcal{R}_k(p^2)\phi(-p)\,, \tag{4.71}$$

where $\mathcal{R}_k(p^2)$ is the regulator. From the way it is inserted into the path integral (i.e. as part of a term quadratic in the field), it acts like a p^2-dependent and infrared cutoff scale k^2-dependent mass. In the Euclidean path integral, such a term suppresses the corresponding field. As we are interesting in our path integral only “seeing” the high momentum modes for now, we would like to suppress all modes below the infrared cut-off scale k^2, meaning that we would like our regulator to only kick in at scales smaller than k^2 (see Figure 4.7). We assume that each field configuration $\phi(x)$ in the path integral can be decomposed into its Fourier modes; the introduction of $\mathcal{R}_k(p^2)$ entails that the Fourier modes with $p^2 > k^2$ are integrated out in the path integral and those with $p^2 < k^2$ are not. Ultimately, we want to remove the regulator, after it has served its purpose of allowing us to do the path integral “momentum-shell wise”. Thus, we require

$$\lim_{k^2/p^2\to 0}\mathcal{R}_k(p^2) = 0\,. \tag{4.72}$$

This condition also implies that the regulator $\mathcal{R}_k$ vanishes for any $p^2 > k^2$, but in particular, $\mathcal{R}_k$ vanishes once all quantum fluctuations are integrated over. Now define

$$\Gamma_k[\Phi] = \sup_J\left(\int \mathrm{d}^4 x J(x)\Phi(x) - \ln Z_k\right) - \Delta S_k[\Phi]\,, \tag{4.73}$$

where we have added a term compared to our original definition for the effective action (4.63) to remove the “auxiliary” mass term we added in (4.70) to suppress low momentum modes. From this definition, we see

$$\Gamma_k \xrightarrow{k^2/\Lambda^2\to 0} \Gamma\,, \tag{4.74}$$

as expected when we effectively remove the regulator and integrate over all the possible field configurations.

In the limit $k^2 \to \Lambda^2 \to \infty$, we require $\Gamma \to S$, so that we can think Γ_k as an "interpolator" between effective action and the classical (or microscopic) action, i.e.

$$\Gamma \to \Gamma_k \to S\,. \tag{4.75}$$

To achieve this, we need

$$\lim_{\Lambda^2\to\infty} \lim_{k^2\to\Lambda^2} \mathcal{R}_k \to \infty\,. \tag{4.76}$$

Exercise: Showing $\Gamma_k \to S$ in the limit $k^2 \to \Lambda^2 \to \infty$

We first write, taking the exponential of the average effective active written in (4.73),

$$\exp\left(-\Gamma_k[\Phi]\right) = \exp\left(-\int \mathrm{d}^4x J_{\mathsf{sup}}(x)\Phi(x)\right) Z_k[J_{\mathsf{sup}}] \exp\left(\Delta S_k[\Phi]\right)\,. \tag{4.77}$$

Inserting Z_k from (4.70) and using that $\Delta S_k[\Phi] - \Delta S_k[\phi] = \Delta S_k[\Phi - \phi]$, we find

$$\exp\left(-\Gamma_k[\Phi]\right) = \int \mathcal{D}\phi\ \exp\left(-S[\phi] - \Delta S_k[\phi - \Phi] + \int \mathrm{d}^4x J_{\mathsf{sup}}(x)(\phi(x) - \Phi)\right)\,. \tag{4.78}$$

Before proceeding, first recall the formula

$$f(x_0) = \lim_{\alpha\to 0} \int_{-\infty}^{+\infty} \mathrm{d}x\ f(x) \exp\left(-\frac{(x-x_0)^2}{\alpha}\right)\,. \tag{4.79}$$

In our case, $f(x)$ is replaced by all the terms in generating function that are not second order in the fields and α is replaced bu $1/\mathcal{R}_k$. With this we can conclude

$$\lim_{\mathcal{R}_k\to\infty} \int \mathcal{D}\phi\ \exp\left(-S[\phi] + \int \mathrm{d}^4x J_{\mathsf{sup}}(x)(\phi(x) - \Phi)\right) \exp\left(-\frac{1}{2}\int \frac{\mathrm{d}^4p}{(2\pi)^4}\mathcal{R}_k(\phi - \Phi)^2\right) \tag{4.80}$$

$$= \exp\left(-S[\Phi] + \int \mathrm{d}^4x J_{\mathsf{sup}}(x)(\Phi - \Phi)\right) \tag{4.81}$$

$$= \exp\left(-S[\Phi]\right)\,. \tag{4.82}$$

Hence, in this particular limit, we do indeed find that the regularized effective action Γ_k becomes the classical action.

Just as we did before for Γ, we can expand

$$\Gamma_k = \sum_i \int \mathrm{d}^4x\ \bar{g}_k^i \mathcal{O}_i\,, \tag{4.83}$$

where the k-dependence is carried by the couplings. We are can compute the β-functions

$$\beta_{g_i} = k\partial_k\left(\bar{g}_i(k) k^{-d_{\bar{g}_i}}\right)\,, \tag{4.84}$$

from the flowing effective action Γ_k. We do so by isolating the appropriate monomial in field space and taking a k-derivative of the flowing effective action Γ_k, i.e.

$$\beta_{g_i} = k\partial_k \Gamma_k\Big|_{\mathcal{O}_i}\,. \tag{4.85}$$

Let us see whether we can derive a useful equation for $k\partial_k\Gamma_k$, from which we could indeed extract the beta functions in this way. Starting from the expression in (4.73), we can write[v]

$$k\partial_k\Gamma_k = \int \mathrm{d}^4x\,(k\partial_k J)\,\Phi - \frac{1}{Z_k}k\partial_k Z_k\Big|_\Phi - \frac{1}{2}\int \frac{\mathrm{d}^4p}{(2\pi)^4}\,(k\partial_k\mathcal{R}_k)\,\Phi(p)\Phi(-p)\,, \tag{4.86}$$

where the supremum in (4.73) forces us to take a derivative of the source field J in the first term. This is because the supremum changes as modes are integrated out. Notice that this will also affect how we deal with the second term, as $Z_k = Z_k[J]$. Indeed, it holds for the latter that

$$\frac{1}{Z_k}k\partial_k Z_k\Big|_\Phi = \frac{1}{Z_k}k\partial_k Z_k\Big|_J + \frac{1}{Z_k}\int \mathcal{D}\phi\; e^{-S[\phi]+\int \mathrm{d}^4x J\phi - \Delta S_k[\phi]}\int \mathrm{d}^4x\,(k\partial_k J)\,\phi \tag{4.87}$$

$$= \frac{1}{Z_k}k\partial_k Z_k\Big|_J + \int \mathrm{d}^4x\,(k\partial_k J)\,\langle\phi\rangle\,. \tag{4.88}$$

We notice that as $\langle\phi\rangle = \Phi$, the second term cancels with the first term of (4.86), giving us the result

$$k\partial_k\Gamma_k = -\frac{1}{Z_k}k\partial_k Z_k\Big|_J - \frac{1}{2}\int \frac{\mathrm{d}^4p}{(2\pi)^4}\,(k\partial_k\mathcal{R}_k)\,\Phi(p)\Phi(-p)\,. \tag{4.89}$$

This expression still makes reference to the classical action $S[\phi]$, which is not useful when we want to use the β_{g_i} to search for asymptotic safety, where we do not know $S[\phi]$. Hence, our next task is to eliminate the dependence on the classical action of the first term above. We start by writing

$$-\frac{1}{Z_k}k\partial_k Z_k\Big|_J = -\frac{1}{Z_k}k\partial_k\int \mathcal{D}\phi\; \exp\left(-S[\phi] + \int \mathrm{d}^4x J\phi - \frac{1}{2}\int \frac{\mathrm{d}^4p}{(2\pi)^4}\phi(p)\mathcal{R}_k\phi(-p)\right)\Bigg|_J \tag{4.90}$$

$$= \frac{1}{Z_k}\int \mathcal{D}\phi\left(\frac{1}{2}\int \frac{\mathrm{d}^4p}{(2\pi)^4}\phi(p)\,(k\partial_k\mathcal{R}_k)\,\phi(-p)\right)e^{-S[\phi]-\int \mathrm{d}^4x\; J\phi - \Delta S_k[\phi]} \tag{4.91}$$

$$= \frac{1}{2}\int \frac{\mathrm{d}^4p}{(2\pi)^4}\,(k\partial_k\mathcal{R}_k)\,\frac{\delta^2 \ln Z_k}{\delta J(p)\delta J(-p)} \tag{4.92}$$

$$= \frac{1}{2}\int \frac{\mathrm{d}^4p}{(2\pi)^4}\,(k\partial_k\mathcal{R}_k)\left(\langle\phi(p)\phi(-p)\rangle - \langle\phi(p)\rangle\langle\phi(-p)\rangle\right)\,, \tag{4.93}$$

where in the last steps we used that $\ln Z_k$, or more generally, $\ln Z$ is the generating functional of connected correlation functions.

We will now use that $\Gamma_k^2 + R_k$ is the inverse of $\frac{\delta^2 \ln Z_k}{\delta J(p)\delta J(-p)}$, where this denotes two functional derivatives of the regularized effective action with respect to the field Φ. As an equation,

$$\frac{\delta^2 \ln Z_k}{\delta J(p)\delta J(-p)} = \left(\frac{\delta^2\Gamma_k}{\delta\Phi^2} + \mathcal{R}_k\right)^{-1} =: \left(\Gamma_k^{(2)} + \mathcal{R}_k\right)^{-1}\,, \tag{4.94}$$

We can show this by rewriting

$$\int \mathrm{d}^4y\,\frac{\delta^2 \ln Z_k}{\delta J(y)\delta J(x')}\left(\Gamma_k^{(2)} + \mathcal{R}_k\right) = \delta(x-x')\,. \tag{4.95}$$

[v]In the following, we use $\big|_\Phi$ to denote an expression taken at fixed Φ.

Starting from the rhs, we write

$$\delta(x-x') = \frac{\delta J(x)}{\delta J(x')} \tag{4.96}$$

$$= \int \mathrm{d}^4 y \frac{\delta J(x)}{\delta \Phi(y)} \frac{\delta \Phi(y)}{\delta J(x')} \tag{4.97}$$

$$= \int \mathrm{d}^4 y \frac{\delta J(x)}{\delta \Phi(y)} \frac{\delta}{\delta J(x')} \frac{\delta \ln Z_k}{\delta J(y)} \,. \tag{4.98}$$

All that remains is to rewrite the first functional derivative into the desired form. We can do this by first noticing that we can use the dependence of the source field $J(x)$ on the regularized action, i.e.,

$$\int \mathrm{d}^4 w J(w)\Phi(w) = \Gamma_k + \ln Z_k + \frac{1}{2}\int \frac{\mathrm{d}^4 p}{(2\pi)^4} \Phi(p)\mathcal{R}_k \Phi(-p)\,, \tag{4.99}$$

to write the source field as a functional derivative of the rhs of the above equation. Doing so, we obtain

$$\frac{\delta J(x)}{\delta \Phi(y)} = \frac{\delta}{\delta \Phi(y)} \left(\frac{\delta \Gamma_k}{\delta \Phi(x)} + \mathcal{R}_k \Phi(-x) \right) = \frac{\delta^2 \Gamma_k}{\delta \Phi^2} + \mathcal{R}_k\,, \tag{4.100}$$

where we used that the regularized generating functional Z_k has no explicit dependence on the field Φ and the product rule in the third term.

Wetterich equation

By combining (4.92) and (4.94), we obtain

$$k\partial_k \Gamma_k = \frac{1}{2}\int \frac{\mathrm{d}^4 p}{(2\pi)^4} \left(k\partial_k \mathcal{R}_k\right) \left(\Gamma_k^{(2)} + \mathcal{R}_k\right)^{-1} \,. \tag{4.101}$$

We make a few important observations on the Wetterich equation.
First, we note of all that all reference to the classical action $S[\phi]$ has disappeared, as hoped.
Secondly, the Wetterich equation does not require us to calculate a path integral, but is instead a functional differential equation.
Third, we observe that the regulator $\mathcal{R}_k$ we had introduced to integrate out high-energy modes, generates an IR cutoff as well as an UV cutoff. This can be best seen considering two parts of the rhs of the Wetterich equation separately. First, note that per our definitions $k\partial_k \mathcal{R}_k(p^2) = 0$ for $p^2 > k^2$, i.e. the integral is cut off at $p^2 > k^2$. (Consequently, we can discard any UV cutoff Λ^2.) Now, note that the second part of the integrand is just a propagator with $\mathcal{R}_k$ playing the role of an auxiliary mass, which is making the integral IR finite. (Even for massless modes, this auxiliary mass is present, as it is not a physical mass.) Thus, the main contribution to the integral, and consequently to the change of Γ_k will come from modes with $p^2 \approx k^2$.
Fourth, the Wetterich equation is formally exact, as we made no approximation in its derivation.

Also note that the structure of the rhs is that of a loop integral, meaning that we have in the Wetterich equation an **exact** one-loop equation. Nevertheless, in practice the results achieved with this flow equation are not exact, because Γ_k contains infinitely many monomials that we cannot all account for at the same time. One consequently needs to **truncate** our ansatz for Γ_k.

Appendix A

Causal Set Theory

A.1 Motivation

We have seen in Section 2 how General Relativity is not perturbatively quantisable. We can conclude that this signifies a breakdown of Quantum Field Theory and more generally, continuous physics. We can thus postulate that spacetime is discrete. In this way, we would consider GR analogously to hydrodynamics, as an effective description in a certain limit. Another motivation for a discrete approach to gravity comes from the calculation by Bekenstein 1973 and Stephen W. Hawking 1976 of the entropy of a black hole. Indeed, they famously find

$$S_{\text{BH}} = \frac{A}{4l_{\text{pl}}^2}\,, \tag{A.1}$$

where A is the area of the black hole and l_{pl} the Planck length. Consider now a tiling of the black hole horizon with l_{pl}^2-sized tiles and put one bit of information on each. The entropy in the Boltzmannian sense is

$$S = \ln W\,, \tag{A.2}$$

where W is the multiplicity of the macrostate, i.e. the number of microstates that recover a certain macrostate. In our case, it holds that $W = 2^n$ for n the number of tiles. This is because there are two possible outcomes for each tile, namely containing a bit of information or containing no bit of information. Inserting this into (A.2)

$$S = n\ln 2 = \frac{A}{l_{\text{pl}}^2}\ln 2 \sim S_{\text{BH}}\,. \tag{A.3}$$

We notice that the entropy has the same scaling as the Bekenstein-Hawking entropy of a black hole, indicating that spacetime could be fundamentally discrete.

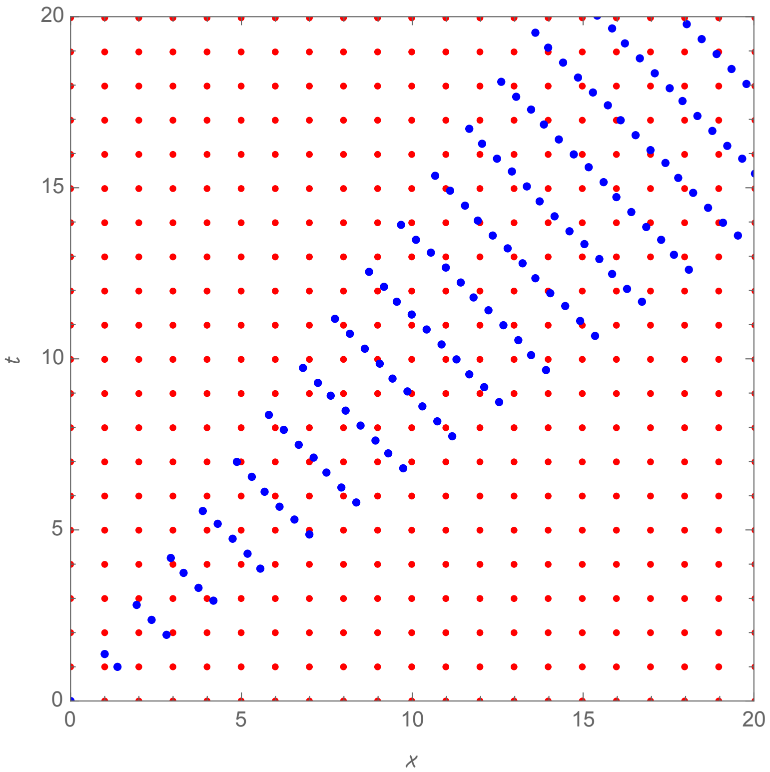


Figure A.1: *Here we plotted the effect of a boost with $\beta = 0.7$ on a regular lattice. The red points constitute a regular lattice whereas the blue points are the boosted points. We see that the regularity of the lattice is lost after performing the boost.*

A.2 Spacetime discreteness and Lorentz invariance

There are observational constraints of Lorentz symmetry. These consider e.g. modifications of the dispersion relation for photons

$$\vec{p}^{\,2} = E^2 \left(1 + \gamma \frac{E}{m_{\text{pl}}} + \cdots \right) , \tag{A.4}$$

by looking at distant γ-ray bursts (see Yang, Bi, and Yin 2024). These modifications are constrained to $|\gamma| > 1$, indicating that any changes would appear only at the Planck scale. Note that while Lorentz invariance is only a local symmetry on a general curved spacetime, it is a global symmetry of Minkowski spacetime. Hence, on a discrete version of Minkowski spacetime, Lorentz invariance should be preserved. As a first attempt, one could consider a regular lattice as a discretisation of Minkowski spacetime. This is shown in Figure A.1. The following becomes clear: Any regular distribution of spacetime points singles out a frame. Hence, the correspondence between a discrete spacetime and a continuous manifold must be based on a random distribution of spacetime points. The main ingredient of this correspondence is based on a Poisson-distribution of discrete spacetime

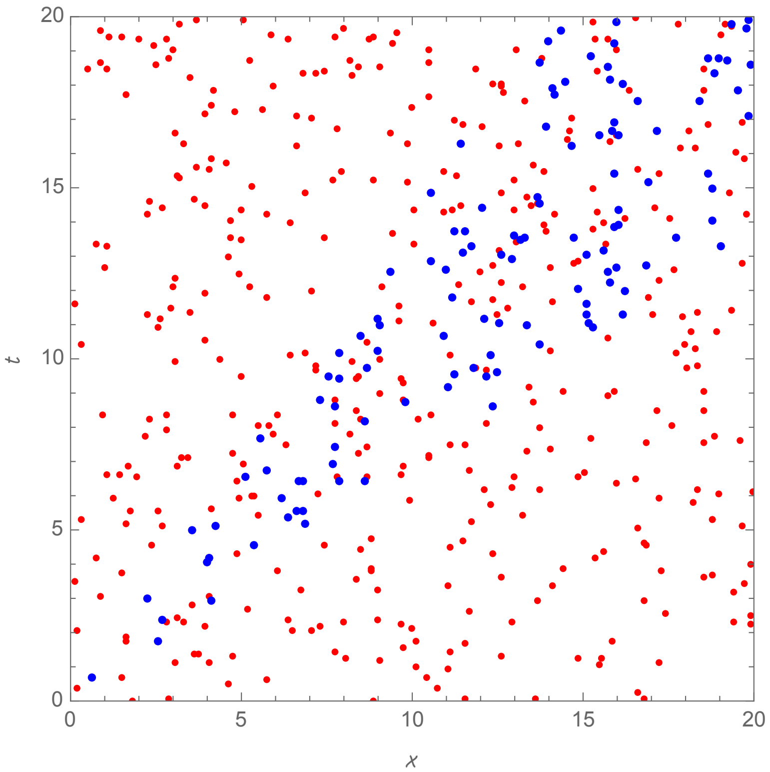


Figure A.2: *Here we plotted the effect of a boost with $\beta = 0.7$ on a random array of points. The red points are found by sampling random real numbers from the interval $[0, 20]^2$ whereas the blue points are the boosted points. Compared to the regular lattice case, we now see that post-boost, the array of points retains its random nature.*

points into a spacetime volume V, i.e.

$$P(n, V) = \frac{1}{n!}\left(\frac{V}{l^4}\right)^n \exp\left(-\frac{V}{l^4}\right) . \tag{A.5}$$

for n the number of points in the volume V. The procedure of randomly choosing points from a spacetime manifold is called a "sprinkling" in the causal set jargon (see Figure A.2). This is the first ingredient needed to construct a causal set.

A.3 Characterising a discrete geometry

From the study of General Relativity, we know that the metric encodes the causal structure of the spacetime, via the sign of the line-element $\mathrm{d}s^2$, as well as distances, via the modulus of the line-element $|\mathrm{d}s^2|$. This fact is made precise by the following theorem.

Theorem (Stephen W Hawking, King, and McCarthy 1976; Malament 1977). *If a causal bijection, i.e. a causal-order preserving bijection, f_b, exists between two d-dimensional spacetimes which are both* *future and past distinguishing*[i], *then these spacetimes are conformally isometric when* $d > 2$.

This theorem lays bare that the causal structure encodes the full information in a metric up to a local conformal factor stretching the metric $g_{\mu\nu} \rightarrow e^{\phi(x)^2} g_{\mu\nu}$. One says that in $d = 4$, the causal structure is (9/10)$^{\text{th}}$ of the metric, as only one of the ten components of the metric is not set by the causal structure itself. The latter is defined through a relation $\prec$, which allows to relate spacetime points to one another based on their causal relation. For example, for $i \neq j$, either $i \prec j$ or $j \prec i$ or i and j are spacelike separated. The causal structure is the second ingredient to construct a causal set.

Definition (Causal set)**.** *A set C of spacetime points i with an order relation $\prec$ (spoken "precedes") is a causal set if it is*

1. *Acyclic:* $i \prec j \wedge j \prec i \implies i = j \,, \forall i, j \in C$,
2. *Transitive:* $i \prec j \wedge j \prec k \implies i \prec k \,, \forall i, j, k \in C$,
3. *Locally finite:* $\forall i, j \in C \,, |\{j \in C \mid i \prec j \prec k\}| < \infty$, *where* $|\cdot|$ *denotes cardinality.*

The third assumption of local finiteness is what differentiates a causal set from causal continuous spacetimes.

A.4 Quantum Causal Set Theory

Our goal in this part will be to motivate the expression

$$\int \mathcal{D}C \; e^{iS[C]} \,, \tag{A.6}$$

as a starting point for a quantum theory of causal sets. First, let us motivate an action $S[C]$. Finding it is part of a larger attempt of reconstructing continuum quantities from a causal set, quantities such as the spacetime dimension, spatial topology, scalar curvature, etc. One can show that

$$\lim_{l \to 0} B\phi(x) = \left(\Box - \frac{1}{2} R(x)\right) \phi(x) \,, \tag{A.7}$$

for a scalar field $\phi(x)$ and an operator B we specify below. From this equation, we see that by setting $\phi(x) \equiv -2$, we find

$$\lim_{l \to 0} B(-2) = R(x) \,. \tag{A.8}$$

[i] This roughly means that if the future and past lightcones of two points are the same, then the points are also the same.

We find the Ricci scalar as the continuum limit of $B(-2)$. The operator B acting on a scalar field evaluated in the point$x \in C$ is defined as

$$B\phi(x) = \frac{4}{\sqrt{6}l^2}\left(-\phi(x) + \left(\sum_{y \in L_1} -9\sum_{y \in L_2} +16\sum_{y \in L_3} -8\sum_{y \in L_4}\right)\phi(y)\right), \tag{A.9}$$

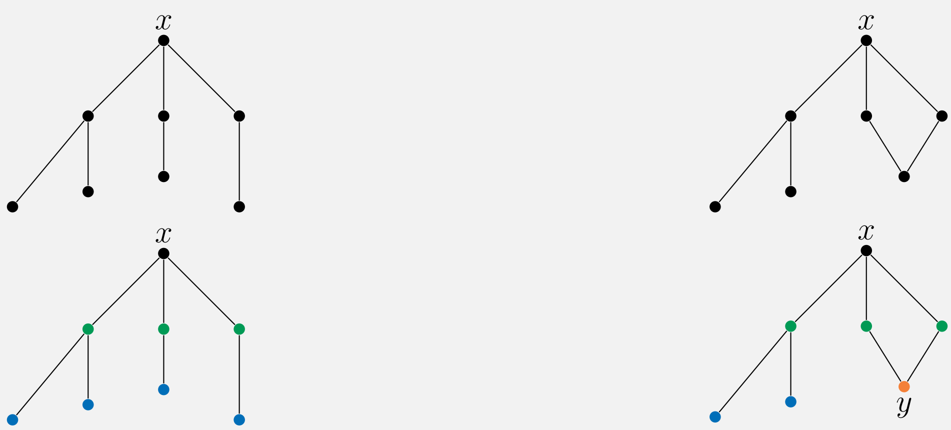
where L_i is the i-th layer. This is defined as

$$L_i(x) = \{y \in C \mid y \prec x \text{ and } |\{z \in C \mid y \prec z \prec x\}| = i - 1\}. \tag{A.10}$$

Thus, L_1 contains all the nearest neighbours to the past of $x \in C$, L_2 contains all the points that form a three-chain with x, and so on.

Exercise: Identifying the layers

We consider two causal sets and their layers in the scheme L_1, L_2, L_3.



We see that in the causal set on the rhs there are two intermediate points between the points x and y. This indicates that y does not satisfy the condition $|\{z \in C \mid y \prec z \prec x\}| = 1$ that points in the second layer L_2 must satisfy.

Of course, now that we have found an object that becomes the scalar curvature in the continuum limit, we have found a discrete analogue of the Einstein-Hilbert action, it is obtained by summing over the Ricci curvature of every point in the causal set. We get

$$S(C) = \sum_{i \in C} B_i(-2) = \frac{4}{\sqrt{6}}(n - N_1 + 9N_2 - 16N_3 + 8N_4), \tag{A.11}$$

where N_i is the total number of i-element intervals in C and n is the overall cardinality of the causal set, i.e. $|C| = n$.

Now we can turn to the measure

$$\int \mathcal{D}C. \tag{A.12}$$

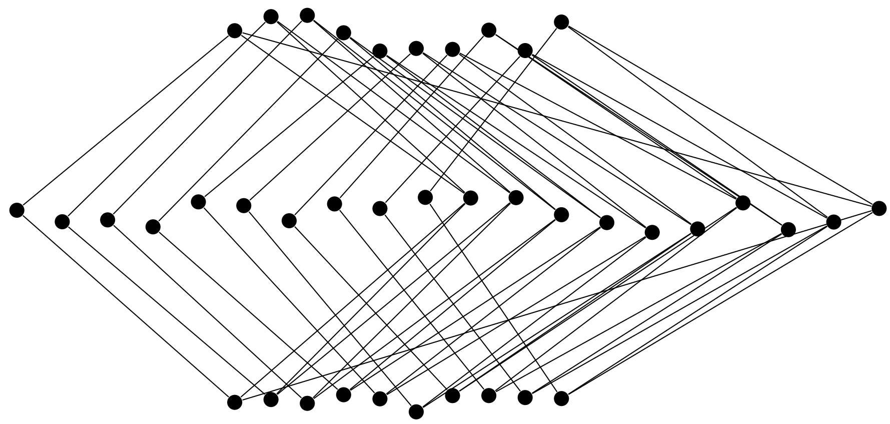


Figure A.3: *This is an example of a Kleitman-Rothschild causal set (or poset, more generally). Its defining feature is its three layers with double the points in the second layer relative to the first and to the third layer.*

Among the causal sets that are integrated over in this path-integral, there will be manifold-like causal sets, i.e. the ones we can obtain by sprinkling into a continuous manifold. But there will also be non manifold-like causal sets integrated over and these turn out to be a majority of the causal sets. An example of such non-manifold-like is the Kleitman-Rothschild causal set (see Figure A.3). The existence of such objects in Causal Set Theory makes one ask how radically should we depart from "well-tested" ingredients, such as General Relativity and Quantum Field Theory, in constructing quantum gravity. Indeed, Causal Set Theory exemplifies how difficult it is to get back well-tested physics at low-energies if we depart quite far.

Bibliography

Ballentine, L. E. (1982). "COMMENT ON 'INDIRECT EVIDENCE FOR QUANTUM GRAVITY'". In: *Phys. Rev. Lett.* 48, p. 522. doi: 10.1103/PhysRevLett.48.522.

Bekenstein, Jacob D. (1973). "Black Holes and Entropy". In: *Phys. Rev. D* 7, pp. 2333–2346. doi: 10.1103/PhysRevD.7.2333.

Buccio, Diego et al. (2024). *Physical running of couplings in quadratic gravity*. arXiv: 2403.02397 [hep-th]. url: https://arxiv.org/abs/2403.02397.

Donoghue, John F. (1994). "Leading quantum correction to the Newtonian potential". In: *Phys. Rev. Lett.* 72, pp. 2996–2999. doi: 10.1103/PhysRevLett.72.2996. arXiv: gr-qc/9310024.

Donoghue, John F. and Gabriel Menezes (2019). "Arrow of Causality and Quantum Gravity". In: *Phys. Rev. Lett.* 123.17, p. 171601. doi: 10.1103/PhysRevLett.123.171601. arXiv: 1908.04170 [hep-th].

Eichhorn, Astrid (2019). *An asymptotically safe guide to quantum gravity and matter*. arXiv: 1810.07615 [hep-th]. url: https://arxiv.org/abs/1810.07615.

Hawking, Stephen W, Andrew R King, and Patrick J McCarthy (1976). "A new topology for curved space–time which incorporates the causal, differential, and conformal structures". In: *Journal of mathematical physics* 17.2, pp. 174–181.

Hawking, Stephen W. (1976). "Black holes and thermodynamics". In: *Phys. Rev. D* 13, pp. 191–197. doi: 10.1103/PhysRevD.13.191.

Held, Aaron and Hyun Lim (2023). "Nonlinear evolution of quadratic gravity in 3+1 dimensions". In: *Phys. Rev. D* 108.10, p. 104025. doi: 10.1103/PhysRevD.108.104025. arXiv: 2306.04725 [gr-qc].

Held, Aaron and Jun Zhang (2023). "Instability of spherically symmetric black holes in quadratic gravity". In: *Phys. Rev. D* 107.6, p. 064060. doi: 10.1103/PhysRevD.107.064060. arXiv: 2209.01867 [gr-qc].

Levi, Michèle (2020). "Effective Field Theories of Post-Newtonian Gravity: A comprehensive review". In: *Rept. Prog. Phys.* 83.7, p. 075901. doi: 10.1088/1361-6633/ab12bc. arXiv: 1807.01699 [hep-th].

Litim, Daniel F. and Francesco Sannino (Dec. 2014). "Asymptotic safety guaranteed". In: *Journal of High Energy Physics* 2014.12. issn: 1029-8479. doi: 10.1007/jhep12(2014)178. url: http://dx.doi.org/10.1007/JHEP12(2014)178.

Lu, H. et al. (2015). “Black Holes in Higher-Derivative Gravity”. In: *Phys. Rev. Lett.* 114.17, p. 171601. doi: 10.1103/PhysRevLett.114.171601. arXiv: 1502.01028 [hep-th].

Malament, David B (1977). “The class of continuous timelike curves determines the topology of spacetime”. In: *Journal of mathematical physics* 18.7, pp. 1399–1404.

Page, Don N and CD Geilker (1981). “Indirect evidence for quantum gravity”. In: *Physical Review Letters* 47.14, p. 979.

Yang, Yu-Ming, Xiao-Jun Bi, and Peng-Fei Yin (Apr. 2024). “Constraints on Lorentz invariance violation from the LHAASO observation of GRB 221009A”. In: *Journal of Cosmology and Astroparticle Physics* 2024.04, p. 060. issn: 1475-7516. doi: 10.1088/1475-7516/2024/04/060. url: http://dx.doi.org/10.1088/1475-7516/2024/04/060.