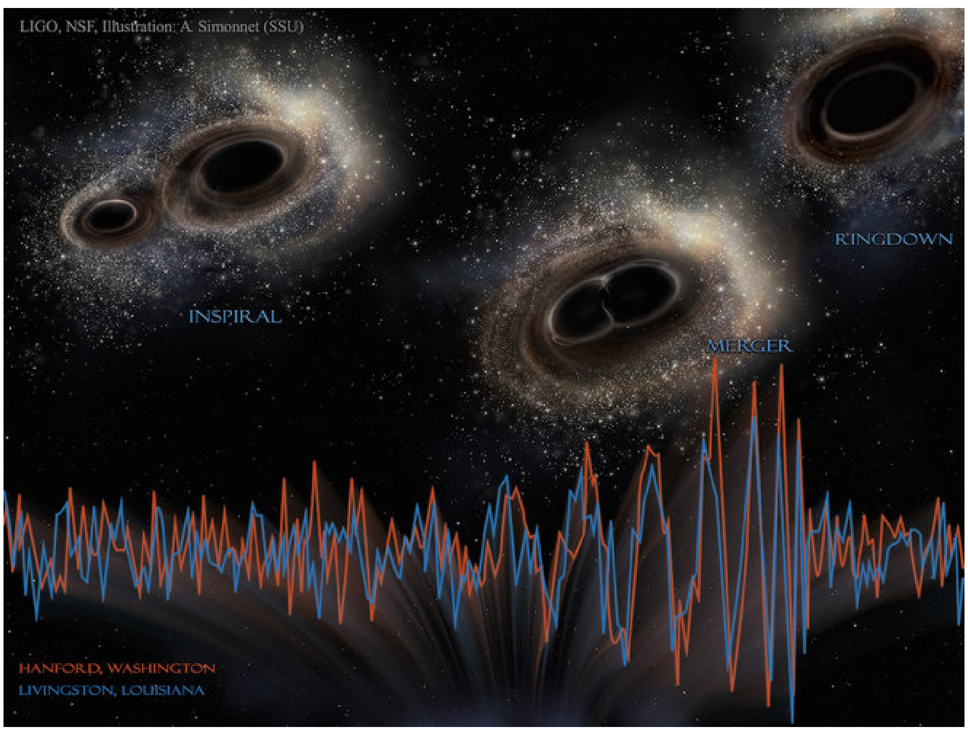
BLACK HOLES II

TU Wien, LVA 136.029

Daniel Grumiller



Notes for readers. These lecture notes were compiled from individual lecture sheets for my lectures "Black Holes II" at TU Wien in March–June 2018. As the name indicates, you are supposed to know the material from the lectures "Black Holes I" by the time you read these notes. Each lecture sheet is in principle self-contained, although there are some references to other lecture sheets. Whenever the equation numbering restarts at (1) this signifies a new lecture sheet. Here is an overview over the whole set of lectures and all lecture sheets:

1. *Horizons and other definitions* (p. 1-2): a brief vocabulary on the most relevant definitions, including the one of black holes

2. *Carter-Penrose diagrams* (p. 3-6): conformal compactifications and methods to construct Carter–Penrose diagrams, with flat space and AdS examples
   Raychaudhuri equation (p. 7-8): geodesic congruences and time-evolution of expansion
   Singularity theorems (p. 9-10): glimpse of singularity theorems by simple example; Hawking's area theorem

3. *Linearized Einstein equations* (p. 11-13): linearizing metric, Riemann tensor, field equations and action
   Gravitational waves (p. 14-15): gravitational waves in vacuum and action on test particles; emission
   QFT aspects of spin-2 particles (p. 16-18): massive and massless spin-2 fields and vDVZ-discontinuity

4. *Black hole perturbations* (p. 19-20, p. 23): scalar perturbations of Schwarzschild and generalizations
   Quasi-normal modes (p. 21-22): definition and applications of QNMs and guide to further literature

5. *Black hole thermodynamics* (p. 24-26): four laws of black hole mechanics and phenomenological aspects

6. *Hawking effect* (p. 27-30): periodicity in Euclidean time as temperature; Hawking–Unruh temperature from Euclidean regularity; semi-classical derivation using Bogoliubov-transformation

7. *Action principle* (p. 31-34): canonical decomposition of metric; Gibbons–Hawking–York boundary term; simple mechanics example

8. *Asymptotically AdS boundary conditions* (p. 35-38): Fefferman–Graham expansion, holographic renormalization, boundary stress tensor and asymptotic symmetries
   Black holes in AdS (p. 39-41): free energy from on-shell action and Hawking–Page phase transition
   Gravity aspects of AdS/CFT (p. 41-42): CFT correlation functions, AdS/CFT dictionary and stress-tensor example

Some selected references are spread throughout these lecture notes; a final list of nine references to other lecture notes or review articles can be found on p. 42.

Notes for lecturers and students. If you want to either study the material found in these lecture notes or provide your own lectures based on them please note that I view exercises as an integral part of digesting this material. You can find 10 sets of three exercises (so 30 exercises in total) on my teaching webpage http://quark.itp.tuwien.ac.at/∼grumil/teaching.shtml. A link to exercises and lecture notes for Black Holes I can be found there as well. The whole lecture series Black Holes I+II is intended for a full academic year.

Acknowledgments. Most figures in sections 2 and 3 were prepared by Patrick Binder, Sebastian Schiffer and Thomas Weigner. I thank all students between 2010 and 2018 for their valuable feedback on various aspects of the lectures Black Holes II, in particular Raphaela Wutte.

1 Horizons and other definitions

On this sheet several basic definitions regarding the causal structure of spacetime and black holes are summarized. For a more detailed account see Wald's book "General Relativity" (chapters 8-9 and parts of 11-12) or the book by Hawking & Ellis "The large scale structure of space-time".

1.1 Aspects of causal structure of spacetime

Time-like/null/causal curve. A 1-dimensional curve (which may or may not be a geodesic) in some spacetime is called time-like (null) [causal] if the tangent vector is time-like (null) [time-like or null] along the whole curve.

Chronological/causal future of a point p. The chronological future $I^+(p)$ [causal future $J^+(p)$] is the set of all events in spacetime that can be reached from p by a time-like [causal] curve. This definition generalizes to sets of points.

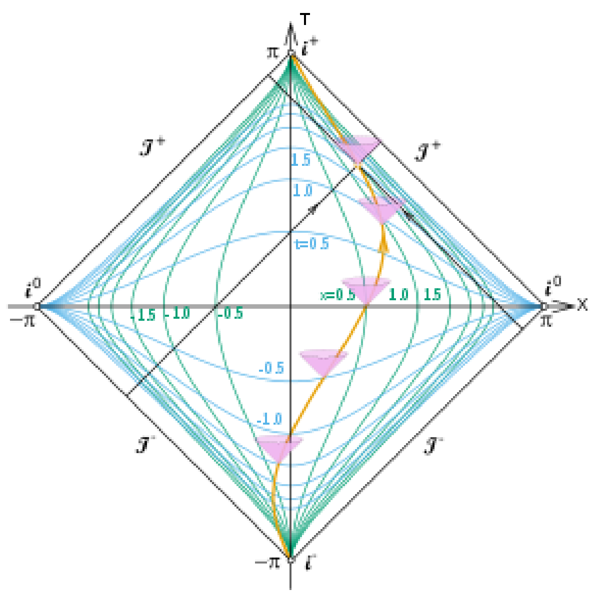
Achronal sets S. A subset S of the spacetime manifold is called achronal if there exists no pair of points $p, q \in S$ such that $q \in I^+(p)$. Equivalently, $I^+(S) \cap S = \{\}$.

Future/past inextendible. A time-like curve is called future (past) inextendible if it has no future (past) endpoint. Analogous definition for causal curves.

Future domain of dependence $D^+(S)$. The future domain of dependence of a closed achronal set S, denoted by $D^+(S)$, is given by the set of all points p in spacetime such that every past inextendible causal curve through p intersects S. Past domain of dependence $D^-(S)$: Exchange "future" $\leftrightarrow$ "past".

Domain of dependence $D(S)$. $D(S) := D^+(S) \cup D^-(S)$.

Asymptotic infinity. Preview of next week on Carter–Penrose diagrams; asymptotic boundaries are denoted by i^+ (future time-like infinity), i^- (past time-like infinity), $\mathscr{I}^+$ (future null infinity), $\mathscr{I}^-$ (past null infinity) and i^0 (spatial infinity.



1.2 Killing, Cauchy, event and apparent horizons

Killing horizon (see last semester). Null hypersurface whose normal is a Killing vector. Useful for stationary black holes, but too restrictive in general.

Cauchy horizon $H(S)$. Let S be a closed achronal set. Its Cauchy horizon $H(S)$ is defined as $H(S) := (\overline{D^+(S)} - I^-[D^+(S)]) \cup (\overline{D^-(S)} - I^+[D^-(S)])$. In plain English, the Cauchy horizon is the boundary of the domain of dependence of S.

Note: A Cauchy horizon is considered as a singularity in the causal structure, since you cannot predict time-evolution beyond a Cauchy horizon.

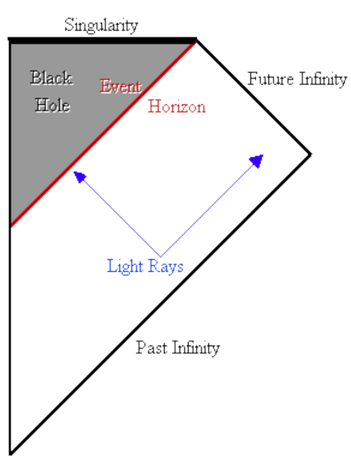
Cauchy surface Σ. A nonempty closed achronal set Σ is a Cauchy surface for some (connected) spacetime manifold iff $H(\Sigma) = \{\}$.

A spacetime M with a Cauchy surface Σ is called "globally hyperbolic".

With the definitions on the first page we are now finally ready to mathematically define the concepts of a black hole region and an event horizon. Note that for astrophysicists these definitions are of limited use since we do not know for sure how our Universe will look like in the infinite future. However, for proving some theorems that apply to isolated black holes it is useful to introduce these definitions.

Black hole region B. In a globally hyperbolic spacetime[1] M the black hole region B is defined by $B := M - J^-(\mathscr{I}^+)$.

Event horizon H. $H := \dot{J}^-(\mathscr{I}^+) \cap M$. In words: the event horizon of a black hole is given by the boundary of the causal past of future null infinity within some spacetime M. See the Carter–Penrose diagram below (again, wait for next week).



Apparent horizon. Wait for later; qualitatively: expansion of null geodesics either negative or zero, i.e., light-rays cannot "escape". Local definition!

[1]Globally hyperbolicity can be too strong, e.g. for charged or rotating black holes, which have a Cauchy horizon as inner horizon. In that case the weaker condition of "strong asymptotic predictability" replaces global hyperbolicity, see the beginning of chapter 12 in Wald's book. Strong asymptotic predictability means that no observer outside a black hole can see a singularity.

2 Carter–Penrose diagrams

In this section Cater–Penrose diagrams (conformal compactifications) are introduced. For a more detailed account in two spacetime dimensions see section 3.2 in hep-th/0204253; see also section 2.4 in gr-qc/9707012.

A simple example of a compactification is the inverse stereographic projection $\mathbb{R}^2 \to S^2$, where infinity is mapped to the North pole on the 2-sphere. Explicitly, for polar coordinates r, ϕ in the plane and standard spherical coordinates θ, φ on the sphere the map reads $r = \cot\frac{\theta}{2}$ and $\phi = \varphi$. Note that $r = \infty$ is mapped to $\theta = 0$. This simple example is a 1-point compactification, meaning that we have to add a single point (spatial infinity) to convert $\mathbb{R}^2$ into something compact, S^2.

2.1 Carter–Penrose diagram of Minkowski space

In Minkowski space we may expect that for a compactification we have to add a whole lightcone, the "lightcone at infinity". We check this now explicitly, applying the coordinate trafo

$$u = \tan\tilde{u} \qquad v = \tan\tilde{v} \qquad \tilde{u}, \tilde{v} \in (-\tfrac{\pi}{2}, \tfrac{\pi}{2}) \tag{1}$$

to the Minkowksi metric in null coordinates ($\mathrm{d}\Omega^2_{S^{D-2}}$ is the metric of S^{D-2})

$$\mathrm{d}s^2 = -\,\mathrm{d}u\,\mathrm{d}v + \tfrac{1}{4}\,(v-u)^2\ \mathrm{d}\Omega^2_{S^{D-2}} \qquad u = t - r \le v = t + r \tag{2}$$

yielding the metric

$$\mathrm{d}s^2 = \Phi^2\,\big(-\,\mathrm{d}\tilde{u}\,\mathrm{d}\tilde{v} + \tfrac{1}{4}\,\sin^2(\tilde{v}-\tilde{u})\ \mathrm{d}\Omega^2_{S^{D-2}}\big) \qquad \Phi^{-1} = \cos\tilde{u}\cos\tilde{v} \tag{3}$$

which is related to a new (unphysical) metric

$$\mathrm{d}\tilde{s}^2 = -\,\mathrm{d}\tilde{u}\,\mathrm{d}\tilde{v} + \tfrac{1}{4}\,\sin^2(\tilde{v}-\tilde{u})\ \mathrm{d}\Omega^2_{S^{D-2}} = \mathrm{d}s^2\,\Phi^{-2} \tag{4}$$

by a Weyl-rescaling

$$\mathrm{d}s^2 = \mathrm{d}\tilde{s}^2\,\Phi^2 \qquad \Leftrightarrow \qquad g_{\mu\nu} = \Phi^2\,\tilde{g}_{\mu\nu}\,. \tag{5}$$

Note that Weyl-rescalings are conformal, i.e., angle-preserving, which in Minkowski signature means **Weyl rescalings preserve the causal structure of spacetime**.

Let us verify this in the Euclidean case, where the angle α between two vectors a^μ and b^μ is given by

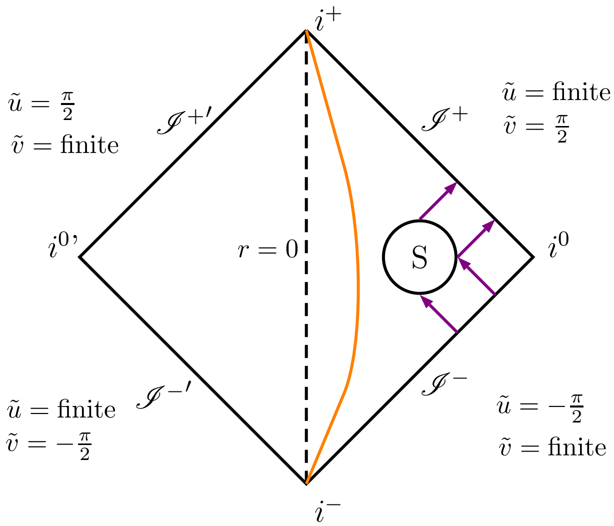
$$\cos\alpha = \frac{g_{\mu\nu}a^\mu b^\nu}{\sqrt{(g_{\mu\nu}a^\mu a^\nu)(g_{\mu\nu}b^\mu b^\nu)}} = \frac{\Phi^2\,\tilde{g}_{\mu\nu}a^\mu b^\nu}{\sqrt{(\Phi^2\,\tilde{g}_{\mu\nu}a^\mu a^\nu)(\Phi^2\,\tilde{g}_{\mu\nu}b^\mu b^\nu)}} = \cos\tilde{\alpha}\,. \tag{6}$$

For Minkowski signature the same calculation applies for vectors that are not null; null vectors are trivially mapped to null vectors under Weyl rescalings (5).

Since Weyl-rescalings preserve the causal structure (but not lengths) we can conveniently compactify spacetimes like Minkowski by adding a lightcone. This means that we consider the conformal Minkowski metric (4) with extended range of coordinates, $\tilde{u}, \tilde{v} \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. The CP-diagram of Minkowski space depicts $\tilde{g}_{\mu\nu}$.

On the backpage the CP-diagram of 2-dimensional Minkowski space is displayed. In such diagrams lines at 45° represent light rays/null geodesics. On any such line either $\tilde{u}$ or $\tilde{v}$ is constant. As an example the diagram shows the scattering of two ingoing into two outgoing lightrays through some interaction (denoted by the S-matrix-symbol S), see the magenta lines. Time-like curves always move within the lightcone, see the orange line.

Higher-dimensional CP-diagrams are similar, but harder to display on paper, since the CP-diagram of any D-dimensional manifold is also D-dimensional. However, often 2-dimensional cuts trough such diagrams convey all relevant info, in particular in the case of spherical symmetry, where each point in the 2-dimensional CP diagram simply corresponds to an S^{D-2}.



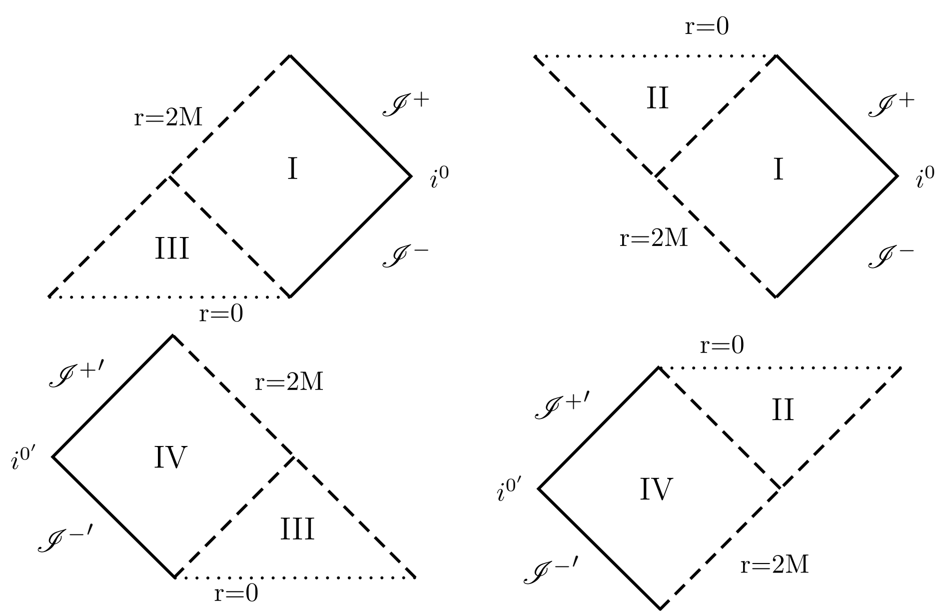
CP diagram of 2d Minkowski (or of higher-dimensional Minkowski if you imagine an S^{D-2} over each point and cut off the diagram at the dashed line corresponding to the origin in spherical coordinates, $r = 0$).
The boundary of the CP-diagram is the light-cone at infinity that was added when compactifying. Its various components correspond to **future (past) timelike infinity** i^+ **(**i^-**), future (past) null infinity** $\mathscr{I}^+$ **(**$\mathscr{I}^-$**) and spatial infinity** i^0.
Note that Minkowski space is globally hyperbolic (exercise: draw some Cauchy hypersurface).

2.2 Carter–Penrose diagram of Schwarzschild

Consider Schwarzschild in outgoing Eddington–Finkelstein (EF) gauge.

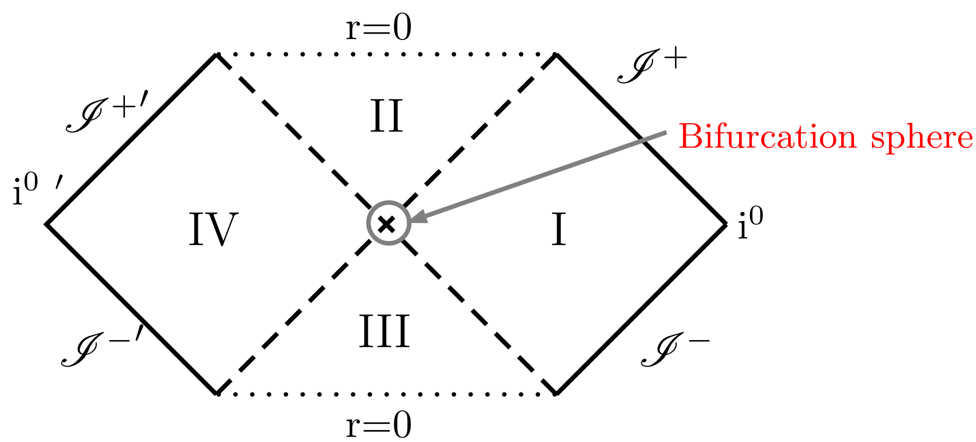
$$\mathrm{d}s^2 = -2\,\mathrm{d}u\,\mathrm{d}r - \big(1-\tfrac{2M}{r}\big)\,\mathrm{d}u^2 + \dots \qquad u = t - r^* \qquad r^* = r + 2M\,\ln\big(\tfrac{r}{2M} - 1\big) \tag{7}$$

EF gauge covers only half of Schwarzschild (ingoing: $-u \to v = t + r^*$). In each EF-patch we have an asymptotic region ($r \to \infty$) that is essentially the same as that of Minkowski space, we have part of the bifurcate Killing horizon and we have the black hole region until we hit the curvature singularity at $r = 0$. Thus, the CP-diagram of an EF-patch is a compactified version of the diagrams we saw last semester, with the compactification working essentially as for Minkowski space.



CP-diagrams for EF-patches. Region I is the external region accessible to the outside observer, region II the black hole region, region III the white hole region and region IV the (unphysical) other external region.

The full Schwarzschild CP-diagram is obtained by gluing together the EF-patches in overlap regions (adding the bifurcation 2-sphere, see Black Holes I).



From the CP-diagram above you can easily apply our definitions of black hole region and event horizon, which you should do as an exercise.

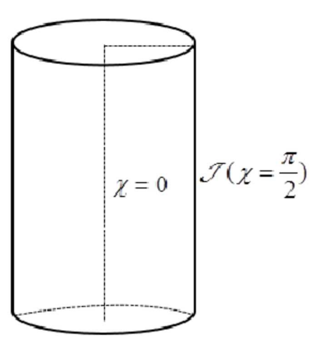
2.3 Carter–Penrose diagram for AdS$_D$

Global Anti-de Sitter (AdS) with AdS-radius ℓ is given by the metric

$$\mathrm{d}s^2 = \ell^2 \big(- \cosh^2\rho \; \mathrm{d}t^2 + \mathrm{d}\rho^2 + \sinh^2\rho \; \mathrm{d}\Omega^2_{S^{D-2}} \big) \qquad \rho \in [0, \infty) \tag{8}$$

which can be rewritten suggestively using a new coordinate $\tan\chi = \sinh\rho$.

$$\mathrm{d}s^2 = \frac{\ell^2}{\cos^2\chi} \big(- \mathrm{d}t^2 + \mathrm{d}\chi^2 + \sin^2\chi \; \mathrm{d}\Omega^2_{S^{D-2}} \big) = \mathrm{d}\tilde{s}^2 \, \Phi^2 \qquad \chi \in [0, \tfrac{\pi}{2}) \tag{9}$$



The compactified metric $\tilde{g}$ differs from the physical metric g by a conformal factor $\Phi^2 = \ell^2/\cos^2\chi$ and allows to add the asymptotic boundary $\chi = \frac{\pi}{2}$. At $\chi = \frac{\pi}{2}$ the compactified metric

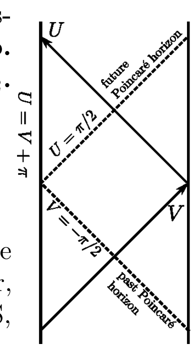
$$\mathrm{d}\tilde{s}^2|_{\chi=\frac{\pi}{2}} = - \mathrm{d}t^2 + \mathrm{d}\Omega^2_{S^{D-2}}$$

describes a $(D-1)$-dimensional cylinder. Thus, **the CP-diagram of AdS$_D$ is a filled cylinder**. **The figure shows the CP-diagram of AdS$_3$.** In higher dimensions the "celestial circle" is replaced by a "celestial sphere" of dimension $D-2$.

Two dimensions are special, since the 0-sphere consists of two disjoint points. **The CP-diagram of AdS$_2$ is a 2d vertical strip. The CP diagram of dS$_2$ is rotated by 90° relative to AdS$_2$.** If instead of global AdS$_D$ we consider Poincaré-patch AdS$_D$,

$$\mathrm{d}s^2 = \frac{\ell^2}{z^2} \big(- \mathrm{d}t^2 + \mathrm{d}z^2 + \mathrm{d}x_1^2 + \dots + \mathrm{d}x^2_{D-2} \big)$$

the metric is manifestly conformally flat so that we get the same CP-diagram as for Minkowski space, namely a triangle. However, that triangle only covers part of the full CP-diagram of global AdS, which for AdS$_2$ is depicted to the right.



2.4 Carter–Penrose diagrams in two spacetime dimensions

Gravity in 2d is described by dilaton gravity theories, see hep-th/0204253 for a review. For all such theories there is a generalized Birkhoff theorem so that all solutions have a Killing vector and the metric in a basic EF-patch reads

$$\mathrm{d}s^2 = -2\,\mathrm{d}u\,\mathrm{d}r - K(r)\;\mathrm{d}u^2 \tag{10}$$

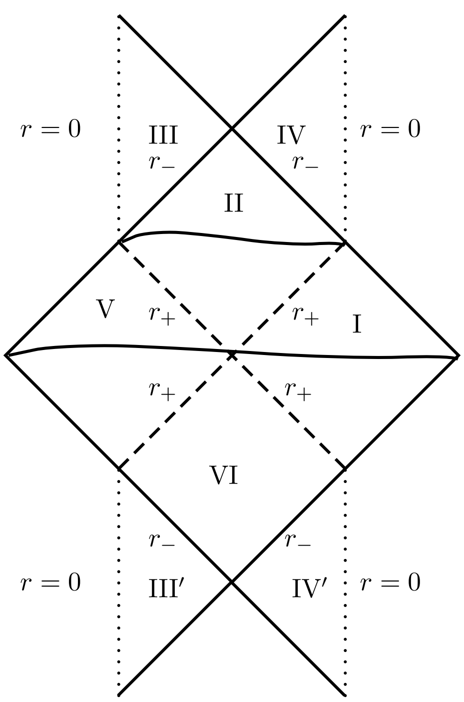
with some arbitrary function $K(r)$ that depends on the specific theory. Non-extremal Killing horizons arise whenever $K(r)$ has a single zero (in case of double or higher zeros the Killing horizons are extremal).

While there is a straightforward detailed algorithm to construct all CP-diagrams in 2d dilaton gravity, in most cases the following simpler recipe works:

1. Identify the asymptotic region (Minkowski, AdS, dS, else) by checking the behavior of $K(r)$ at large radii, $r \to \infty$
2. Identify the number and types of Killing horizons by finding all zeros (as well as their multiplicities) of $K(r)$
3. Identify curvature singularities by calculating $K''(r)$ and checking whether it remains finite; check if singularities reachable with geodesics of finite length
4. Use the info above to "guess" the CP-diagram of a basic EF-patch
5. Copy three mirror images of the CP-diagram of the basic EF-patch
6. Glue together all EF-patches on overlap regions to get full CP-diagram
7. If applicable continue full CP-diagram periodically

As an example we consider Reissner–Nordström, whose 2d part is (10) with

$$K(r) = 1 - \frac{2M}{r} + \frac{Q^2}{r^2} \qquad\qquad r_\pm = M \pm \sqrt{M^2 - Q^2}\,, \quad M > |Q|\,. \tag{11}$$



CP diagram of Reissner–Nordström. Applying the recipe yields 1. asymptotic flatness for $r \to \infty$, 2. two non-extremal Killing horizons for $M > |Q|$ at $r = r_\pm$, 3. a curvature singularity at $r = 0$, 4. a basic EF-patch similar to Schwarzschild, but with an additional Killing horizon, 5. corresponding mirror flips, and 6. the CP-diagram displayed on the left. Concerning 7., one could identify region III with III' and IV with IV' or declare them to be different and get several copies of the CP-diagram appended above and below. Note, however, that the **inner horizon** $r = r_-$ **is a Cauchy horizon**. Indeed, the domain of dependence of the achronal set reaching from i^0 in region I to $i^{0\,\prime}$ in region V is given by the union of regions I, II, V and VI, but excludes regions III and IV beyond the Cauchy horizon. **Cauchy horizons are believed to be unstable.** If true, then regions III and IV are merely artifacts.

Note: can finally check incompleteness of geodesics at singularity and completeness at asymptotic boundary, e.g. null geodesics $\mathrm{d}u/\,\mathrm{d}r = -2/K(r)$.

3 Raychaudhuri equation and singularity theorems

In cosmology and theoretical GR we are often interested in the movement of nearby bits of matter (primordial fluctuations during inflation, stars in a galaxy, galaxies in a cluster, test-particles in some black hole background etc.). Besides practical applications, these considerations are of importance for singularity theorems, as we shall see. The equations that describe the acceleration of nearby test-particles are known as "Raychaudhuri equations", and our first task is to derive them.

3.1 Geodesic congruences

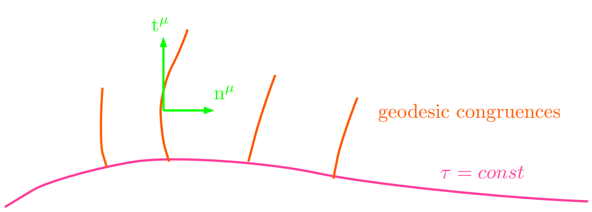
A congruence is a set of curves such that exactly one curve goes through each point in the manifold. A geodesic congruence is a congruence where all curves are geodesics. For concreteness we assume that all geodesics on our congruence are time-like. Consider a single geodesic with tangent vector t^μ, normalized such that $t^2 = -1$. We define a velocity tensor B as the covariant derivative of the tangent vector.

$$B_{\mu\nu} := \nabla_\nu t_\mu \tag{12}$$

Since in Riemannian geometry geodesics are also autoparallels, we can use the autoparallel equation $t^\mu \nabla_\mu t^\nu = 0$ to deduce that the velocity tensor projects to zero when contracted with the tangent vector.

$$B_{\mu\nu} t^\nu = 0 = B_{\nu\mu} t^\nu \tag{13}$$

Consider a timelike geodesic congruence and introduce the normal vector field n^μ, describing infinitesimal displacement between nearby geodesics.



Orange lines denote members of a timelike geodesic congruence. The pink line is some Cauchy surface at some constant value of time τ. Green arrows denote one example of the tangent vector t^μ and the normal vector n^μ.

The normal vector by definition commutes with the tangent vector, so that the Lie-derivative of one such vector with respect to the other vanishes, e.g. $\mathcal{L}_t n^\mu = t^\nu \nabla_\nu n^\mu - n^\nu \nabla_\nu t^\mu = 0$. Using this property yields a chain of equalities:

$$t^\nu \nabla_\nu n^\mu = n^\nu \nabla_\nu t^\mu = n^\nu B^\mu{}_\nu \tag{14}$$

The equalities (14) let us interpret the tensor $B^\mu{}_\nu$ as measuring the failure of the normal vector n^μ to be transported parallel along the tangent vector t^μ. Thus, an observer following some geodesic would deduce that nearby geodesics are stretched and rotated by the linear map $B^\mu{}_\nu$.

It is useful to decompose the tensor $B_{\mu\nu}$ into its algebraically irreducible components. To this end we define a projector [see also section 11.1 in Black Holes I lecture notes, just before Eq. (11.8); D is the spacetime dimension]

$$\Pi_{\mu\nu} := g_{\mu\nu} + t_\mu t_\nu = \Pi_{\nu\mu} \qquad \Pi_{\mu\nu} t^\nu = 0 \qquad \Pi^\mu{}_\nu \Pi^\nu{}_\lambda = \Pi^\mu{}_\lambda \qquad \Pi^\mu{}_\mu = D-1 \tag{15}$$

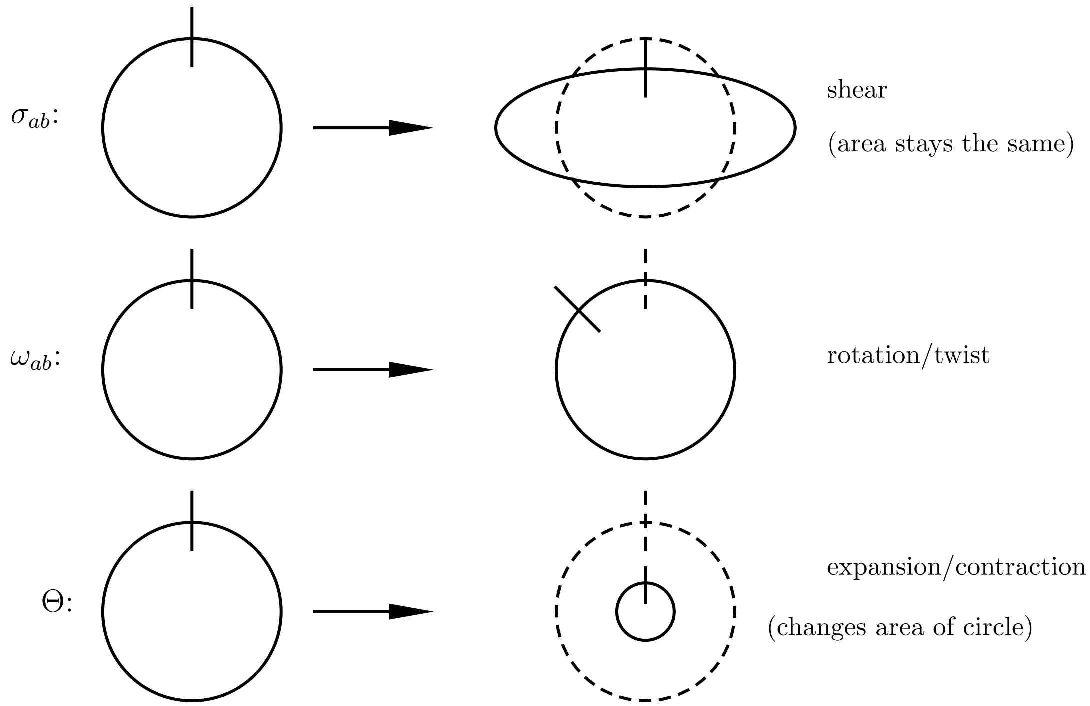
and split B into symmetric traceless part (shear σ), antisymmetric part (twist ω) and trace part (expansion Θ)

$$\sigma_{\mu\nu} := B_{(\mu\nu)} - \tfrac{1}{D-1}\,\Theta\,\Pi_{\mu\nu} \qquad \omega_{\mu\nu} := B_{[\mu\nu]} \qquad \Theta := B^\mu{}_\mu \tag{16}$$

so that

$$B_{\mu\nu} = \sigma_{\mu\nu} + \omega_{\mu\nu} + \tfrac{1}{D-1}\,\Theta\,\Pi_{\mu\nu}\,. \tag{17}$$

Below is a simple picture of the various contributions to the deformation tensor $B_{\mu\nu}$, starting with a circular ring of geodesics and some reference observer denoted by a line.



3.2 Raychaudhuri equation

We are interested in acceleration, so we consider the derivative of the deformation tensor B along the tangent vector t and manipulate suitably.

$$\begin{aligned} t^\lambda \nabla_\lambda B_{\mu\nu} &= t^\lambda \nabla_\lambda \nabla_\nu t_\mu = t^\lambda \nabla_\nu \nabla_\lambda t_\mu + t^\lambda \left[\nabla_\lambda,\, \nabla_\nu\right] t_\mu \\ &= \nabla_\nu \big(t^\lambda \nabla_\lambda t_\mu\big) - (\nabla_\nu t^\lambda)(\nabla_\lambda t_\mu) - t^\lambda R^\alpha{}_{\mu\lambda\nu} t_\alpha = -B^\lambda{}_\nu B_{\mu\lambda} - R^\alpha{}_{\mu\lambda\nu} t^\lambda t_\alpha \end{aligned} \tag{18}$$

The equation above describes the acceleration of all deformation types.

Often one is interested particularly in the acceleration associated with expansion, which is obtained by taking the trace of (18).

$$t^\lambda \nabla_\lambda B^\mu{}_\mu = -B^{\mu\nu} B_{\nu\mu} - R_{\mu\nu} t^\mu t^\nu \tag{19}$$

Defining $\mathrm{d}/\,\mathrm{d}\tau := t^\mu \nabla_\mu$ and expanding the quadratic term in B in terms of its irreducible components (17) yields the **Raychaudhuri equation:**

$$\boxed{\frac{\mathrm{d}\Theta}{\mathrm{d}\tau} = -\frac{1}{D-1}\,\Theta^2 - \sigma_{\mu\nu}\sigma^{\mu\nu} + \omega_{\mu\nu}\omega^{\mu\nu} - R_{\mu\nu} t^\mu t^\nu} \tag{20}$$

A key aspect of the right hand side of the Raychaudhuri equation (20) is that the first and second term are non-positive. The third term vanishes in many situations (twist-free congruences), while the last term is non-positive if the Einstein equations are fulfilled and the strong energy condition holds for all unit timelike vectors t,

$$T_{\mu\nu} t^\mu t^\nu \geq -\tfrac{1}{2}\, T \qquad \Rightarrow \qquad R_{\mu\nu} t^\mu t^\nu = \kappa\big(T_{\mu\nu} - \tfrac{1}{2}\, g_{\mu\nu} T\big)\, t^\mu t^\nu \geq 0\,. \tag{21}$$

Caveat: all local energy conditions are violated by quantum effects; while most of them are expected to hold for "reasonable" classical matter, the strong energy condition (21) is already violated by a cosmological constant. So take classical energy conditions with a grain of salt.

3.3 Glimpse of singularity theorems

There is a number of singularity theorems that can be proven through the same type of scheme: assume some convexity condition (like some energy condition) and some trapping condition (like negativity of expansion). Then use something like the Raychaudhuri equation to deduce the existence of a singularity. The conclusion is that, given certain conditions, the existence of a black hole predicts the existence of a singularity. Thus, classically singularities are an unavoidable feature of spacetimes that contain black holes.

It is not the intention of these lecture to prove such theorems in generality, but we shall at least prove a simpler theorem that allows to deduce the singularity in a timelike geodesic congruence (which is not necessarily a singularity in spacetime).

Theorem. Let t^μ be the tangent vector field in a timelike geodesic congruence that is twist-free and assume $R_{\mu\nu}t^\mu t^\nu \geq 0$. If the expansion Θ associated with this congruence takes the negative value Θ_0 at any point of a geodesic, then the expansion diverges to $-\infty$ along that geodesic within a proper time $\tau \leq (D-1)/|\Theta_0|$.

Proof. The Raychaudhuri equation (20) together with absence of twist, $\omega_{\mu\nu} = 0$, and the convexity property $R_{\mu\nu}t^\mu t^\nu \geq 0$ establishes the differential inequality

$$\frac{\mathrm{d}\Theta}{\mathrm{d}\tau} \leq -\frac{1}{D-1}\,\Theta^2 \qquad \Rightarrow \qquad \frac{\mathrm{d}\Theta^{-1}}{\mathrm{d}\tau} \geq \frac{1}{D-1} \tag{22}$$

which is easily solved.

$$\Theta^{-1}(\tau) \geq \Theta^{-1}(\tau_0) + \frac{\tau-\tau_0}{D-1} \tag{23}$$

Assuming that the initial value at $\tau_0 = 0$ is such that $\Theta(0) = \Theta_0 < 0$ (by assumptions of the theorem such a τ_0 must exist and with no loss of generality we shift it to $\tau_0 = 0$) the right hand side of (23) has a zero at some finite $\tau \leq (D-1)/|\Theta_0|$. This means that $1/\Theta$ goes to zero from below, so that Θ tends to $-\infty$. □

More generally, Hawking, Penrose and others have proved that given some convexity property (e.g. ensured by some energy condition and the fulfillment of the Einstein equations) together with the existence of some trapped surface implies the existence of at least one incomplete geodesic (usually also some condition on the causal structure is required, like the absence of closed timelike curves). By definition this means that there is a singularity. The lesson is, whenever you have a black hole you have a singularity. Thus, the singularities inside Schwarzschild or Kerr are not an artifact of a highly symmetric situation but a generic feature of black holes.

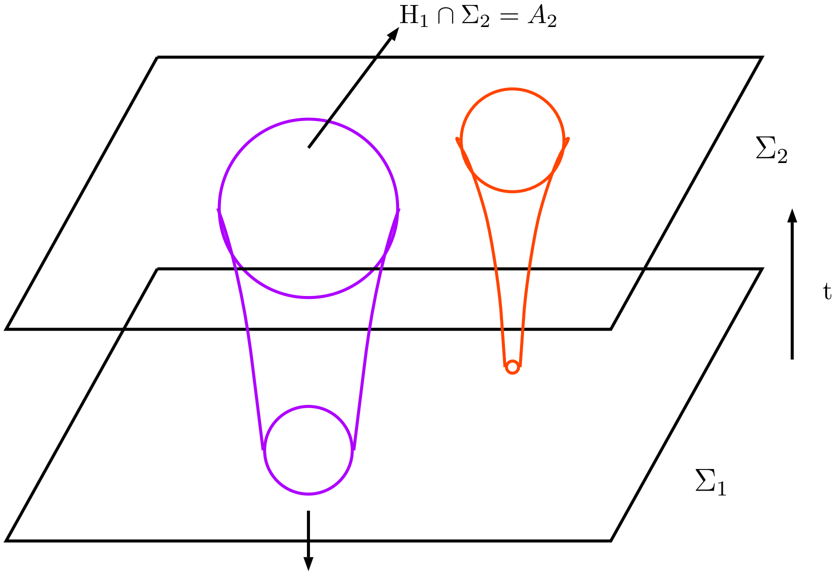
3.4 Remarks on other theorems, especially the area theorem

There is a number of useful theorems, for instance Penrose's theorem that future event horizons have no future end points or the Schoen–Yau/Witten theorem of positivity of energy. If you are interested in them you are strongly encouraged to consult the Hawking & Ellis book or reviews (e.g. 1302.3405 or physics/0605007).

Perhaps the most remarkable one is **Hawking's area theorem**. We are not going to prove it, but here are at least the assumptions, a version of the theorem itself, an idea of how to prove it and some interpretation what it means.

Assume that the Einstein equations hold and that $T_{\mu\nu}$ obeys some energy condition (e.g. the "weak energy condition", $T_{\mu\nu}t^\mu t^\nu \geq 0$ for all timelike vectors t). Assume further cosmic censorship (which is satisfied, for instance, if spacetime is globally hyperbolic, i.e., there is a Cauchy surface). Finally, assume there is an event horizon and that spacetime is asymptotically flat. Then **the area of the event horizon is monotonically increasing as a function of time.**

Implication of Hawking's area theorem: black holes grow but do not shrink!



Idea of proof. It is sufficient to show that each area element a is monotonically increasing in time. Using the expansion Θ it is easy to show

$$\frac{\mathrm{d}a}{\mathrm{d}\tau} = \Theta\, a\,. \tag{24}$$

Thus, Hawking's area law holds if $\Theta \geq 0$ everywhere on the event horizon. The second part of the proof is to show that whenever $\Theta < 0$ there must be a singularity, so that either one of the assumptions of the theorem fails to hold or we get a contradiction to Penrose's theorem that the event horizon has no future endpoint. Either way, the conclusion is that $\Theta < 0$ cannot hold on the event horizon, which proves Hawking's area theorem.

Hawking's area theorem can be expressed as a formula e.g. as follows. Let H by the event horizon and $\Sigma_{1,2}$ two Cauchy surfaces at times $\tau_{1,2}$ with $\tau_2 > \tau_1$. Then Hawking's area theorem states

$$H \cap \Sigma_2 \geq H \cap \Sigma_1\,. \tag{25}$$

Yet another way to express the same content (in a very suggestive way) is to simply call the area "A" and to write Hawking's area theorem as a convexity condition reminiscent of the second law of thermodynamics,

$$\boxed{\delta A \geq 0\,.} \tag{26}$$

The inequality (26) is also known as "second law of black hole mechanics". We shall see later that the similarity to the second law of thermodynamics is not just incidental. Note that we have encountered already the zeroth law (constancy of surface gravity for stationary black holes) in Black Holes I, and we shall learn about the first law a bit later. Also various versions of the third law can be proven for black holes (which means the impossibility to reach an extremal black hole starting with a non-extremal one within finite time).

4 Linearized gravity

In many instances (not just in gravity but also in quantum field theory) one is interested in linearizing perturbations around a fixed background, which considerably simplifies the classical and quantum analysis. While this approach is only justified if the linearized perturbation is small enough, there are numerous applications where this assumption holds. Examples include gravitational waves, holographic applications and perturbative quantization of gravity. In this section we develop the basic tools to address all these issues.

4.1 Linearization of geometry around fixed background

Assume that the metric can be meaningfully split into background $\bar{g}_{\mu\nu}$ and fluctuations $h_{\mu\nu}$. You can think of $\bar{g}$ as some classical background (e.g. Minkowski space, AdS, dS, FLRW or some black hole background) and of h either as a classical perturbation (e.g. a gravitational wave on your background) or as a variation of the metric (e.g. when checking the variational principle or in holographic contexts) or as a quantum fluctuation (e.g. when semi-classically quantizing gravity).

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu} \tag{1}$$

For calculations we generally need various geometric quantities, like the inverse metric, the Christoffel symbols, the Riemann tensor etc., so we consider them now to linear order in h. Note that $h_{\mu\nu} = h_{\nu\mu}$ is a symmetric tensor.

Let us start with the inverse metric. The identity $g^{\mu\nu} g_{\nu\lambda} = \delta^\mu_\lambda$ yields

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

In all linearized expressions we raise and lower indices with the background metric $\bar{g}$, so that e.g. $h^{\mu\nu} = \bar{g}^{\mu\alpha}\bar{g}^{\nu\beta}h_{\alpha\beta}$. All quantities with bar on top have their usual meaning and are constructed from the background metric $\bar{g}$, e.g. $\bar{\Gamma}^\alpha{}_{\beta\gamma} = \frac{1}{2}\,\bar{g}^{\alpha\mu}\left(\bar{g}_{\beta\mu,\gamma} + \bar{g}_{\gamma\mu,\beta} - \bar{g}_{\beta\gamma,\mu}\right)$. We denote the difference between full and background expression with δ, for example $\delta g_{\mu\nu} = g_{\mu\nu} - \bar{g}_{\mu\nu} = h_{\mu\nu}$ and $\delta g^{\mu\nu} = g^{\mu\nu} - \bar{g}^{\mu\nu} = -h^{\mu\nu}$.

The determinant of the metric expands as explained in Black Holes I. (We suppress from now on $\mathcal{O}(h^2)$ as it is understood that all equations below hold only at linearized level.)

$$\sqrt{-g} = \sqrt{-\bar{g}}\,\Big(1 + \frac{1}{2}\,\bar{g}^{\mu\nu}h_{\mu\nu}\Big) \tag{3}$$

The Christoffel symbols expand as follows

$$\delta\Gamma^\alpha{}_{\beta\gamma} = \Gamma^\alpha{}_{\beta\gamma} - \bar{\Gamma}^\alpha{}_{\beta\gamma} = \frac{1}{2}\,\bar{g}^{\alpha\mu}\left(\bar{\nabla}_\beta h_{\gamma\mu} + \bar{\nabla}_\gamma h_{\beta\mu} - \bar{\nabla}_\mu h_{\beta\gamma}\right). \tag{4}$$

The result (4) implies that the **variation of the Christoffels, $\delta\Gamma$, is a tensor**.

The linearized Riemann tensor can be expressed concisely in terms of (4).

$$\boxed{\delta R^\alpha{}_{\beta\mu\nu} = \bar{\nabla}_\mu\,\delta\Gamma^\alpha{}_{\beta\nu} - \bar{\nabla}_\nu\,\delta\Gamma^\alpha{}_{\beta\mu}} \tag{5}$$

While the results above are all we need for now, it is useful to provide more explicit results for the linearized Ricci-tensor

$$\delta R_{\mu\nu} = \bar{\nabla}_\alpha\,\delta\Gamma^\alpha{}_{\mu\nu} - \bar{\nabla}_\nu\,\delta\Gamma^\alpha{}_{\mu\alpha} = \frac{1}{2}\left(\bar{\nabla}^\alpha\bar{\nabla}_\mu h_{\alpha\nu} + \bar{\nabla}^\alpha\bar{\nabla}_\nu h_{\alpha\mu} - \bar{\nabla}_\mu\bar{\nabla}_\nu h^\alpha_\alpha - \bar{\nabla}^2 h_{\mu\nu}\right) \tag{6}$$

and the linearized Ricci-scalar

$$\delta R = -\bar{R}^{\mu\nu}h_{\mu\nu} + \bar{\nabla}^\mu\bar{\nabla}^\nu h_{\mu\nu} - \bar{\nabla}^2 h^\mu_\mu\,. \tag{7}$$

4.2 Linearization of Einstein equations

Consider the vacuum Einstein equations $R_{\mu\nu} = 0$ and assume some solution thereof for the background metric, $\bar{g}_{\mu\nu}$ such that $\bar{R}_{\mu\nu} = 0$ (e.g. $\bar{g}$ could be Minkowski space or the Kerr solution). Classical perturbations around that background then have to obey the linearized Einstein equations $\delta R_{\mu\nu} = 0$, viz.

$$\bar{\nabla}^\alpha \bar{\nabla}_\mu h_{\alpha\nu} + \bar{\nabla}^\alpha \bar{\nabla}_\nu h_{\alpha\mu} - \bar{\nabla}_\mu \bar{\nabla}_\nu h^\alpha_\alpha - \bar{\nabla}^2 h_{\mu\nu} = 0\,. \tag{8}$$

Before attempting to solve these equations it is useful to decompose the perturbations h as follows.

$$h_{\mu\nu} = h^{TT}_{\mu\nu} + \bar{\nabla}_{(\mu}\xi_{\nu)} + \frac{1}{D}\,\bar{g}_{\mu\nu}\,h \tag{9}$$

The first contribution on the right hand side of (9) is called "transverse-traceless part" (TT-part) since it obeys the conditions

$$\bar{\nabla}^\mu h^{TT}_{\mu\nu} = 0 = h_\mu^{\mu\,TT}\,. \tag{10}$$

The second contribution on the right hand side of (9) is called "gauge part" since it can be compensated by an infinitesimal diffeomorphism of the background metric, $\mathcal{L}_\xi \bar{g}_{\mu\nu} = \bar{\nabla}_\mu \xi_\nu + \bar{\nabla}_\nu \xi_\mu$. The last contribution on the right hand side of (9) is called "trace part", since up to a gauge term the trace of $h_{\mu\nu}$ is given by h. Alternatively, one can call the three contributions (in this order) tensor, vector and scalar part.

In $D \geq 3$ spacetime dimensions the tensor $h_{\mu\nu}$ has $D(D+1)/2$ algebraically independent components, with D of them residing in the gauge part and 1 of them in the trace part. This means at this stage the TT-part has $(D+1)(D-2)/2$ algebraically independent components, which corresponds to the correct number of massive spin-2 polarizations. However, in Einstein gravity gravitons are massless which reduces the number of polarizations. As we shall see below **there are $D(D-3)/2$ gravity wave polarizations in D-dimensional Einstein gravity.**

For simplicity we assume from now on that the background metric is flat so that $\bar{R}^\alpha{}_{\beta\gamma\delta} = 0$. We evaluate for this case the linearized Einstein equations (8) separately for the TT-part[1]

$$\text{on flat background:} \quad \bar{\nabla}^2 h^{TT}_{\mu\nu} = 0 \tag{11}$$

and the trace part $(\bar{g}_{\mu\nu}\bar{\nabla}^2 + (D-2)\bar{\nabla}_\mu\bar{\nabla}_\nu)h = 0$. The gauge part trivially solves the linearized Einstein equations (8).

Thus, on a flat background the TT-part obeys a wave equation (11), essentially of the same type as a vacuum Maxwell-field in Lorenz-gauge. We show now that the same wave equation can be obtained by suitable gauge fixing of the original $h_{\mu\nu}$, namely by imposing harmonic gauge, a.k.a. de-Donder gauge

$$\bar{\nabla}_\mu h^\mu_\nu = \frac{1}{2}\,\partial_\nu h^\mu_\mu\,. \tag{12}$$

The gauge choice (12) fixes D of the $D(D+1)/2$ components of $h_{\mu\nu}$, but we still have residual gauge freedom, i.e., gauge transformations

$$h_{\mu\nu} \to \tilde{h}_{\mu\nu} = h_{\mu\nu} + \bar{\nabla}_\mu\xi_\nu + \bar{\nabla}_\nu\xi_\mu \quad \text{such that} \quad \bar{\nabla}_\mu \tilde{h}^\mu_\nu = \frac{1}{2}\,\partial_\nu \tilde{h}^\mu_\mu \tag{13}$$

that preserve de-Donder gauge. The last equality in (13) establishes $\bar{\nabla}^2\xi_\mu = 0$, so that we have D independent residual gauge transformations. In conclusion, the number of physical degrees of freedom contained in linearized perturbations $h_{\mu\nu}$ in Einstein gravity is given by $D(D+1)/2 - 2D = D(D-3)/2$. Inserting de-Donder gauge (12) into the linearized Einstein equations (8) yields $\bar{\nabla}^2 h_{\mu\nu} = 0$, as promised.

[1] The reason why this makes sense is because TT-, gauge- and trace-part decouple in the quadratic action (16) below. Hence, also the linearized field equations decouple.

4.3 Linearization of Hilbert action

We can use the linearization not only at the level of field equations but also at the level of the action.

As a first task we fill in a gap that was left open in Black Holes I when deriving the Einstein equations from varying the Hilbert action. We drop here all bars on top of the metric and denote the fluctuation by δg instead of h. Using the formulas for the variation of the determinant (3) and the Ricci scalar (7) yields

$$\delta I_{\text{EH}} = \frac{1}{16\pi G}\,\delta\int \mathrm{d}^D x\sqrt{-g}\,R = \frac{1}{16\pi G}\int \mathrm{d}^D x\sqrt{-g}\,\Big(\big(\tfrac{1}{2}\,g^{\mu\nu}\,R - R^{\mu\nu}\big)\,\delta g_{\mu\nu} + \nabla^\mu\big(\nabla^\nu\delta g_{\mu\nu} - g^{\alpha\beta}\nabla_\mu\delta g_{\alpha\beta}\big)\Big)\,. \tag{14}$$

Setting to zero the terms in the first line for arbitrary variations yields the vacuum Einstein equations. The terms in the second line are total derivative terms and vanish upon introducing a suitable boundary action and suitable boundary conditions on the metric (we shall learn more about this later in these lectures).

As second task we vary the action (14) again to obtain an expression quadratic in the fluctuations (again dropping total derivative terms). Since $\delta^2 g_{\mu\nu} = 0$ and the Einstein equations hold for the background we only need to vary the Einstein tensor.

$$\delta G_{\mu\nu} = \delta R_{\mu\nu} - \tfrac{1}{2}\,\bar{R}\delta g_{\mu\nu} - \tfrac{1}{2}\,\bar{g}_{\mu\nu}\,\delta R = \delta R_{\mu\nu} - \tfrac{1}{2}\,\bar{g}_{\mu\nu}\,\delta R \tag{15}$$

Plugging (15) together with (6) and (7) into the second variation of the action (using again h instead of δg) establishes the quadratic action (up to boundary terms)

$$16\pi G\, I^{(2)}_{\text{EH}} = \int \mathrm{d}^D x\sqrt{-\bar{g}}\,h^{\mu\nu}\,\delta G_{\mu\nu} = \int \mathrm{d}^D x\sqrt{-\bar{g}}\,\tfrac{1}{2}\,h^{\mu\nu}\,\big(\bar{\Box}_{\mu\nu}{}^{\alpha\beta}\,h_{\alpha\beta}\big) \tag{16}$$

with the wave operator

$$\bar{\Box}_{\mu\nu}{}^{\alpha\beta} = \delta^\beta_\nu\bar{\nabla}^\alpha\bar{\nabla}_\mu + \delta^\beta_\mu\bar{\nabla}^\alpha\bar{\nabla}_\nu - \bar{g}^{\alpha\beta}\,\bar{\nabla}_\mu\bar{\nabla}_\nu - \delta^\alpha_\mu\delta^\beta_\nu\bar{\nabla}^2 - \bar{g}_{\mu\nu}\,\bar{\nabla}^\alpha\bar{\nabla}^\beta + \bar{g}_{\mu\nu}\bar{g}^{\alpha\beta}\,\bar{\nabla}^2\,. \tag{17}$$

The quadratic action (16) has a number of uses for semi-classical gravity and holography. The field equations for h associated with the action (16), $\bar{\Box}_{\mu\nu}{}^{\alpha\beta}\,h_{\alpha\beta} = 0$, are equivalent to the linearized Einstein equations (8). Thus, the action (16) is a perturbative action for the gravitational wave (plus gauge) degrees of freedom.

4.4 Backreactions and recovering Einstein gravity

In this subsection we work schematically, omitting factors and indices. In the presence of matter sources T the quadratic action reads $I^{(2)} \sim \int(\frac{1}{G}\,h\partial^2 h + hT)$. However, in contrast to electrodynamics where the photon is not charged, the graviton is charged under its own gauge group, i.e., gravitons have energy and thus interact with themselves. One can take this effect into account perturbatively by calculating the energy-momentum tensor associated with the quadratic fluctuations, which schematically is of the form $T^{(2)} \sim \frac{1}{G}\,\partial h\partial h$. Thus, taking into account backreactions we are led to a cubic action $I^{(3)} \sim \int(\frac{1}{G}h\partial^2 h + \frac{1}{G}\,h\partial h\partial h + hT)$. However, the cubic term also contributes to the stress tensor, $T^{(3)} \sim \frac{1}{G}\,h\partial h\partial h$ and so forth. Continuing this perturbative expansion yields an action

$$I^{(\infty)} \sim \int\big[\tfrac{1}{G}\,\big(h\partial^2 h + h\partial h\partial h + h^2\partial h\partial h + h^3\partial h\partial h + \dots\big) + hT\big]\,. \tag{18}$$

It was shown by Boulware and Deser that the whole sum can be rewritten as $\frac{1}{G}\,\sqrt{-g}\,R$, so that even if one had never heard of Riemannian geometry in principle one could derive the Hilbert action of Einstein gravity by starting with a massless spin-2 action (16), adding a source and taking into account consistently backreactions.

5 Gravitational waves

5.1 Gravitational waves in vacuum

Let us stick to $D = 4$ and solve the gravitational wave equation on a Minkowski background together with de-Donder gauge,

$$\partial^2 h_{\mu\nu} = 0 = \partial_\mu h^\mu_\nu - \tfrac{1}{2}\,\partial_\nu h^\mu_\mu\,. \tag{19}$$

Linearity of the wave equation allows us to use the superposition principle and build the general solution in terms of plane waves

$$h_{\mu\nu} = \epsilon_{\mu\nu}(k)\, e^{ik_\mu x^\mu} \qquad\qquad k^2 = 0 \qquad\qquad k_\mu \epsilon^\mu_\nu = \tfrac{1}{2}\, k_\nu \epsilon^\mu_\mu\,. \tag{20}$$

The first equality contains the symmetric polarization tensor $\epsilon_{\mu\nu}$ that has to obey the third equality to be compatible with de-Donder gauge. The second equality ensures that the wave equation holds. The general solution is then some arbitrary superposition of plane waves (20), exactly as for photons in electrodynamics.

The four residual gauge transformations are now used to set to zero the components $\epsilon_{0i} = 0$ and the trace $\epsilon^\mu_\mu = 0$. Thus, the de-Donder condition (20) simplifies to transversality, $k^\mu \epsilon_{\mu\nu} = 0$. [With these choices the polarization tensor is transverse and traceless, so that only h^{TT} in (9) contributes.] For concreteness assume now that the gravitational wave propagates in z-direction, $k^\mu = \omega\,(1,0,0,1)^\mu$. Then transversality implies $\epsilon_{00} = \epsilon_{0x} = \epsilon_{0y} = \epsilon_{0z} = \epsilon_{xz} = \epsilon_{yz} = \epsilon_{zz} = 0$. Together with symmetry, $\epsilon_{\mu\nu} = \epsilon_{\nu\mu}$, and traceleceness, $\epsilon^\mu_\mu = 0$, the polarization tensor

$$\epsilon_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \epsilon_+ & \epsilon_\times & 0 \\ 0 & \epsilon_\times & -\epsilon_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_{\mu\nu} =: \epsilon^+_{\mu\nu} + \epsilon^\times_{\mu\nu} \tag{21}$$

is characterized by two real numbers, corresponding to the two polarizations of gravitational waves or, equivalently, to the two helicity states of massless spin-2 particles. They are called "plus-polarization" (ϵ_+) and "cross-polarization" ($\epsilon_\times$).

5.2 Gravitational waves acting on test particles

With a single test-particle it is impossible to detect a gravitational wave, so let us assume there are two massive test-particles, one at the origin (A) and the other (B) at some finite distance L_0 along the x-axis. Let us further assume there is a planar gravitational wave propagating along the z-direction with $+$-polarization, $h_{\mu\nu} = \epsilon^+_{\mu\nu} f(t-z)$ with $\epsilon_+ = 1$. The perturbed metric then reads

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \big(1 + f(t-z)\big)\,\mathrm{d}x^2 + \big(1 - f(t-z)\big)\,\mathrm{d}y^2 + \mathrm{d}z^2 \qquad f \ll 1\,. \tag{22}$$

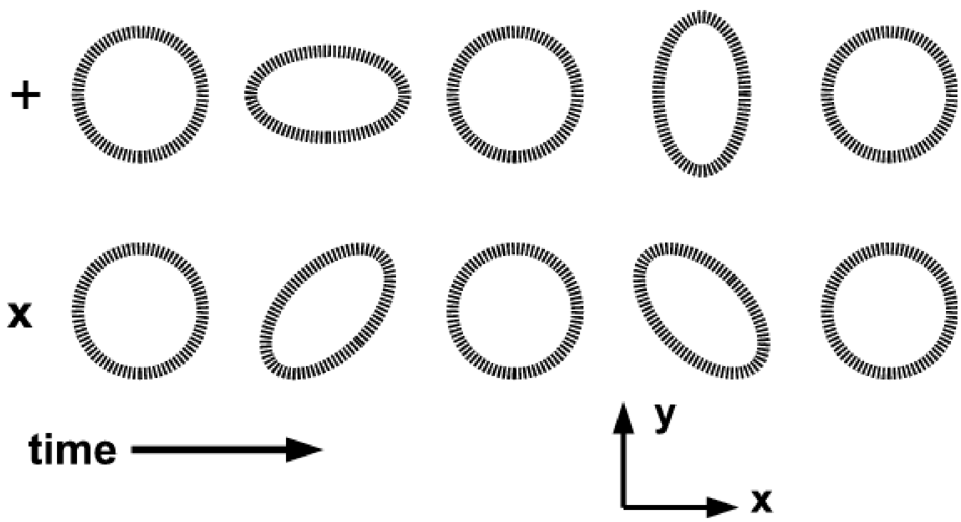
Assuming both test-particles are at rest originally, $u^\mu_A = u^\mu_B = (1,0,0,0)$, we can solve the geodesic equation to linearized order.

$$\mathrm{d}u^\mu/\,\mathrm{d}\tau = -\delta\Gamma^\mu{}_{00} = 0 \tag{23}$$

The last equality is checked easily by explicitly calculating the relevant Christoffel symbols for the metric (22). Since the right hand side in (23) vanishes the test-particles remain at rest and the coordinate distance between A and B does not change. However, the proper distance between them changes (we keep $y = z = 0$).

$$L(t) = \int\limits_0^{L_0} \mathrm{d}x\,\sqrt{1 + f(t)} \qquad \Rightarrow \qquad \frac{L(t) - L_0}{L_0} \approx \frac{1}{2}\, f(t) \tag{24}$$

For periodic functions f the proper distance thus oscillates periodically around its mean value L_0. This is an effect that in principle can be measured, e.g. with LIGO.



Effects of plus and cross polarized gravitational waves on ring of test-particles

5.3 Gravitational wave emission

Like light-waves, gravitational waves need a source. In the former case the source consists of accelerated charges, producing dipole (and higher multipole) radiation, in the latter case the source consists of energy, producing quadrupole (and higher multipole) radiation. The first step is to generalize the wave equation (19) (defining $\tilde{h}_{\mu\nu} := h_{\mu\nu} - \frac{1}{2}\,\eta_{\mu\nu}\,h^\alpha_\alpha$ so that de-Donder gauge reads $\partial^\mu \tilde{h}_{\mu\nu} = 0$) to include an energy-momentum tensor as source

$$\partial^2 \tilde{h}_{\mu\nu} = -16\pi G\, T_{\mu\nu} \tag{25}$$

which for consistency has to obey the conservation equation $\partial^\mu T_{\mu\nu} = 0$.

Up to the decoration with an additional index this is precisely the same situation as in electrodynamics, where the inhomogeneous Maxwell-equations in Lorenz-gauge read $\partial^2 A_\mu = -4\pi\, j_\mu$ and the source has to obey the conservation equation $\partial^\mu j_\mu = 0$. Using the retarded Green function yields

$$\tilde{h}_{\mu\nu}(t,\,\vec{x}) = 4G \int \mathrm{d}^3x'\, \frac{T_{\mu\nu}(t - |\vec{x} - \vec{x}'|,\,\vec{x}')}{|\vec{x} - \vec{x}'|}\,. \tag{26}$$

Thus, we can basically apply nearly everything we know from electrodynamics to gravitational waves. We shall not do this here in great detail, but consider merely one example, the multipole expansion. Taylor-expanding around $\vec{x}' = 0$ the factor $|\vec{x} - \vec{x}'| = r(1 - \vec{x}\cdot\vec{x}'/r^2 + \dots)$ in (26) yields

$$\frac{\tilde{h}_{\mu\nu}(t,\,\vec{x})}{4G} = \frac{1}{r}\int T_{\mu\nu} + \frac{x^i}{r^3}\int x'_i T_{\mu\nu} + \frac{3x^i x^j - r^2\delta^{ij}}{2r^5}\int x'_i x'_j T_{\mu\nu} + \dots \tag{27}$$

The quantities $\int T^{00} = \int \mathrm{d}^3x'\, T^{00}(t - |\vec{x} - \vec{x}'|, \vec{x}') = M$ and $\int T^{0i} = \int \mathrm{d}^3x'\, T^{0i}(t - |\vec{x} - \vec{x}'|, \vec{x}') = P^i$ are mass and momentum of the source. A few lines of calculation establish a formula for $\tilde{h}^{ij}$ in terms of the second time-derivative of the quadrupole moment $Q^{ij}(t) := \int \mathrm{d}^3x' x'^i x'^j T^{00}(t,\,\vec{x}')$ of the source.

$$\tilde{h}^{ij}(t,\,\vec{x}) = \frac{2G}{r}\,\frac{\mathrm{d}^2 Q^{ij}(t)}{\mathrm{d}t^2}\Big|_{t\to t-|\vec{x}-\vec{x}'|} \tag{28}$$

In the far-field approximation (28) describes the dominant part of gravitational radiation.

6 Quantum field theory aspects of spin-2 particles

There are undeniable analogies between Maxwell's theory (a theory of massless spin-1 fields), with the linearized gauge symmetry

$$A_\mu \to A_\mu + \partial_\mu \xi \tag{29}$$

and linearized Einstein gravity on Minkowski background (a theory of massless spin-2 fields), with the linearized gauge symmetry

$$\boxed{h_{\mu\nu} \to h_{\mu\nu} + \partial_{(\mu}\xi_{\nu)}\,.} \tag{30}$$

(This analogy extends to spins higher than 2.) In the remainder of this section we work exclusively in four spacetime dimensions for sake of specificity.

6.1 Gravitoelectromagnetism

As we have shown in section 4.2 in a suitable gauge $h_{\mu\nu}$ obeys the same wave equation as A_μ. In fact, given some observer worldline u^μ one can do a split analogous to electromagnetism into electric part and magnetic part of the Weyl tensor (the Ricci tensor vanishes for vacuum solutions), which in $D = 4$ reads

$$E_{\mu\nu} = C_{\mu\alpha\nu\beta}\, u^\alpha u^\beta \qquad\qquad B_{\mu\nu} = \frac{1}{2}\,\epsilon_{\mu\alpha}{}^{\lambda\gamma}\, C_{\nu\beta\lambda\gamma}\, u^\alpha u^\beta\,. \tag{31}$$

If you want to read more on this formulation see for instance in `gr-qc/9704059`.

6.2 Massive spin-2 QFT

We can gain some insights from looking at the quantum field theory of spin-1 particles (massless or massive QED) and extrapolating results to massless or massive spin-2 particles. (If you are unfamiliar with QED just skip the remainder of this section.) A particular goal of this subsection is to derive that positive charges repel each other while positive masses attract each other just from the spin of the associated exchange particle (spin-1 for electromagnetism, spin-2 for gravity).

To avoid issues with gauge redundancies consider for the moment the massive case. The effective action for massive spin-1 particles is given by

$$W(j) = -\frac{1}{2}\int \frac{\mathrm{d}^4k}{(2\pi)^4}\, j^{\mu\,*}(k)\,\Delta_{\mu\nu}(k)\, j^\nu(k) \tag{32}$$

where j are external currents and $\Delta_{\mu\nu}$ is the propagator,

$$\Delta_{\mu\nu}(k) = \frac{\eta_{\mu\nu} + k_\mu k_\nu/m^2}{k^2 + m^2 - i\epsilon} \tag{33}$$

with the photon mass m and $i\epsilon$ is the prescription to obtain the Feynman propagator. Current conservation $\partial_\mu j^\mu = 0$ implies transversality $k_\mu j^\mu = 0$ so that the second term in the numerator of (33) drops out, yielding

$$W(j) = -\frac{1}{2}\int \frac{\mathrm{d}^4k}{(2\pi)^4}\, j^{\mu\,*}(k)\,\frac{1}{k^2 + m^2 - i\epsilon}\, j_\mu(k)\,. \tag{34}$$

Consider now the situation where the sources are stationary charges so that $j^0 \neq 0$ but $j^i = 0$ (assume further that j^0 is real). Then the result above simplifies to

$$W(j^0) = \frac{1}{2}\int \frac{\mathrm{d}^4k}{(2\pi)^4}\,(j^0)^2\,\frac{1}{k^2 + m^2 - i\epsilon}\,. \tag{35}$$

Actually, the only aspect of interest to us is the sign in (35): it is positive, meaning that there is a positive potential energy between charges of the same sign. Thus, equal charges repel each other.

Since we intend to generalize the considerations above to massive spin-2 particles we need to know their propagator. To this end let us rederive the massive photon propagator (33) using transversality of polarization vectors, $k^\mu \epsilon_\mu^I(k) = 0$, where I runs over all possible polarizations (for massive spin-1 particles $I = 1, 2, 3$) and with no loss of generality we choose $k^\mu = m\,(1, 0, 0, 0)$ and $\epsilon_\mu^I = \delta_\mu^I$. On general grounds, the amplitude for creating a state with momentum k and polarization I at the source is proportional to $\epsilon_\mu^I(k)$, and similarly the amplitude for annihilating a state with momentum k and polarization I at the sink is proportional to $\epsilon_\nu^I(k)$. The numerator in (33) (which determines the residue of the poles) should thus be given by the sum $\sum_I \epsilon_\mu^I(k)\epsilon_\nu^I(k)$. Suppose we did not know the result for the residue. Then we can argue that by Lorentz invariance the result must be given by the sum of two terms, one proportional to $g_{\mu\nu}$ and the other proportional to $k_\mu k_\nu$. Transversality fixes the relative coefficient so that the numerator (and hence the residue) must be proportional to

$$D_{\mu\nu} = \eta_{\mu\nu} + k_\mu k_\nu / m^2 \,. \tag{36}$$

The overall normalization is determined to be $+1$, e.g. from considering the component $\mu = \nu = 1$. This concludes our derivation of the residue of the pole in the massive spin-1 propagator (33). The location of the pole itself just follows from the wave equation; the $i\epsilon$ prescription is the least obvious aspect, but standard since Feynman's time. If you are unfamiliar with it consult some introductory QFT book, like Peskin & Schroeder.

We do now the same calculation for massive spin-2 particles, where the analog of the effective action (32) reads

$$W(T) = -\frac{1}{2}\int \frac{\mathrm{d}^4k}{(2\pi)^4}\, T^{\mu\nu}(k)\,\Delta_{\mu\nu\alpha\beta}(k)\,T^{\alpha\beta}(k)\,. \tag{37}$$

The source is now the energy-momentum tensor $T^{\mu\nu}$. Source and propagator have twice as many indices as compared to the spin-1 case.

Our first task is to determine the propagator $\Delta_{\mu\nu\alpha\beta}(k)$. We use the spin-2 polarization tensor $\epsilon_{\mu\nu}$, which has to be transverse, traceless and symmetric.

$$k^\mu \epsilon_{\mu\nu} = 0 \qquad \epsilon^\mu{}_\mu = 0 \qquad \epsilon_{\mu\nu} = \epsilon_{\nu\mu} \tag{38}$$

This means that we have 5 independent components in $\epsilon_{\mu\nu}$ corresponding to the 5 spin-2 helicity states. We introduce again a label I to discriminate between these 5 helicity states, $\epsilon_{\mu\nu}^I(k)$ and allow for k-dependence (fixing the normalization e.g. by $\sum_I \epsilon_{12}^I \epsilon_{12}^I = 1$). It is then a straightforward exercise [exploiting the properties (38)] to perform the sum over all helicities

$$\sum_{I=1}^{5} \epsilon_{\mu\nu}^I \epsilon_{\alpha\beta}^I = D_{\mu\alpha}D_{\nu\beta} + D_{\mu\beta}D_{\nu\alpha} - \tfrac{2}{3}\,D_{\mu\nu}D_{\alpha\beta} \tag{39}$$

where $D_{\mu\nu}$ is the same expression as in (36). The overall normalization was fixed again by considering a specific example, e.g. evaluating (39) for $\mu = \alpha = 1$ and $\nu = \beta = 2$. This means that the massive spin-2 (Feynman) propagator is given by

$$\Delta_{\mu\nu\alpha\beta}(k) = \frac{D_{\mu\alpha}D_{\nu\beta} + D_{\mu\beta}D_{\nu\alpha} - \frac{2}{3}\,D_{\mu\nu}D_{\alpha\beta}}{k^2 + m^2 - i\epsilon}\,. \tag{40}$$

Our second task is to consider the interaction between two sources of energy. For simplicity assume that only $T^{00} \neq 0$ and all other components of $T^{\mu\nu}$ vanish.

Then inserting the massive spin-2 propagator (40) into the effective action (37) yields [using transversality $k_\mu T^{\mu\nu} = 0$ only the η-term in (36) contributes]

$$W(T^{00}) = -\frac{1}{2}\int\frac{\mathrm{d}^4k}{(2\pi)^4}\,(T^{00}(k))^2\,\frac{1+1-\frac{2}{3}}{k^2+m^2-i\epsilon}\,. \tag{41}$$

Since all numerator terms in the integrand are positive, the overall sign of the potential energy $W(T^{00})$ is opposite to that of the potential energy $W(j_0)$ in the spin-1 case (35). Thus, **gravity is attractive for positive energy because the exchange particle has spin-2.**

It is remarkable that we were able to conclude the attractiveness of gravity merely from the statement that its exchange particle has spin-2. Of course, there is a gap in the logic above: we have proved this statement so far only for massive gravitons, but Einstein gravity has massless gravitons.

6.3 Massless spin-2 QFT and vDVZ-discontinuity

It may be tempting to conclude that the difference between a massless spin-2 particle and a massive one is negligible if the mass in sufficiently small. Actually this conclusion is correct, but in a highly non-trivial way, which we address here.

Let us consider first the massless spin-2 propagator, which we can read off from the results in section 5.1 [or from (17) together with a gauge-fixing term].

$$\Delta^0_{\mu\nu\alpha\beta}(k) = \frac{\eta_{\mu\alpha}\eta_{\nu\beta}+\eta_{\mu\beta}\eta_{\nu\alpha}-\eta_{\mu\nu}\eta_{\alpha\beta}}{k^2+m^2-i\epsilon} + \text{possibly } k_\lambda\text{-terms } (\lambda=\mu,\nu,\alpha,\beta) \tag{42}$$

The main difference to the massive spin-2 propagator (40) is that the factor $-\frac{2}{3}$ in the last term is replaced by -1 here, which causes a discontinuity, as it persists for arbitrarily small non-zero masses. This effect is called **van Dam–Veltman–Zakharov discontinuity**.

Should we care about this discontinuity? Consider the interaction between two particles with stress tensors $T^{\mu\nu}_{1,2}$ exchanging a massive spin-2 particle in the limit of vanishing mass versus them exchanging a massless spin-2 particle:

$$\text{massive } (m\to 0):\qquad T_1^{\mu\nu}\Delta_{\mu\nu\alpha\beta}T_2^{\alpha\beta} = \frac{1}{k^2}\left(2T_1^{\mu\nu}T_{2\,\mu\nu} - \tfrac{2}{3}\,T_1T_2\right) \tag{43}$$

$$\text{massless:}\qquad T_1^{\mu\nu}\Delta^0_{\mu\nu\alpha\beta}T_2^{\alpha\beta} = \frac{1}{k^2}\left(2T_1^{\mu\nu}T_{2\,\mu\nu} - T_1T_2\right) \tag{44}$$

Thus, for the gravitational interaction of massless particles ($T_1 = T_2 = 0$) there is no difference between the exchange of (tiny) massive and massless spin-2 particles, but for massive particles ($T_1 \neq 0 \neq T_2$) there is a difference by a factor of order unity. This factor of order unity should have shown up in the classical tests (light-bending and perihelion shift). So can we conclude from the vDVZ-discontinuity that experimentally the graviton must be exactly massless?

The answer is no. While Einstein gravity predicts massless gravitons, we cannot be sure experimentally whether or not the graviton is exactly massless or has a tiny non-zero mass. The issue why the vDVZ-discontinuity does not contradict this statement was resolved by Vainshtein. His key insight was that massive spin-2 theories with some central object of mass M come with an intrinsic distance scale, given by $r_V = (GM)^{1/5}/m^{4/5}$ also known as "Vainshtein radius" (G is Newton's constant and m the graviton mass). The difference between Einstein gravity and massive spin-2 theories is negligible inside the Vainshtein radius, which can be arbitrarily large if m is tiny. The approximations we made above using massive spin-2 exchange are only valid outside the Vainshtein radius; within the Vainshtein radius the higher order terms in the expansion analogous to (18) are not negligible.

7 Black hole perturbations and quasi-normal modes

In our perturbative treatment around some fixed background we have focussed so far on maximally symmetric backgrounds, i.e., Minkowski space or (A)dS. Another obvious set of interesting backgrounds is provided by black holes. If you have a large black hole and you throw some small perturbation (say, a spaceship) into it or scatter some wave on the black hole, then you expect the black hole to be only slightly modified. It turns out that the black starts to ring like a bell when perturbed, but with damped oscillations, so-called quasi-normal modes. They are characteristic for a black hole in much the same way as normal modes are characteristic for systems described by a bunch of harmonic oscillators.

The purpose of this section is to develop the theory of black hole perturbations and in particular derive equations for quasi-normal modes. Possible applications include gravitational wave emission of black hole binaries, stability investigations of black holes, scattering and absorption of waves by black holes, late-time behavior of the gravitational field after black hole formation, radiation generated by objects falling into black holes and various holographic applications.

7.1 Scalar perturbations of Schwarzschild black holes

In order to illuminate the main concepts and tools let us consider the simplest black hole in four spacetime dimensions, the Schwarzschild black hole.

$$\mathrm{d}s^2 = -(1-2M/r)\,\mathrm{d}t^2 + \mathrm{d}r^2/(1-2M/r) + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right) \tag{1}$$

Let us further assume that we perturb this black hole by switching on a free scalar field ϕ propagating on that background, i.e., obeying the Klein–Gordon equation.

$$\nabla^2\phi = 0 \qquad \Rightarrow \qquad \partial_\mu\big(\sqrt{-g}g^{\mu\nu}\partial_\nu\phi\big) = 0 \tag{2}$$

Exploiting spherical symmetry we decompose the scalar field into spherical harmonics

$$\phi_{lm} = \frac{\psi_l(t,\,r)}{r}\,Y_{lm}(\theta,\,\varphi) \tag{3}$$

where we have pulled out a convenient factor $1/r$. Using the tortoise coordinate r_* (see Black Holes I),

$$r_* = r + 2M\,\ln\big(r/(2M)-1\big) \tag{4}$$

the functions ψ_l obey wave equations

$$\Big(-\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial r_*^2} - V_l(r)\Big)\,\psi_l = 0 \tag{5}$$

with the effective potential

$$V_l(r) = \Big(1-\frac{2M}{r}\Big)\Big(\frac{l(l+1)}{r^2} + \frac{2M}{r^3}\Big)\,. \tag{6}$$

Since the wave equation (5) is linear it is useful to decompose the functions $\psi_l(t,\,r)$ into plane waves $\hat\psi_l(r,\omega)e^{-i\omega t}$. Note that ω is complex in general. All that remains to be solved is a second order ODE.

$$\Big(\frac{\mathrm{d}^2}{\mathrm{d}r_*^2} + \omega^2 - V_l(r)\Big)\,\hat\psi_l(r,\,\omega) = 0 \tag{7}$$

Thus, we have reduced the problem to scattering on the potential (6), displayed in Fig. 1 on the next page. This type of problem you have encountered in basic lectures on quantum mechanics. Equation (7) is known as Regge–Wheeler equation.

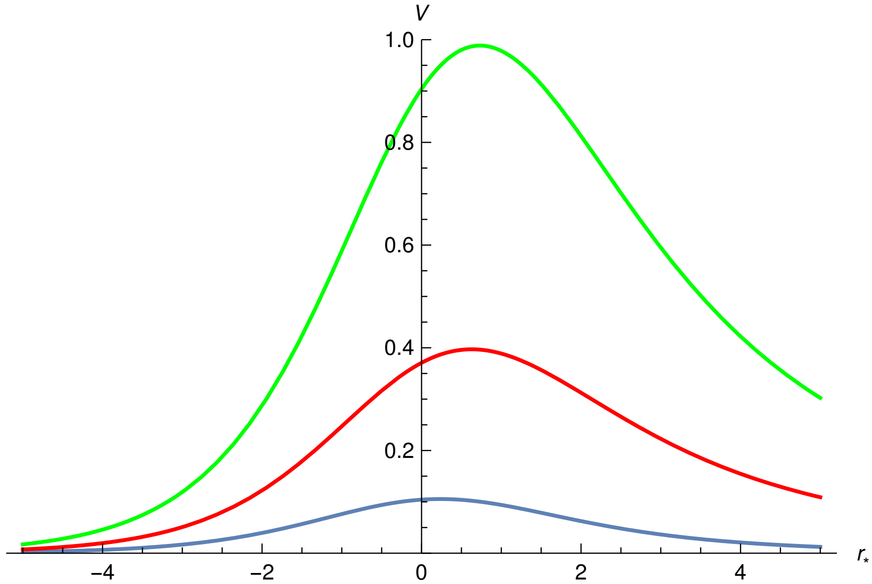


Figure 1: Potential V_l [width $M = 1/2$] for $l = 0$ (lowest/blue curve), $l = 1$ (middle/red curve) and $l = 2$ (upper/green curve) as function of r_*. The horizon is at $r_* \to -\infty$, the asymptotic region at $r_* \to +\infty$. Note that the potential is positive everywhere and has a maximum close to the photon sphere $r = 3M$. The form of the potential implies that there are no discrete normalizable bound states.

Let us thus apply general insights from quantum mechanical potential scattering for barrier potentials displayed above. If the waves have short wave-length as compared to the Schwarzschild radius we expect them to be easily transmitted through the potential barrier. Waves with wavelengths of order of the Schwarzschild radius will be partly transmitted and partly absorbed. Waves with long wavelengths will be reflected almost completely by the potential barrier. Moreover, since the effective potential V_l vanishes for $r_* \to \pm\infty$ (i.e., for $r \to 2M$ and $r \to \infty$) we know that two linearly independent solutions to (7) asymptotically behave as $\hat{\psi}_l \sim e^{\pm i\omega r_*}$. Thus, we can make the following ansatz for infalling modes

$$\lim_{r_* \to -\infty} \hat{\psi}_l = e^{-i\omega r_*} \qquad \lim_{r_* \to \infty} \hat{\psi}_l = \left(A_R(\omega) e^{i\omega r_*} + e^{-i\omega r_*}\right) \frac{1}{A_T(\omega)} \tag{8}$$

having normalized the mode conveniently at the horizon (where it is ingoing!) and introduced reflection and transmission amplitudes A_R and A_T, respectively, to parametrize the asymptotic amplitudes. A second set of modes is given by the complex conjugate of (8). Since the Wronskian of two linearly independent solutions of (7) is constant, evaluation at $r_* = \pm\infty$ yields a quadratic relation between transmission and reflection amplitudes (which again should look familiar from quantum mechanics),

$$|A_R|^2 + |A_T|^2 = 1\,. \tag{9}$$

The squares of the amplitudes are reflection and transmission probabilities.

Note that the modes defined by (8) can be interpreted as waves propagating from $\mathscr{I}^-$ towards the future event horizon (transmission), partly being scattered to $\mathscr{I}^+$ (reflection). The complex conjugate of these modes corresponds to waves approaching $\mathscr{I}^+$ and emanating partly from the past event horizon ("transmission") and partly from $\mathscr{I}^-$ ("reflection"). Finally, note that one can also define two additional sets of modes where the limits $r_* \to \pm\infty$ are exchanged as compared to the definition (8) and its complex conjugate. They describe waves that are only ingoing at the future horizon or only outgoing at the past horizon. Particularly the modes that are only ingoing at the future horizon have interesting physics applications.

7.2 Quasi-normal modes

Quasi-normal modes are perturbations of a black hole that are ingoing on the future event horizon and outgoing at infinity. They are of particular interest, since they capture the response of a black hole under small perturbations ("ringing"). Before discussing them it is useful to recall basic features of normal modes.

Many physical systems are well-described by harmonic oscillators. Compact systems governed by harmonic motion can conveniently be decomposed in terms of normal modes

$$\psi(t,r) = \sum_{n=1}^{\infty} \psi_n(r)\, e^{-i\omega_n t} \tag{10}$$

where all ω_n are real. The function $\psi(t,r)$ describes the state of the system. Perhaps the simplest example is a string of finite length with fixed endpoints. Non-compact systems require a bit more care. Consider the wave equation (5) (dropping the subscripts l) with vanishing potential $V = 0$. Then the spectrum is continuous, but we can still use as basic building blocks plane waves and represent general solutions as (continuous) superpositions of these plane waves. Thus, for non-compact systems plane waves (again with real frequencies) are analogs of normal modes for compact systems, which is the essence of Fourier analysis.

For the definition of quasi-normal modes consider the wave equation (5) with non-negative finite $V \geq 0$, assuming compact support, i.e., $V(r_*) = 0$ for $|r_*| > r_0$. All solutions ψ are then bounded and we can employ a Laplace transformation.

$$\tilde{\psi}(s,r_*) = \int\limits_0^\infty \mathrm{d}t\, e^{-st}\, \psi(t,r_*) \tag{11}$$

The Laplace transform $\tilde{\psi}(s,r_*)$ of solutions $\psi(t,r_*)$ obeys

$$s^2\tilde{\psi} - \tilde{\psi}'' + V\,\tilde{\psi} = s\,\psi(0,r_*) + \partial_t\psi(0,r_*) =: j(s,r_*) \tag{12}$$

where the right hand side contains the initial data $\psi(0,r_*)$ and $\partial_t\psi(0,r_*)$, and prime denotes derivative with respect to r_*. Boundedness of ψ implies analyticity of $\tilde{\psi}$ in the complex half-plane $\mathrm{Re}(s) > 0$. The homogeneous version of (12) reads

$$s^2\tilde{\psi} - \tilde{\psi}'' + V\,\tilde{\psi} = 0\,. \tag{13}$$

To solve (12) we consider its Green function.

$$G(s,r_*,r_*') = \frac{1}{W(s)}\big(f_+(s,r_*)f_-(s,r_*')\theta(r_*-r_*') + f_+(s,r_*')f_-(s,r_*)\theta(r_*'-r_*)\big) \tag{14}$$

Here $f_\pm$ are suitable solutions to the homogeneous equation (13) and $W(s)$ is the Wronskian of $f_\pm$. The general solution to (12) is then given by

$$\tilde{\psi}(s,r_*) = \int\limits_{-\infty}^{\infty} \mathrm{d}r_*'\, G(s,r_*,r_*')\, j(s,r_*') \tag{15}$$

It remains to be clarified what "suitable solutions" means. We need to select a unique pair of solutions $f_\pm$ compatible with all our assumptions. The Laplace transformation (11) guarantees that $\tilde{\psi}$ is bounded as function of r_*. Compact support of the potential implies that for $|r_*| > r_0$ solutions to (13) behave as

$$f \sim e^{\pm s r_*}\,. \tag{16}$$

Thus, we have the following unique pair of linearly independent and bounded solutions:

$$f_\pm = e^{\mp s r_*} \quad \text{for } \pm r_* > r_0\,. \tag{17}$$

For the relevant domain $\text{Re}(s) > 0$ the solution f_+ (the solution f_-) is decaying at large positive (large negative) r_*.

Now we are ready to define quasi-normal mode frequencies s_n. They are defined as complex number for which the two solutions $f_\pm$ become linearly dependent, i.e., the Wronskian vanishes.

$$f_+(s_n, r_*) = A_n(s_n)\, f_-(s_n, r_*) \tag{18}$$

The corresponding solutions $f_+(s_n, r_*)$ are referred to as quasi-eigenfunctions. At this stage we should worry about their existence. After all, by construction we have a unique regular Green function (14) for $\text{Re}(s) > 0$ so that in this part of the complex plane it is impossible to obey (18) since necessarily both solutions $f_\pm$ are linearly independent there. However, one can show that $f_\pm$ have unique analytic continuations into the full complex plane. Moreover, there is a theorem (see A. Bachelot and A. Motet-Bachelot, *Ann.Inst.H.Poincare Phys.Theor.* **59** (1993) 3) that for non-negative potentials V with compact support there is always a countable number of zeros of the Wronskian in the half-plane $\text{Re}(s) < 0$.

To appreciate the physical significance of quasi-eigenfunctions consider the inverse Laplace trafo for some positive $a > 0$

$$\psi(t, r_*) = \frac{1}{2\pi i} \int\limits_{-\infty}^{\infty} \text{d}s\, e^{(a+is)t}\, \hat{\psi}(a+s, r_*) \tag{19}$$

where the complex line integral along s can be suitably deformed to show the following behavior of the function ψ:

$$\psi(t, r_*) \sim \sum_n A_n\, e^{s_n t}\, f_+(s_n, r_*) \tag{20}$$

where

$$s_n = \kappa_n + i\omega_n \qquad \text{with } \kappa_n < 0 \tag{21}$$

and the sum extends over all quasi-normal mode frequencies. The functions $\psi(t, r_*)$ in (20) are called quasi-normal modes.

The result (20), (21) shows that quasi-normal modes decay exponentially in time. Thus, at very late times the behavior of the system under consideration is dominated by the lowest lying quasi-normal mode, i.e., by the mode with the quasi-normal frequency that has the largest (or least negative) real part $\kappa_{n_{\min}}$.

In the case of Schwarzschild the potential V given in (6) does not have compact support, so one needs to extend the discussion above to cases of non-compactly supported V that decay at infinity. While it is non-trivial to make this statement more precise, it is plausible that for sufficiently fast fall-off to zero there will be again quasi-normal modes. Indeed, for Schwarzschild black holes quasi-normal modes do exist and were constructed numerically by Nollert, *Phys. Rev.* **D** (1993) 5253 and Andersson *Class. Quant. Grav.* **10** (1993) L61 and in the limit of large damping analytically by Motl and Neitzke, hep-th/0301137.

While there are numerous other applications, the main applications of quasi-normal modes within black hole physics include gravitational wave emission in the "ring-down" phase after a black hole merger (i.e., exponential decay towards a stationary black hole governed by the lowest-lying quasi-normal modes) and tests of the AdS/CFT correspondence (where the black hole quasi-normal frequencies coincide with poles of the retarded Green function of the dual CFT, see the paper by Birmingham, Sachs and Solodukhin hep-th/0112055; see also the more recent work by Janik, Jankowski and Soltanpanahi 1603.05950). For an older review article on quasi-normal modes and more details see gr-qc/9909058.

7.3 Generalizations

In these lectures we considered only spin-0 (=scalar) perturbations around Schwarzschild black holes, but the same techniques work for perturbations of higher spin (including gravitational perturbations), while different techniques may be needed for other black holes (in particular Kerr, the phenomenologically most important black hole). Rather then deriving such generalizations below we merely quote some key results, whose derivation conceptually is along the lines of the previous sections.

A generalization of the Regge–Wheeler potential (6) to arbitrary spin s is given by

$$V_l^s(r) = \Big(1 - \frac{2M}{r}\Big)\Big(\frac{l(l+1)}{r^2} + \frac{2M(1-s^2)}{r^3}\Big)\,. \tag{22}$$

For $s = 0$ it coincides with (6), for $s = 1$ (photons) the last term vanishes and for $s = 2$ (gravitons) the last term changes its sign as compared to scalar perturbations. Note, however, that there are two kinds of gravitational perturbations: axial ones (which induce rotation; they are parity odd) and polar ones (which do not induce rotation; they are parity even); this nomenclature was introduced by Chandrasekhar, see *The mathematical theory of black holes* Oxford Science Publications (1985). It turns our that the axial perturbations are indeed governed by the Regge–Wheeler potential (22) with $s = 2$, while the polar ones are governed by the so-called Zerilli-potential ($n := \frac{1}{2}(l-1)(l+2)$)

$$V_l^z(r) = 2\Big(1 - \frac{2M}{r}\Big)\frac{n^2(n+1)r^3 + 3n^2Mr^2 + 9nM^2r + 9M^3}{r^3(nr+3M)^3}\,. \tag{23}$$

For Kerr black holes (see Black Holes I: $\Sigma = r^2 + a^2\cos^2\theta$, $\Delta = r^2 - 2Mr + a^2$)

$$\mathrm{d}s^2 = -\frac{\Delta}{\Sigma}\left(\mathrm{d}t - a\sin^2\theta\,\mathrm{d}\varphi\right)^2 + \frac{\Sigma}{\Delta}\,\mathrm{d}r^2 + \Sigma\,\mathrm{d}\theta^2 + \frac{\sin^2\theta}{\Sigma}\left((r^2+a^2)\,\mathrm{d}\varphi - a\,\mathrm{d}t\right)^2 \tag{24}$$

peturbations are solutions to the Teukolsky equation, which you can find e.g. in this paper by Fiziev, 0908.4234 (see also Refs. therein). This is not the place to go into details of (or even display in its full glory) the Teukolsky equation. We merely focus on one physically important detail concerning scalar perturbations.

Again the ansatz (8) works for infalling modes (where now the tortoise coordinate is given by $\mathrm{d}/\mathrm{d}r_* = \Delta/(r^2+a^2)\,\mathrm{d}/\mathrm{d}r$), except that in the limit $r_* \to -\infty$ there is a shift $\omega \to \omega - m\Omega$, where m is the magnetic quantum number of the perturbation and $\Omega = a/(r_+^2 + a^2)$ is the angular velocity of the outer horizon, with $r_+ = M + \sqrt{M^2 - a^2}$, see exercise sheet 10 of Black Holes I. Constancy of the Wronskian now leads to a condition slightly different from (9), namely

$$|A_R|^2 + \Big(1 - \frac{m\Omega}{\omega}\Big)\,|A_T|^2 = 1\,. \tag{25}$$

If the inequality

$$m\Omega > \omega \tag{26}$$

holds, plugging this inequality into relation (25),

$$|A_R|^2 - 1 = \Big(\frac{m\Omega}{\omega} - 1\Big)\,|A_T|^2 > 0 \tag{27}$$

shows that $|A_R| > 1$ in this case. Thus, the reflected amplitude is bigger than the incoming one! This is called "superradiant scattering" and allows to extract energy from a Kerr black hole.

For further aspects of black hole perturbation theory — like black hole stability or gravitational waves from black hole binaries — see chapter 4 of Frolov & Novikov, *Black holes physics*, Kluwer Academic Publishers (1998) and Refs. therein.

8 Black hole thermodynamics

The insight that black holes have a (Hawking) temperature and a (Bekenstein–Hawking) entropy has profoundly influenced our understanding of black holes and our path on the road towards quantum gravity. Despite of their classical simplicity, captured by "no hair" theorems, quantum mechanically black holes are not only complicated, but arguably the most complex entities that could possibly exist in our (or any other) Universe. Sometimes the analogy is made that understanding the thermodynamics of black holes quantum mechanically could play the same role for the development of quantum gravity as the quantum mechanical understanding of the Hydrogen atom in the development of quantum mechanics. Regardless of whether this turns out to be true, black hole thermodynamics certainly is a cornerstone in reasonable attempts to quantize gravity and has found applications in AdS/CFT and black hole analogs.

The reason why there are no astrophysical applications in the current phase of our Universe is the smallness of the Hawking temperature for black holes whose mass is larger than the mass of our Sun (which applies to all astrophysical black holes detected so far and must be true if the black hole results from gravitational collapse of a star, see the beginning of Black Holes I).

In this section we work out classical aspects of black hole thermodynamics, starting with the four laws.

See 1402.5127 and Refs. therein for more on black hole thermodynamics.

8.1 Four laws of black hole mechanics and thermodynamics

In Black Holes I we derived a version of the zeroth law of black hole mechanics (surface gravity κ is constant for stationary black holes) and in Black Holes II we mentioned the proof idea of the second law of black hole mechanics. The third law states that it is impossible to reach a black hole state of vanishing surface gravity from an initial black hole with non-vanishing surface gravity in finite time (see one of the exercises). We focus now on the missing item, the first law of black hole mechanics.

As a preparation we consider Smarr's formula

$$M = \frac{\kappa A}{4\pi} + 2\Omega J \tag{1}$$

for Kerr black holes with mass M, angular momentum J, event horizon area A, surface gravity κ and angular velocity of the horizon Ω. Smarr's formula can be derived using the Komar integrals we introduced in Black Holes I for the Killing vector $\partial_t + \Omega\partial_\varphi$, but since we know already all the results for Kerr black holes we can easily verify (1) simply by expressing all quantities in terms of outer and inner horizon radii. Recalling

$$M = \frac{r_+ + r_-}{2} \qquad J(= aM) = \frac{r_+ + r_-}{2}\,\sqrt{r_+ r_-} \qquad A = 4\pi\left(r_+^2 + r_+ r_-\right)$$
$$\kappa = \frac{r_+ - r_-}{2(r_+^2 + r_+ r_-)} \qquad \Omega = \frac{\sqrt{r_+ r_-}}{r_+^2 + r_+ r_-} \tag{2}$$

allows to verify that (1) indeed holds for all values of $r_\pm$.

We state now a simplified version of the first law. Take a stationary (Kerr) black hole of mass M and angular momentum J and perturb it infinitesimally by changing to mass $M + \delta M$ and angular momentum $J + \delta J$. Then the change of the area δA is related linearly to δM and δJ through the first law as follows,

$$\boxed{\delta M = \frac{\kappa}{8\pi}\,\delta A + \Omega\,\delta J} \tag{3}$$

where κ is surface gravity and Ω the angular velocity of the horizon. The proof of the first law can be found in the paper by Bardeen, Carter and Hawking.[1]

Instead of an actual proof we present here a slick derivation that is due to Gibbons. Black hole uniqueness implies that M, J and A cannot be independent from each other since (Kerr) black holes are uniquely specified by providing two of these numbers. Thus, either of them must be a function of the two others. For instance, $M = M(J, A)$. Now, A and J have both dimensions of mass squared (we are in four spacetime dimensions right now). This implies that the function $M(J, A)$ is homogeneous of degree $\frac{1}{2}$. Euler's theorem for homogeneous functions establishes

$$J\,\frac{\partial M}{\partial J} + A\,\frac{\partial M}{\partial A} = \frac{1}{2}\,M = \frac{\kappa}{8\pi}\,A + \Omega J \tag{4}$$

where the last equality follows from Smarr's formula (1). We can rewrite (4) suggestively

$$J\,\Big(\frac{\partial M}{\partial J} - \Omega\Big) + A\,\Big(\frac{\partial M}{\partial A} - \frac{\kappa}{8\pi}\Big) = 0 \tag{5}$$

and then argue that both terms in (5) have to vanish separately since the coefficients J and A are arbitrary and independent from each other. If you buy this argument then you obtain the desired result

$$\frac{\partial M}{\partial J} = \Omega \qquad\qquad \frac{\partial M}{\partial A} = \frac{\kappa}{8\pi} \tag{6}$$

which establishes the first law (3).

More general black holes may also depend on the electric charge and be immersed in something other than Minkowski space, e.g. in (A)dS space. In all these cases there is a first law of the form

$$\delta M = \frac{\kappa}{8\pi}\,\delta A + \text{work terms}\,. \tag{7}$$

We summarize now the four laws of black hole mechanics and contrast them with the four laws of thermodynamics.

	black hole mechanics	thermodynamics
0^{th}	$\kappa = \text{const.}$	$T = \text{const.}$
1^{st}	$\delta M = \frac{\kappa}{8\pi}\,\delta A + \text{work terms}$	$\delta E = T\,\delta S + \text{work terms}$
2^{nd}	$\delta A \geq 0$	$\delta S \geq 0$
3^{rd}	$\kappa \to 0$ impossible	$T \to 0$ impossible

Comparing left and right columns it is tempting to identify surface gravity with temperature, $\kappa \sim T$, area with entropy, $A \sim S$ and mass with energy, $M \sim E$. Actually, we know that at least the last identification is correct, thanks to Einstein's most famous formula $E = M$ (in units of $c = 1$). Moreover, note that there are non-trivial consistency checks of this identification — for example, κ plays the role of T not only in the 0^{th} law, but also in the 1^{st} and 3^{rd} law. Should we therefore take the analogy displayed in the table above seriously? The naive answer is yes, the more sophisticated answer is no (see the footnote on this page for the reason) and the correct answer is again yes. However, to show this we need to take into account quantum fluctuations on black hole backgrounds in order to derive the Hawking effect, the Hawking–Unruh temperature and the Bekenstein–Hawking entropy.

[1] This paper is not only nice, but also remarkable since it contains the statement "In fact the effective temperature of a black hole is absolute zero. One way of seeing this is to note that a black hole cannot be in equilibrium with black body radiation at any non-zero temperature, because no radiation could be emitted from the hole whereas some radiation would always cross the horizon into the black hole." that was famously falsified by its last author about a year later.

8.2 Phenomenological aspects of black hole thermodynamics

Before we delve into semi-classical aspects associated with the Hawking effect we address phenomenological aspects of black hole thermodynamics. Let us start with the Schwarzschild black hole. According to our previous discussion we have

$$T \sim \frac{1}{M} \qquad S \sim M^2 \quad \Rightarrow \quad S \sim \frac{1}{T^2} \tag{8}$$

where the similarity signs remind us that we do not know the factors of order unity in these identifications (all we know is $\kappa A = 8\pi TS$). Interesting observations:

1. For stellar mass black holes the temperature is tiny, $T \sim 10^{-38} \approx 10^{-6}\,$Kelvin $\ll T_{\text{CMB}} \approx 3\,$Kelvin. (Once all factors are considered the result for a stellar mass black hole is $T \approx 61.7\,$nanoKelvin.) Thus, we do not expect to ever detect Hawking radiation from stellar mass black holes (nor from heavier ones).

2. For a stellar mass black hole the entropy is ridiculously large, $S \sim 10^{76}$, which means that we have a googolplex-like number of microstates, $N \sim e^{10^{76}}$.

3. The Bekenstein–Hawking entropy is not extensive in the usual way, i.e., it does not scale like the volume of the black hole but rather like its area. This observation is the seed of the holographic principle, which states that quantum gravity in, say, four spacetime dimensions is equivalent to some quantum field theory in three spacetime dimensions (where then the area is reinterpreted as volume of the lower-dimensional theory).

4. The Schwarzschild black hole has negative specific heat.

$$C = \frac{\mathrm{d}M}{\mathrm{d}T} \sim -\frac{1}{T^2} < 0 \tag{9}$$

This statement just rephrases the fact that the more a Schwarzschild black hole radiates (and hence the more it reduces its mass) the warmer it gets. Thus, by itself the Schwarzschild black hole is thermodynamically unstable, but we should not worry too much about this given how tiny the specific heat is. It is possible to stabilize the Schwarzschild black hole by putting it into a box (either literally or by providing AdS asymptotics, see below).

Charged (Reissner–Nordström) or rotating (Kerr or Kerr–Newman) black holes have additional interesting features. There are now work terms present associated with changes of the charge or angular momentum. Moreover, we can have extremal solutions where temperature vanishes, but which are macroscopically large and thus have a huge entropy. For example,

$$S_{\text{extremal Kerr}} \sim A = 4\pi \left(r_+^2 + r_+ r_-\right) = 8\pi r_+^2 = 8\pi M^2 \gg 1\,. \tag{10}$$

No analog condensed matter system is known which at zero temperature has such a large degeneracy of states.

Finally, let us briefly consider black holes in AdS; for simplicity consider Schwarzschild-AdS, whose metric is given by (ℓ is the AdS_4 radius)

$$\mathrm{d}s^2 = -\Big(\frac{r^2}{\ell^2} + 1 - \frac{2M}{r}\Big)\,\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{\frac{r^2}{\ell^2} + 1 - \frac{2M}{r}} + r^2\,\mathrm{d}\Omega^2_{S^2}\,. \tag{11}$$

Calculating the specific heat in the limit of small masses recovers the negative sign of (9), $C \sim -M^2 + \mathcal{O}(M^4/\ell^2)$. Interestingly, in the limit of large masses specific heat is positive, $C \sim \ell^{4/3} M^{2/3} + \mathcal{O}(\ell^{8/3}/M^{2/3})$. This suggests that there could be a phase transition at some finite value of the mass, $M/\ell \sim \mathcal{O}(1)$, which indeed exists and is known as Hawking–Page phase transition.

9 Hawking effect

Black holes at finite surface gravity κ emit radiation that to leading order approximation is thermal. This is known as Hawking effect. The purpose of this section is to calculate the Hawking temperature in terms of surface gravity, i.e., to determine the precise $\mathcal{O}(1)$ coefficient in the relation $\kappa \sim T$. We shall do this in two different ways, by Euclidean continuation and by a semi-classical calculation of scalar field fluctuations on a black hole background.

9.1 Periodicity in Euclidean time is inverse temperature

Quantum mechanically unitary time evolution of some state $|\psi(0)\rangle$ is generated by some Hermitean Hamiltonian H,

$$|\psi(t)\rangle = e^{iHt}|\psi(0)\rangle\,. \tag{1}$$

Quantum statistically, the partition function is defined by a trace over the Boltzmann factor $e^{-\beta H}$, where $\beta = T^{-1}$ is inverse temperature,

$$Z = \operatorname{tr}\big(e^{-\beta H}\big) = \sum_\psi \langle\psi(0)|e^{-\beta H}|\psi(0)\rangle = \sum_\psi e^{-\beta E_\psi} \tag{2}$$

where the sum is over a complete set of states $|\psi(0)\rangle$ and E_ψ are energy eigenvalues. The key observation here is that the Boltzmann factor can be reinterpreted as time evolution of the state $|\psi(0)\rangle$ over the imaginary time period $-i\beta$, thus yielding

$$Z = \sum_\psi \langle\psi(0)|\psi(-i\beta)\rangle = \sum_\psi \langle\psi(+i\beta)|\psi(0)\rangle\,. \tag{3}$$

Given the expressions (3) for the partition function it is suggestive to impose periodicity in the imaginary part of time,

$$t \sim t - i\beta \quad \Rightarrow \quad \tau \sim \tau + \beta \quad \text{where } \tau = it\,. \tag{4}$$

Periodicity in Euclidean time τ is identical to inverse temperature β.

Actually, for those who know a bit of QFT let us be more concrete and consider the Green function of a free theory at finite temperature,

$$G(x-y) = \frac{\sum_\psi \langle\psi|T(\phi(x)\phi(y))|\psi\rangle e^{-\beta E_\psi}}{\sum_\psi e^{-\beta E_\psi}} = \frac{1}{Z}\operatorname{tr}\big(e^{-\beta H}T(\phi(x)\phi(y))\big) \tag{5}$$

where the $|\psi\rangle$ are eigenstates of H with eigenvalues E_ψ and T denotes time-ordering. We then get the following chain of identities (assuming $x^0 > 0$ we can drop time ordering in the first step)

$$\begin{aligned} G(x^0,\vec{x};0,\vec{y}) &= \frac{1}{Z}\operatorname{tr}\big(e^{-\beta H}\phi(x^0,\vec{x})\phi(0,\vec{y})\big) = \frac{1}{Z}\operatorname{tr}\big(\phi(0,\vec{y})e^{-\beta H}\phi(x^0,\vec{x})\big) \\ &= \frac{1}{Z}\operatorname{tr}\big(e^{-\beta H}e^{\beta H}\phi(0,\vec{y})e^{-\beta H}\phi(x^0,\vec{x})\big) = \frac{1}{Z}\operatorname{tr}\big(e^{-\beta H}\phi(-i\beta,\vec{y})\phi(x^0,\vec{x})\big) \\ &= \frac{1}{Z}\operatorname{tr}\big(e^{-\beta H}T(\phi(x^0,\vec{x})\phi(-i\beta,\vec{y}))\big) = G(x^0,\vec{x};-i\beta,\vec{y}) \end{aligned} \tag{6}$$

Perhaps the least obvious step is the penultimate equality, where we applied time-ordering in presence of imaginary time. Comparing the initial and the final expressions shows periodicity of the finite temperature Green function in Euclidean time with period $\beta = T^{-1}$. Thus, in a quantum field theory the defining signature of a thermal state at temperature T is periodicity in Euclidean time, a conclusion we also reached above. This is also known as KMS condition.

Thus, if you construct a physical state and can show that it has to be periodic in Euclidean time τ with period β, i.e., $\tau \sim \tau + \beta$, you can deduce it is a thermal state at temperature $T = 1/\beta$.

9.2 Hawking temperature from Euclidean regularity

Consider now a D-dimensional spacetime with a non-extremal Killing horizon with surface gravity $\kappa > 0$. As we have shown in the last semester, near the horizon we can universally approximate the spacetime as two-dimensional Rindler spacetime together with some transversal space,

$$\mathrm{d}s^2 = -\kappa^2 r^2 \,\mathrm{d}t^2 + \mathrm{d}r^2 + g_{ij}^{\text{trans}} \,\mathrm{d}x^i \,\mathrm{d}x^j \tag{7}$$

where $i, j = 2, 3, \dots, D$. For instance, for Schwarzschild $g_{ij}^{\text{trans}} \,\mathrm{d}x^i \,\mathrm{d}x^j$ is the metric of the round two-sphere. Continuing (7) to Euclidean signature, $\tau = it$, yields

$$\mathrm{d}s^2 = r^2 \,\mathrm{d}(\kappa\tau)^2 + \mathrm{d}r^2 + \dots \tag{8}$$

where we displayed only the (Euclidean) Rindler part of the metric (see also exercise 9.3). They key observation is that the space defined by the metric (8) locally is just flat Euclidean space in polar coordinates. Globally, however, the metric in general has a conical singularity at $r \to 0$. The only way to avoid this singularity is to make $\kappa\tau$ periodic with period 2π.

We have just derived that regularity of a Killing horizon in Euclidean signature implies Euclidean time is periodic with period $2\pi/\kappa$. Thus, given the considerations of the previous subsection we arrive at an important conclusion. **Spacetimes with a Killing horizon at surface gravity $\kappa > 0$ are thermal states with Hawking–Unruh temperature**

$$\boxed{T = \frac{\kappa}{2\pi}} \tag{9}$$

Note that this conclusion applies to all types of Killing horizons, including event horizons of stationary black holes, cosmological horizons and acceleration horizons.

An important consequence of (9) is that together with the four laws it fixes the numerical factor in the **Bekenstein–Hawking entropy law**

$$\boxed{S_{\text{BH}} = \frac{A_{\text{horizon}}}{4}\,.} \tag{10}$$

9.3 Semi-classical aspects of Hawking radiation

This subsection is again directed towards students familiar with basic aspects of QFT. As in our discussion of black hole perturbations consider a scalar field ϕ on a fixed (black hole) background. Since the Klein–Gordon equation is second order in derivatives we obtained two linearly independent solutions (for each value of the angular quantum number l), so in total the solution was

$$\phi(x) = \sum_i \big(a_i \psi_i(x) + a_i^* \psi_i^*(x)\big) \tag{11}$$

where the sum extends over a complete basis, a_i denote the amplitudes and $\psi_i(x)$ are solutions to the Klein–Gordon equation on a black hole background.

In QFT the amplitudes are replaced by creation and annihilation operators,

$$\phi(x) = \sum_i \big(a_i \psi_i(x) + a_i^\dagger \psi_i^*(x)\big) \tag{12}$$

obeying the Heisenberg algebra (all commutators not displayed vanish)

$$[a_i,\, a_j^\dagger] = \delta_{ij}\,. \tag{13}$$

The QFT Hilbert space is the usual Fock space that starts from a vacuum $|0\rangle$ defined by the conditions

$$a_i|0\rangle = 0 \qquad \forall i \tag{14}$$

together with normalization $\langle 0|0\rangle = 1$. Non-vacuum states in this Fock space are generated by acting on the vacuum with creation operators $a_i^\dagger$.

Let us now choose a different basis of solutions, $\tilde\psi_i$, defined by

$$\tilde\psi_i = \sum_j \left(A_{ij}\psi_j + B_{ij}\psi_j^*\right) \tag{15}$$

subject to the normalization conditions[1]

$$A^\dagger A - B^\dagger B = \mathbb{1} \qquad\qquad AB^T = BA^T\,. \tag{16}$$

The resulting annihilation operators also transform correspondingly,

$$\tilde a_i = \sum_j \left(a_j A_{ji} + a_j^\dagger B_{ji}^*\right) \tag{17}$$

Such a change of basis is known as Bogoliubov-transformation with Bogoliubov coefficients A_{ij} and B_{ij}. Note that for $B_{ij} = 0$ this basis change preserves the vacuum, in the sense that the conditions (14) are identical to the similar conditions with a_i replaced by $\tilde a_i$. However, this is no longer true when $B_{ij} \neq 0$! A consequence of this is that **the original vacuum becomes an excited state with respect to the new basis.**

To show this important statement more explicitly consider the number operator for the $i^{\rm th}$ mode in the original basis,

$$N_i = a_i^\dagger a_i \tag{18}$$

and consider its expectation value in the original vacuum,

$$\langle 0|N_i|0\rangle = \langle 0|a_i^\dagger a_i|0\rangle = 0 \tag{19}$$

which vanishes. Now take instead the number operator for the $i^{\rm th}$ mode in the new basis

$$\tilde N_i = \tilde a_i^\dagger \tilde a_i = \sum_j \left(a_j^\dagger A_{ji}^* + a_j B_{ji}\right)\sum_k \left(a_k A_{ki} + a_k^\dagger B_{ki}^*\right) \tag{20}$$

and consider its expectation value in the original vacuum (in the new vacuum it vanishes by construction),

$$\begin{aligned}\langle 0|\tilde N_i|0\rangle &= \sum_{j,k}\langle 0|a_j B_{ji} a_k^\dagger B_{ki}^*|0\rangle = \sum_{j,k}\langle 0|a_j a_k^\dagger|0\rangle B_{ji}B_{ki}^* \\ &= \sum_{j,k}\langle 0|[a_j,\, a_k^\dagger]|0\rangle B_{ji}B_{ki}^* = \sum_j B_{ji}B_{ij}^\dagger = \left(B^\dagger B\right)_{ii} \neq 0 \end{aligned} \tag{21}$$

Let us now apply Bogoliubov transformations to a scalar field propagating on a black hole background. The key observation is that a mode that has positive frequency at late times (near $\mathscr{I}^+$)

$$\psi_\omega \sim e^{-i\omega(t-r_*)} \tag{22}$$

in general is a mixture of positive and negative frequency modes at early times (near $\mathscr{I}^-$). Similarly, positive frequency modes near $\mathscr{I}^-$ form a mixture of positive and negative frequency modes near $\mathscr{I}^+$. We saw this explicitly when discussing solutions to the Regge–Wheeler equation a few lectures ago. In terms of Bogoliubov

[1] These conditions leave invariant the symplectic inner product $\langle\psi_i,\psi_j\rangle = \delta_{ij} = -\langle\psi_i^*,\psi_j^*\rangle$ and $\langle\psi_i,\psi_j^*\rangle = 0 = \langle\psi_i^*,\psi_j\rangle$.

coefficients it can be shown that the map between the two vacua at $\mathscr{I}^\pm$ is given by (see for instance section 7.3 in the black holes lecture notes gr-qc/9707012 or section 8.2 in the textbook "Introduction to Quantum Effects in Gravity" by Mukhanov and Winitzki)

$$B_{\omega,\tilde{\omega}} = e^{-\pi\omega/\kappa}\, A_{\omega,\tilde{\omega}} \tag{23}$$

where κ is surface gravity of the black hole horizon.

Inserting the result (23) into the left Bogoliubov relation (16) yields a chain of equalities,

$$\delta_{\omega,\tilde{\omega}} = \sum_\lambda \big(A_{\omega,\lambda}A^*_{\tilde{\omega},\lambda} - B_{\omega,\lambda}B^*_{\tilde{\omega},\lambda}\big) = \Big(e^{\pi(\omega+\tilde{\omega})/\kappa} - 1\Big)\sum_\lambda B_{\omega,\lambda}B^*_{\tilde{\omega},\lambda} \tag{24}$$

that establishes

$$\big(BB^\dagger\big)_{\omega,\omega} = \frac{1}{e^{2\pi\omega/\kappa}-1}\,. \tag{25}$$

Since the B-coefficients are non-zero we have thus **particle creation by black holes**.

To check what spectrum we obtain we calculate the vacuum expectation value of the number operator in the vacuum near $\mathscr{I}^+$, using (21) and (25).

$$\langle 0_{\mathscr{I}^+}|N_\omega|0_{\mathscr{I}^+}\rangle = \frac{1}{e^{2\pi\omega/\kappa}-1} \tag{26}$$

This is nothing but the **Planck distribution for a black body at the Hawking temperature** (9)**.**

There are alternative semi-classical derivations. A nice one is given in Unruh's paper, *Phys. Rev.* **D14** (1976) 870. Another efficient method is to derive the existence of a vacuum expectation value of the stress tensor from anomalies. Let me sketch here one such derivation that works for black holes that are effectively two-dimensional (including Schwarzschild). Imposing conformal gauge

$$\mathrm{d}s^2 = e^{2\Omega}\, 2\,\mathrm{d}x^+\,\mathrm{d}x^- \tag{27}$$

covariant conservation of the vev of the stress tensor, $\nabla_\mu\langle T^{\mu\nu}\rangle = 0$, viz.

$$\partial_+\langle T_{--}\rangle + \partial_-\langle T_{+-}\rangle - 2(\partial_-\Omega)\langle T_{+-}\rangle = 0 \tag{28}$$

allows to determine the flux component $\langle T_{--}\rangle$ from the trace component $\langle T_{+-}\rangle$ up to an integration constant, since on static backgrounds $\partial_+ = -\partial_- = \partial_z/\sqrt{2} = \xi(r)\partial_r/\sqrt{2}$ (the same remarks and calculations apply to $\langle T_{++}\rangle$, which for brevity we do not display). A straightforward (but for these lecture notes too lengthy) 1-loop calculation yields the trace anomaly $\langle T^\mu_\mu\rangle \propto R$, which leads to $\langle T_{+-}\rangle = \partial_z^2\Omega/(24\pi)$ and establishes

$$\langle T_{--}\rangle = \frac{1}{24\pi}\big(\partial_z^2\Omega - (\partial_z\Omega)^2\big) + t_- \tag{29}$$

where the integration constant t_- is fixed by the regularity requirement $\langle T_{--}\rangle = 0$ at the horizon (otherwise infinite blueshift factors would render the flux component singular at the horizon in global coordinates). By virtue of $\Omega = \frac12\ln\xi$, with ξ being the Killing norm, this constant is fixed as (for the second equality recall exercise 8.2 of Black Holes I)

$$t_- = \frac{(\partial_r\xi)^2\big|_{r=r_{\text{horizon}}}}{96\pi} = \frac{\kappa^2}{24\pi} = \frac{\pi}{6}\Big(\frac{\kappa}{2\pi}\Big)^2 = \frac{\pi}{6}\,T^2 \tag{30}$$

where T is the Hawking temperature (9). The result (30) gives the asymptotic energy flux and is compatible with the two-dimensional version of the Stefan–Boltzmann law. See section 6 in hep-th/0204253 for more on this derivation and on details how to calculate the trace anomaly using heat kernel methods.

10 Action principle and boundary issues

The first variation of the action should vanish on all solutions to the equations of motion allowed by the boundary conditions. Interestingly, this does not happen automatically. In particular, it does not happen for the Einstein–Hilbert action with the most common boundary conditions (asymptotically flat, asymptotically (A)dS). To resolve this issue we need to first understand what the issue is and how it arises. This, in turn necessitates to take a closer look at the variational principle of Einstein gravity in the presence of (actual or asymptotic) boundaries. In order to be able to do so we need to introduce such boundaries, which in turn requires techniques to decompose "bulk quantities" (such as the metric or the Riemann tensor) into "boundary quantities" plus extra stuff. In this section we give these words a precise mathematical meaning, starting with a canonical decomposition of the metric and related quantities.

10.1 Canonical decomposition of the metric

The canonical decomposition of a D-dimensional metric into a $(D-1)$-dimensional metric and a normal vector was already used in our derivation of the Raychaudhuri equations. Such a decomposition is useful in initial value formulations/Hamiltonian formulations of gravity. For our purposes we need a slightly different decomposition, where the normal vector is not time-like (as it would be for Raychaudhuri's equation or the initial value formulation) but rather spacelike. Thus, our primary data are some D-dimensional metric $g_{\mu\nu}$ (often referred to as "bulk metric") and some spacelike normal vector n^μ, normalized to unity, $n^\mu n_\mu = +1$.

With these data we can define a $(D-1)$-dimensional metric (often referred to as "boundary metric", "induced metric" or "first fundamental form"),

$$h_{\mu\nu} := g_{\mu\nu} - n_\mu n_\nu \tag{1}$$

which is still a D-dimensional symmetric tensor, but projects out the normal component,

$$h_{\mu\nu} n^\nu = 0 \qquad\qquad h^\mu_\mu = D - 1\,. \tag{2}$$

It is also useful to define the projected velocity with which the normal vector changes (often referred to as "extrinsic curvature" or "second fundamental form"),

$$K_{\mu\nu} := h^\alpha_\mu h^\beta_\nu \, \nabla_\alpha n_\beta = \frac{1}{2}\left(\mathcal{L}_n h\right)_{\mu\nu} \tag{3}$$

which can be recast as (one half of) the Lie-variation of the boundary metric along the normal vector. Note that also extrinsic curvature is a symmetric tensor and has vanishing contraction with the normal vector,

$$K_{\mu\nu} = K_{\nu\mu} \qquad\qquad K_{\mu\nu} n^\mu = 0 \tag{4}$$

We shall also need the contraction (or trace) of extrinsic curvature,

$$K := K^\mu_\mu = \nabla_\mu n^\mu\,. \tag{5}$$

Projection with the boundary metric yields a boundary-covariant derivative

$$\mathcal{D}_\mu := h^\nu_\mu \nabla_\nu \tag{6}$$

that leads to standard (pseudo-)Riemann tensor calculus at the boundary when acting on tensors projected to the boundary.

Note that in a canonical context extrinsic curvature also can be interpreted as velocity of the boundary metric, since in that case $\mathcal{L}_n h \sim \dot{h}$, where dot denotes derivative with respect to time, so that derivative of the Lagrange density with respect to extrinsic curvature yields the canonical momentum density. Beware: in such a context there are also various sign changes as compared to these lecture notes since the normal vector in that case would be normalized to -1 instead of $+1$.

10.2 Boundary action for Dirichlet boundary value problem

Often a Dirichlet boundary value problem is desired where the metric is fixed at the boundary $\partial\mathcal{M}$, while its normal derivative is free to fluctuate,

$$\delta g_{\mu\nu}\big|_{\partial\mathcal{M}} = 0 \qquad\qquad n^\alpha\nabla_\alpha\delta g_{\mu\nu}\big|_{\partial\mathcal{M}} \neq 0\,. \tag{7}$$

We show now that the Einstein–Hilbert action is incompatible with such a boundary value problem.

As we have shown previously [see section 4.3, Eq. (14)], first variation of the Einstein–Hilbert action leads to the Einstein equations in the bulk plus total derivative terms,

$$\delta I_{\text{EH}}\big|_{\text{EOM}} = \frac{1}{16\pi G}\int_{\mathcal{M}} \mathrm{d}^D x\sqrt{-g}\,\nabla^\mu\big(\nabla^\nu\delta g_{\mu\nu} - g^{\alpha\beta}\nabla_\mu\delta g_{\alpha\beta}\big) \tag{8}$$

where the subscript 'EOM' indicates that we drop terms that vanish when the bulk equations of motion hold. Using Stokes theorem the total derivative terms in (8) are converted into boundary terms,

$$\delta I_{\text{EH}}\big|_{\text{EOM}} = \frac{1}{16\pi G}\int_{\partial\mathcal{M}} \mathrm{d}^{D-1} x\sqrt{-h}\,n^\mu\big(\nabla^\nu\delta g_{\mu\nu} - g^{\alpha\beta}\nabla_\mu\delta g_{\alpha\beta}\big)\,. \tag{9}$$

Using $n^\mu\nabla^\nu\,\delta g_{\mu\nu} = n^\mu(h^{\nu\alpha} + n^\nu n^\alpha)\nabla_\alpha\,\delta g_{\mu\nu}$ and $n^\mu h^{\nu\alpha}\nabla_\alpha\,\delta g_{\mu\nu} = n^\mu h^{\nu\alpha}\nabla_\alpha[(h^\gamma_\mu + n_\mu n^\gamma)(h^\beta_\nu + n_\nu n^\beta)\,\delta g_{\beta\gamma}] = -K^{\mu\nu}\,\delta g_{\mu\beta} + K n^\mu n^\nu \delta g_{\mu\nu}$ + total boundary derivative, the result (10) can be reformulated as[1]

$$\delta I_{\text{EH}}\big|_{\text{EOM}} = -\frac{1}{16\pi G}\int_{\partial\mathcal{M}} \mathrm{d}^{D-1} x\sqrt{-h}\,\big(h^{\mu\nu}n^\alpha\nabla_\alpha\delta g_{\mu\nu} + (K^{\mu\nu} - K n^\mu n^\nu)\,\delta g_{\mu\nu}\big)\,. \tag{10}$$

The first term in (9) generically is non-zero for the Dirichlet boundary value problem (7). Thus, the Einstein–Hilbert action is inconsistent with (7).

To resolve this issue we add suitable boundary terms to the bulk action, since they do not affect the bulk equations of motion, but may convert the result for the variation (9) into something compatible with the boundary value problem (7). Specifically, we need a boundary term that preserves diffeomorphisms along the boundary and that is capable of canceling the normal derivative of the fluctuations of the metric in (9). Like in the bulk, we can do a derivative expansion of the boundary action,

$$I_{\partial\mathcal{M}} = \frac{1}{16\pi G}\int_{\partial\mathcal{M}} \mathrm{d}^{D-1} x\sqrt{-h}\,\big(b_0 + b_1\mathcal{R} + b_2 K + \dots\big) \tag{11}$$

where the ellipsis refers to terms with higher derivatives (e.g. $K^{\mu\nu}K_{\mu\nu}$ or $K\mathcal{R}$) and $\mathcal{R}$ is the boundary Ricci scalar (constructed from the boundary metric $h_{\mu\nu}$ and the boundary covariant derivative (6)). It is now easy to see that terms intrinsic to the boundary (like the boundary cosmological constant b_0 or the boundary Einstein–Hilbert term $b_1\mathcal{R}$) will not help us, since they cannot produce normal derivatives $n^\mu\nabla_\mu$. Thus, we set $b_0 = b_1 = 0$, focus on the term $b_2 K$ and vary it. Using the definition (5) as well as $\delta n_\mu = \frac{1}{2}\,n_\mu n^\alpha n^\beta\,\delta g_{\alpha\beta}$ yields

$$\delta K = \frac{1}{2}\,h^{\mu\nu}n^\alpha\nabla_\alpha\delta g_{\mu\nu} - \frac{1}{2}\,K n^\mu n^\nu\,\delta g_{\mu\nu} + \text{total boundary derivative} \tag{12}$$

Comparing with the variation (10) we deduce that we should choose $b_2 = 2$ to get consistency with the Dirichlet conditions (7).

[1] We assume here that the boundary $\partial\mathcal{M}$ has no boundary; if this assumption is relaxed the total derivative term is converted into a 'corner' contribution $1/(16\pi G)\int_{\partial^2\mathcal{M}}\mathrm{d}^{D-2}x\sqrt{|\sigma|}\,n^\mu_\sigma n^\nu\,\delta g_{\mu\nu}$, where n^μ_σ is the outward pointing unit normal of the corner.

The full action for Einstein gravity (at this stage of our discussion) compatible with a Dirichlet boundary value problem (7) thus consists of the bulk action I_{EH} plus a boundary action I_{GHY}, known as Gibbons–Hawking–York boundary term.

$$\boxed{I = I_{\text{EH}} + I_{\text{GHY}} = \frac{1}{16\pi G}\int_{\mathcal{M}} \mathrm{d}^D x\sqrt{-g}\big(R - 2\Lambda\big) + \frac{1}{8\pi G}\int_{\partial\mathcal{M}} \mathrm{d}^{D-1}x\sqrt{-h}\,K} \tag{13}$$

Its first variation (assuming a smooth boundary) is given by

$$\delta I = -\frac{1}{16\pi G}\int_{\mathcal{M}} \mathrm{d}^D x\sqrt{-g}\,\big(R^{\mu\nu} - \frac{1}{2}\,g^{\mu\nu}R + \Lambda g^{\mu\nu}\big)\,\delta g_{\mu\nu} \\ - \frac{1}{16\pi G}\int_{\partial\mathcal{M}} \mathrm{d}^{D-1}x\sqrt{-h}\,\big(K^{\mu\nu} - h^{\mu\nu}K\big)\,\delta g_{\mu\nu} \tag{14}$$

The tensor multiplying the variation $\delta g_{\mu\nu}$ at the boundary is known as **Brown–York stress tensor**,

$$T^{\mu\nu}_{\text{BY}} := \frac{1}{8\pi G}\big(K^{\mu\nu} - h^{\mu\nu}K\big)\,. \tag{15}$$

It is important to realize that further boundary terms can be added to the action (13) without spoiling the Dirichlet boundary value problem (7), for instance by choosing $b_0 \neq 0$ or $b_1 \neq 0$ in (11). As we shall see later in these lecture notes these terms are actually necessary in many applications. The reason for this is that even though we have a well-defined Dirichlet boundary value problem we still may not have a well-defined action principle, in the sense that there could be allowed variations of the metric that do not lead to a vanishing first variation (14) on some solutions of the equations of motion. We show now an example for this.

10.3 Action principle in mechanics

Before dealing in the next section with Einstein gravity we consider a much simpler example where the same boundary issues can arise, namely a classical field theory in 0+1 dimensions, also known as mechanics.

Consider specifically the conformal mechanics Hamiltonian

$$H(q,\,p) = \frac{p^2}{2} + \frac{1}{q^2} \tag{16}$$

in the bulk action (chosen on purpose with a $-q\dot{p}$-term to make it more similar to Einstein–Hilbert)

$$I_{\text{bulk}} = \int\limits_0^{t_c} \mathrm{d}t\,\big(-q\dot{p} - H(q,\,p)\big) \tag{17}$$

and a Dirichlet boundary problem, $q(0) = q_0$, $q(t_c) = q_c$. The first variation of the action (17) leads to a boundary term $-q\delta p$, so we introduce a mechanics version of the Gibbons–Hawking–York boundary term

$$I_{\text{GHY}} = qp\big|_0^{t_c}\,. \tag{18}$$

The variation of the full action $I = I_{\text{bulk}} + I_{\text{GHY}}$ yields

$$\delta I = \int\limits_0^{t_c} \mathrm{d}t\,\Big[\Big(-\dot{p} - \frac{H(q,\,p)}{\partial q}\Big)\,\delta q + \Big(\dot{q} - \frac{H(q,\,p)}{\partial p}\Big)\,\delta p\Big] + p\,\delta q\big|_{t=t_c} - p\,\delta q\big|_{t=0}\,. \tag{19}$$

Assuming the initial value q_0 is finite we have $\delta q\big|_{t=0} = 0$ and the last term drops. The bulk terms yield the (Hamilton) equations of motion. Thus, the first variation of the action (19) vanishes on-shell if it were true that $p\,\delta q\big|_{t=t_c} = 0$. For finite t_c and vanishing δq this is obviously the case, but we are interested in the limit $t_c \to \infty$ to mimic typical gravity systems where the range of the coordinates is non-compact.

Now comes the key observation: if we consider $t_c \to \infty$ the correct boundary value is $q_c \to \infty$ (if you look at the form of the potential in (16) you can see this — a ball in that potential just rolls all the way to infinity given infinite amount of time). Thus, finite variations,

$$\lim_{t_c\to\infty} \delta q|_{t=t_x} = \mathcal{O}(1) \tag{20}$$

preserve the asymptotic boundary condition that q_c tends to infinity. But if we allow such variations then **the action I does not have a well-defined variational principle since (19) does not vanish for all variations that preserve our boundary conditions.**

The resolution of this profound problem is to add another boundary term to the action (or to "holographically renormalize it") that does not spoil our Dirichlet boundary value problem. The most general such action is given by

$$\Gamma = \lim_{t_c\to\infty} \big(I_{\text{bulk}} + I_{\text{GHY}} - S(q,\,t)\big|^{t_c}\big) \tag{21}$$

where the counterterm $S(q,\,t)$ needs to be chosen such that the problem above goes away, i.e., the first variation of the full action Γ,

$$\delta\Gamma\big|_{\text{EOM}} = \lim_{t_c\to\infty} \Big(p - \frac{\partial S}{\partial q}\Big)\,\delta q\Big|^{t_c} \tag{22}$$

has to vanish on-shell for all variations preserving our boundary conditions, including finite variations δq.

Thus, we are looking for some function depending on the boundary values that is on-shell equivalent to the momentum, so that the term in parenthesis vanishes in (22). Actually, classical mechanics provides us with a natural candidate, namely Hamilton's principal function which is a solution to the Hamilton–Jacobi equation,

$$H\Big(q,\,\frac{\partial S}{\partial q}\Big) + \frac{\partial S}{\partial t} = 0\,. \tag{23}$$

For the potential (16) the solution is given by the expansion (if you want to see the exact solution look at (11) in `0711.4115`)

$$S(q,\,t) = \frac{q^2}{2t} + \mathcal{O}(1/t)\,. \tag{24}$$

Solving the equations of motion for large time yields

$$p = \frac{q}{t} + \mathcal{O}(1/t^2)\,. \tag{25}$$

Plugging these asymptotic expansions into the variation (22) establishes

$$\delta\Gamma\big|_{\text{EOM}} = \lim_{t_c\to\infty}\Big(\frac{q_c}{t_c} + \mathcal{O}(1/t_c^2) - \frac{q_c}{t_c} + \mathcal{O}(1/t_c)\Big)\,\delta q\Big|^{t_c} = \lim_{t_c\to\infty}\mathcal{O}(1/t_c)\,\delta q\Big|^{t_c} = 0\,. \tag{26}$$

Thus, the action (21) with (17) and (18) has a well-defined variational principle.

Let us finally address another issue with the unrenormalized action. Evaluating I on-shell yields a result that diverges in the limit $t_c \to \infty$. This is problematic insofar as the on-shell action provides the leading order contribution to the semi-classical partition function, which should not be singular. Fortunately, this problem is solved here automatically once we use the action Γ that has a well-defined variational principle. Indeed, evaluating Γ on-shell shows that the result is always finite, even when the upper boundary tends to infinity, $t_c \to \infty$.

11 Gravity aspects of AdS/CFT

In this final section we consider Einstein gravity with asymptotically AdS boundary conditions, where we encounter the same issues as in the conformal mechanics example in the previous section. The solution will again be the same: the addition of suitable boundary terms, a procedure known as "holographic renormalization". This allows us, among other things, to calculate a renormalized Brown–York stress tensor that remains finite. We shall also address asymptotic symmetries, an important concept in gravity and gauge theories far beyond applications in AdS/CFT, which are the focus of this section. Another interesting application of our holographically renormalized action is the calculation of the free energy that permits us to discuss the Hawking–Page phase transition between black holes in AdS and thermal AdS. Finally, we address higher point boundary correlation functions on the gravity side and provide a first glimpse into the AdS/CFT correspondence.

11.1 Asymptotically AdS boundary conditions

We are interested in spacetimes that asymptote to AdS, but would like to make precise what this means. Recalling the line-element for Poincaré-patch AdS (see the last formula in section 2.3 and also recall the cylindrical shape of the Penrose diagram displayed therein), we can define asymptotically AdS spacetimes as metrics with an asymptotic expansion of the form

$$\mathrm{d}s^2\big|_{\text{aAdS}} = \frac{\ell^2}{z^2}\,\mathrm{d}z^2 + \Big(\frac{\ell^2}{z^2}\,\gamma^{(0)}_{\mu\nu} + \gamma^{(2)}_{\mu\nu} + \dots\Big)\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \tag{1}$$

where ℓ is the AdS radius, $\mu,\nu = 0,1,\dots,D-1$, assuming $D\geq 3$, and the ellipsis refers to terms that vanish when $z\to 0$ is approached. We may further restrict $\gamma^{(0)}_{\mu\nu}=\eta_{\mu\nu}$, as it is the case for Poincaré-patch or global AdS. The quantity $\gamma^{(2)}_{\mu\nu}$ may depend arbitrarily on x^μ, which are often referred to as "boundary coordinates". To fully specify our boundary conditions we have to declare if/how $\gamma^{(0)}$ and $\gamma^{(2)}$ are allowed to vary. We postulate

$$\delta\gamma^{(0)}_{\mu\nu} = 0 \qquad\qquad \delta\gamma^{(2)}_{\mu\nu} = \text{arbitrary}\,. \tag{2}$$

The fixed metric $\gamma^{(0)}_{\mu\nu}$ is often called "boundary metric". We call any metric consistent with the expansion (1) and the boundary conditions (2) "locally asymptotically AdS" or just "asymptotically AdS".

There are numerous generalizations of (1), (2): we could change the coordinates; we could switch on mixed terms $\mathrm{d}z\,\mathrm{d}x^\mu$; we could consider non-flat boundary metrics; we could relax or alter the conditions (2) by allowing fluctuations of the boundary metric; we could consider terms that are subleading as compared to $\gamma^{(0)}$ but more dominant than $\gamma^{(2)}$; there could be subleading terms logarithmic in z; etc. As always, boundary conditions are a choice, and the precise choice is dictated by the physical questions one would like to address. The boundary conditions above are useful often enough, so we restrict to them. (For a variety of choices in three dimensional Einstein gravity see 1608.01308.)

Since it can be technically simpler to use Gaussian normal coordinates let us recast (1) in this form, using the coordinate transformation $\rho = -\ell\ln z$.

$$\boxed{\mathrm{d}s^2\big|_{\text{aAdS}} = \mathrm{d}\rho^2 + \Big(e^{2\rho/\ell}\,\gamma^{(0)}_{\mu\nu} + \gamma^{(2)}_{\mu\nu} + \dots\Big)\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu} \tag{3}$$

In many applications the asymptotic boundary $\rho\to\infty$ is replaced by a finite cutoff surface $\rho=\rho_c\gg\ell$. It is then of interest to calculate extrinsic curvature at such a boundary. In Gaussian normal coordinates (3) the normal vector has as only non-vanishing component $n_\rho = n^\rho = 1$ and we obtain

$$K_{\mu\nu} = \nabla_\mu n_\nu = -\Gamma^\rho{}_{\mu\nu} = \frac{1}{\ell}\,e^{2\rho_c/\ell}\,\gamma^{(0)}_{\mu\nu} + \dots \qquad\qquad K_{\rho\mu} = 0 = K_{\rho\rho}\,. \tag{4}$$

11.2 Holographic renormalization

Let us reconsider the on-shell action, the Brown–York stress tensor and the variational principle for asymptotically AdS spacetimes, starting from the action (13) in section 10.2 and using the results (3)-(4). Since we do not have matter, the bulk Ricci scalar is constant, $R = 2D/(D-2)\Lambda = -D(D-1)/\ell^2$. The bulk volume form is determined from $\sqrt{-g} = e^{\rho/\ell(D-1)}\sqrt{-\gamma^{(0)}}\,\big(1+\mathcal{O}(e^{-2\rho/\ell})\big)$. The boundary volume form is given by the same expression, but evaluated at the cutoff surface $\rho = \rho_c$, $\sqrt{-h} = e^{\rho_c/\ell(D-1)}\sqrt{-\gamma^{(0)}}\,\big(1+\mathcal{O}(e^{-2\rho_c/\ell})\big)$. Finally, trace of extrinsic curvature (4) evaluates to $K = (D-1)/\ell\,\big(1+\mathcal{O}(e^{-2\rho_c/\ell})\big)$. Plugging these results into the Einstein–Hilbert action with Gibbons–Hawking–York boundary term yields

$$I = \Big(-\frac{D-1}{8\pi G\,\ell^2}\int\limits_{\rho_0}^{\rho_c}\mathrm{d}\rho\, e^{\rho/\ell(D-1)} + \frac{D-1}{8\pi G\,\ell}\,e^{\rho_c/\ell(D-1)}\Big)\int \mathrm{d}x^{D-1}\sqrt{-\gamma^{(0)}} + \dots \tag{5}$$

which in the large ρ_c limit evaluates to something infinite

$$I\big|_{\rho_c\gg 1} = e^{\rho_c/\ell(D-1)}\,\frac{D-2}{8\pi G\,\ell}\int \mathrm{d}x^{D-1}\sqrt{-\gamma^{(0)}}\,\big(1+\dots\big)\,. \tag{6}$$

Similarly, the Brown–York stress tensor [(15) in section 10.2] is infinite as the cutoff tends to infinity,

$$T_{\mu\nu}^{\textrm{BY}} = \frac{1}{8\pi G}\,\big(K_{\mu\nu} - h_{\mu\nu}K\big) = e^{2\rho_c/\ell}\,\frac{2-D}{8\pi G\,\ell}\,\gamma_{\mu\nu}^{(0)} + \mathcal{O}(1)\,. \tag{7}$$

Worst of all, the variational principle is not well-defined for some variations that preserve our boundary conditions (2). Indeed, evaluating the variation (14) in section 10.2 on-shell and setting to zero $\delta\gamma_{\mu\nu}^{(0)}$ yields

$$\delta I\big|_{\textrm{EOM}} = e^{(D-3)\rho_c/\ell}\,\frac{D-2}{16\pi G\,\ell}\int \mathrm{d}x^{D-1}\sqrt{-\gamma^{(0)}}\,\gamma^{\mu\nu}_{(0)}\,\delta\gamma_{\mu\nu}^{(2)} + \dots \neq 0\,. \tag{8}$$

Thus, we recover the same type of problems that we encountered in the simple mechanics model in section 10.3. It is suggestive that the resolution could also be the same: simply add suitable boundary terms that do not violate our Dirichlet boundary value problem. Adding such boundary terms is known as "holographic renormalization" (holographic, since we are adding boundary terms, and renormalization, since we convert infinite quantities like on-shell action and Brown–York stress tensor into finite ones). The full action is the given by

$$\boxed{\Gamma = I_{\textrm{EH}} + I_{\textrm{GHY}} + I_c} \tag{9}$$

where the holographic counterterm is of the form ($\mathcal{R}$ is the boundary Ricci scalar)

$$I_c = \int_{\partial\mathcal{M}} \mathrm{d}^{D-1}x\sqrt{-h}\,\big(c_0 + c_2\,\mathcal{R} + \dots\big) \tag{10}$$

We have again adhered to the principle to write down all possible terms compatible with the symmetries (boundary diffeomorphisms and boundary Lorentz invariance) and displayed explicitly the first two terms in a derivative expansion; depending on the dimension one might need more than these two terms. The coefficients c_0, c_2 etc. are fixed such that all the problems encountered above go away. We could also determine them from solving a Hamilton–Jacobi equation (as in section 10.3), but often it is more efficient to simply start with an Ansatz like (10) and determine the coefficients by direct calculation.

For sake of specifity we consider in the remainder of this section the simplest case more explicitly, namely Einstein gravity with negative cosmological constant in three spacetime dimensions, $D = 3$.

11.3 Renormalized action and boundary stress tensor

Consider the variation of the action (9) with (10).

$$\lim_{\rho_c\to\infty}\delta\Gamma\big|_{\text{EOM}} = \frac{1}{16\pi G\,\ell}\int \mathrm{d}x^2\sqrt{-\gamma^{(0)}}\,\gamma^{\mu\nu}_{(0)}\,\delta\gamma^{(2)}_{\mu\nu} + \int \mathrm{d}x^2\sqrt{-\gamma^{(0)}}\,\big(c_0\,\frac{1}{2}\,\gamma^{\mu\nu}_{(0)}\,\delta\gamma^{(2)}_{\mu\nu}\big) \tag{11}$$

The first line is just copied-and-pasted from (8) for $D=3$ and the second line comes from varying the holographic counterterm (10). Note that for $D=3$ the term proportional to c_2 has vanishing variation, so that only c_0 remains (and possible higher derivative terms, which however are not needed). Choosing $c_0=-1/(8\pi G\,\ell)$ the right hand side of the variation (11) vanishes. Therefore, the holographically renormalized action for Einstein gravity in AdS_3 reads

$$\boxed{\Gamma_{\text{AdS}_3}=\frac{1}{16\pi G}\int_{\mathcal{M}}\mathrm{d}^3x\sqrt{-g}\,\Big(R+\frac{2}{\ell^2}\Big)+\frac{1}{8\pi G}\int_{\partial\mathcal{M}}\mathrm{d}^2x\sqrt{-h}\,\Big(K-\frac{1}{\ell}\Big)}\,. \tag{12}$$

By construction, the holographically renormalized action (12) has a well-defined variational principle, i.e.,

$$\delta\Gamma_{\text{AdS}_3}\big|_{\text{EOM}}=0 \tag{13}$$

for all variations that preserve the boundary conditions (2)-(3).

Interestingly, the addition of the holographic counterterm also solves the other two problems we had. In particular, the on-shell action is finite, $\Gamma_{\text{AdS}_3}\big|_{\text{EOM}}=\mathcal{O}(1)$. The holographically renormalized Brown–York stress tensor

$$T^{\text{BY-ren}}_{\mu\nu}=\frac{1}{8\pi G}\Big(K_{\mu\nu}-h_{\mu\nu}K+h_{\mu\nu}\,\frac{1}{\ell}\Big)=-\frac{1}{8\pi G\,\ell}\,\gamma^{(2)}_{\mu\nu} \tag{14}$$

is also a finite quantity. Note that it is the subleading contribution in the asymptotic expansion (3) that contributes to the boundary stress tensor. Thus, this expansion coefficient of the metric captures state-dependent information.

11.4 Asymptotic symmetries and glimpse of AdS/CFT

Let us consider all asymptotic Killing vectors, by which we mean all vector fields ξ with the property

$$\mathcal{L}_\xi g_{\mu\nu}=\mathcal{O}(\delta g_{\mu\nu}) \tag{15}$$

where the left hand side is the Lie derivative of any metric compatible with the asymptotically AdS_3 boundary conditions and the right hand side is any fluctuation of the metric compatible with the boundary conditions (plus our gauge conditions to Gaussian normal coordinates; the latter condition could be relaxed).

In components this means

$$\rho\rho:\quad \partial_\rho\xi^\rho=0 \tag{16}$$

$$\rho\pm:\quad \partial_\pm\xi^\rho+g_{\pm\pm}\partial_\rho\xi^\pm+g_{+-}\partial_\rho\xi^\mp=0 \tag{17}$$

$$\pm\pm:\quad \xi^\mu\partial_\mu g_{\pm\pm}+2g_{\pm\pm}\partial_\pm\xi^\pm+2g_{+-}\partial_\pm\xi^\mp=\mathcal{O}(1) \tag{18}$$

$$\pm\mp:\quad \xi^\mu\partial_\mu g_{+-}+g_{+-}\big(\partial_+\xi^++\partial_-\xi^-\big)+g_{++}\partial_-\xi^++g_{--}\partial_+\xi^-=\mathcal{O}(1) \tag{19}$$

where we used light-cone coordinates $x^\pm$ for the boundary metric, $\gamma^{(0)}_{+-}=1$, $\gamma^{(0)}_{\pm}=0$. We solve these equations from bottom to top. The last one implies $\xi^\rho=-\ell/2(\partial\xi)$. The one above it yields $\partial_\pm\xi^\mp=\mathcal{O}(e^{-2\rho/\ell})$. The next one determines the subleading terms in $\xi^\pm$. The top one sets all subleading terms in ξ^ρ to zero.

Thus, we end up with the following set of asymptotic Killing vectors:

$$\xi = \varepsilon^+(x^+)\partial_+ + \varepsilon^-(x^-)\partial_- - \frac{\ell}{2}\big(\partial_+\varepsilon^+ + \partial_-\varepsilon^-\big)\partial_\rho + \mathcal{O}\big(e^{-2\rho/\ell}\big) \tag{20}$$

Modulo small diffeomorphisms (coordinate changes that do not affect the boundary metric or the physical state), they are labeled by two functions, $\varepsilon^\pm(x^\pm)$. Since AdS_3 topologically is a cylinder we can introduce Fourier modes for these two functions, $\varepsilon^\pm_n = ie^{inx^\pm}\partial_\pm$, and determine their Lie-bracket algebra.

$$[\varepsilon^\pm_n,\, \varepsilon^\pm_m] = (n-m)\,\varepsilon^\pm_{n+m} \qquad\qquad n, m \in \mathbb{Z} \tag{21}$$

The algebra (21) consists of two copies of the so-called Witt algebra. Note that this algebra is infinite-dimensional. Thus, we have infinitely many asymptotic Killing vectors in AdS_3.

If you are familiar with CFT_2 you have seen already the algebra (21) and its centrally extended version

$$\boxed{[L^\pm_n,\, L^\pm_m] = (n-m)\,L^\pm_{n+m} + \frac{c^\pm}{12}\,n(n^2-1)\,\delta_{n+m,\,0}} \tag{22}$$

which consists of two copies of the Virasoro algebra with central charges $c^\pm$. You may wonder whether or not there is a central extension on the gravity side. To address this question we note that in a CFT_2 the Fourier modes of the stress tensor flux components $T_{\pm\pm}$ are essentially the Virasoro modes $L^\pm_n$ (and the trace component vanishes due to scale symmetry, $T_{\mu\nu} = 0$). The Virasoro algebra (22) implies that the stress-tensor of a CFT_2 transforms anomalously under (anti-)holomorphic coordinate transformations

$$\delta_\varepsilon T = \varepsilon T' + 2T\varepsilon' + \frac{c}{24\pi}\,\varepsilon''' \tag{23}$$

where we suppressed all $\pm$-indices. We check now if we recover the transformation behavior (23) (a.k.a. infinitesimal Schwarzian derivative) on the gravity side.

As we saw in (14) the role of the boundary stress tensor is played by the subleading term $\gamma^{(2)}_{\mu\nu}$. We check now how it transforms under the asymptotic Killing vectors (20). Insertion into (18) yields

$$\delta_{\varepsilon^\pm}\gamma^{(2)}_{\pm\pm} = \varepsilon^\pm\gamma^{(2)}_{\pm\pm}{}' + 2\gamma^{(2)}_{\pm\pm}\varepsilon^\pm{}' + 2e^{2\rho/\ell}\partial_\pm\xi^\mp\,. \tag{24}$$

Comparison with the infinitesimal Schwarzian derivative (23) shows that all terms match on left and right hand sides, except for the last one which we still need to evaluate. For this we need to determine the subleading terms of $\xi^\pm$, which we obtain from (17).

$$\xi^\pm = \varepsilon^\pm(x^\pm) - \frac{\ell^2}{4}\,e^{-2\rho/\ell}\,\partial^2_\mp\varepsilon^\mp(x^\mp) + \mathcal{O}(e^{-4\rho/\ell}) \tag{25}$$

Plugging (25) back into (24) the last term therein indeed becomes a triple derivative, $-\frac{\ell^2}{2}\,\partial^3_\pm\varepsilon^\pm$, just like the last term in the infinitesimal Schwarzian derivative (23). Reading off the value of the central charge requires to take into account the normalization of the holographically renormalized stress tensor (14), establishing

$$\boxed{c^\pm = \frac{3\ell}{2G}\,.} \tag{26}$$

The results (22) with (26) appeared first in seminal work by Brown and Henneaux. The conclusion of their analysis is that gravity in AdS_3 with asymptotically AdS boundary conditions is dual to a CFT_2, in the sense that the physical phase space falls into representations of two copies of the Virasoro algebra (22) with central charges given by (26). In retrospect, this was an important precursor for the AdS/CFT correspondence.

11.5 Black holes in AdS and Hawking–Page phase transition

Black holes in AdS exist in any spacetime dimension greater or equal to two. In three dimensions they are known as BTZ black holes. They are solutions to the classical field equations descending from the action (12). The BTZ metric ($\varphi \sim \varphi + 2\pi$)

$$\mathrm{d}s^2_{\text{BTZ}} = -\frac{(r^2 - r_+^2)(r^2 - r_-^2)}{r^2\ell^2}\, \mathrm{d}t^2 + \frac{r^2\ell^2\, \mathrm{d}r^2}{(r^2 - r_+^2)(r^2 - r_-^2)} + r^2\left(\mathrm{d}\varphi - \frac{r_+r_-}{r^2\ell}\, \mathrm{d}t\right)^2 \tag{27}$$

has two Killing horizons at $r = r_\pm$. The outer one, $r = r_+ \geq r_-$, is an event horizon. Note that the metric (27) is not only asymptotically AdS_3 [as exercise you can bring it into Fefferman–Graham form (3)], but also locally AdS_3. Thus, the metric (27) differs from the vacuum solution (global AdS_3; in the above coordinates $r_+^2 = -1$ and $r_- = 0$) only by global properties but is locally indistinguishable from it. BTZ is an orbifold of AdS_3; this means it is a quotient space of global AdS_3 by a subgroup of its isometries. For more details on BTZ see gr-qc/9302012. Mass M, angular momentum J, temperature T and angular velocity Ω are given by

$$M = \frac{r_+^2 + r_-^2}{8G\ell^2} \qquad J = \frac{r_+r_-}{4G\ell} \qquad T = \frac{r_+^2 - r_-^2}{2\pi r_+\ell^2} \qquad \Omega = \frac{r_-}{r_+\ell}\,. \tag{28}$$

The Bekenstein–Hawking entropy S is compatible with the first law and a Smarr-like relation,

$$S = \frac{2\pi r_+}{4G} \qquad \mathrm{d}M = T\, \mathrm{d}S + \Omega\, \mathrm{d}J \qquad M = \frac{1}{2}\, TS + \Omega J\,. \tag{29}$$

We are interested in thermodynamical stability of black holes. There are two kinds of instabilities that could arise: perturbative instabilities (e.g. from negative specific heat) and non-perturbative instabilities (from instanton tunneling to a saddle point with lower free energy). Both of them can be checked by considering the Euclidean path integral

$$Z = \int \mathcal{D}g\, \exp\big(-\Gamma_{\text{E}}[g_{\text{E}}]\big) \tag{30}$$

where the left hand side is the Euclidean partition function and the right hand side is the path integral (with some suitable measure $\mathcal{D}g_{\text{E}}$) for the Euclidean version of the holographically renormalized action (12). Around each classical saddle point g_c we expand perturbatively, $g_{\text{E}} = g_c + \delta g$, the Euclidean action

$$\Gamma_{\text{E}}[g_{\text{E}}] = \Gamma_{\text{E}}[g_c + \delta g] = \Gamma_{\text{E}}[g_c] + \delta\Gamma[g_c; \delta g] + \frac{1}{2}\, \delta^2\Gamma[g_c; \delta g] + \dots \tag{31}$$

and the path integral (30)

$$Z = \sum_c e^{-\Gamma_c} \times \int \mathcal{D}\delta g\, e^{-\frac{1}{2}\,\delta^2\Gamma[g_c;\delta g]} \times \dots \tag{32}$$

where we defined the Euclidean on-shell action as $\Gamma_c := \Gamma_{\text{E}}[g_c]$ and used the well-defined variational principle (13). The sum in the Euclidean partition function (32) extends over all classical solutions c compatible with some boundary conditions. The leading contribution to the Euclidean partition function is thus given by (minus the exponential of) the Euclidean on-shell action evaluated for the most dominant saddle point (i.e., the saddle point with the lowest valued for the action). The first subleading corrections are captured by the Gaussian path integral, the middle term in (32); higher order corrections are denoted by the ellipsis. Perturbative stability means that the Gaussian integral converges. Non-perturbative stability of a given saddle point means there is no other saddle point with lower action to which that saddle point could tunnel. The latter statement can also be expressed as evaluation of free energy

$$F = -T\, \ln Z = T\,\Gamma_c + \dots \tag{33}$$

and verifying which classical solution leads to the lowest free energy. Let us check this now for (Euclidean) BTZ.

We consider a thermodynamical ensemble where we keep fixed the temperature T and the angular velocity Ω. Thus, in the sum over c in the partition function (32) all solutions to Einstein gravity contribute that have a given value for these two observables. Since each BTZ black hole has a unique set of values for T and Ω (28) different BTZ black holes cannot compete with each other; however, we always have the possibility to consider thermal AdS_3 as second saddle point, since we can put global AdS_3 at any temperature and at any angular velocity by identifying $(t_{\text{E}}, \varphi) \sim (t_{\text{E}}, \varphi + 2\pi) \sim (t_{\text{E}} + \beta, \varphi + \beta\Omega)$, where t_{E} is Euclidean time.

$$\mathrm{d}s^2_{\text{thAdS}} = \mathrm{d}\rho^2 + \cosh^2(\rho/\ell)\ \mathrm{d}t_{\text{E}}^2 + \ell^2 \sinh^2(\rho/\ell)\ \mathrm{d}\varphi^2 \tag{34}$$

While there might be additional saddles competing for the lowest free energy, our main goal here is to verify in which parameter range BTZ black holes are the dominant saddle as compared to thermal AdS. Thus, we need to evaluate the Euclidean on-shell action

$$\Gamma_{\text{E}} = \frac{1}{16\pi G}\Big(\int \mathrm{d}^3x\sqrt{g_{\text{E}}}\,\frac{4}{\ell^2} - 2\int \mathrm{d}^2x\sqrt{\gamma_{\text{E}}}\,\big(K - \frac{1}{\ell}\big)\Big) \tag{35}$$

for both saddles, where we inserted the on-shell relation $R = -6/\ell^2$ and the required relative minus sign as compared to the Lorentzian action (12). Inserting BTZ coordinates (27) and integrating over the whole outside region yields

$$\begin{aligned}\Gamma_{\text{E}}[g_{\text{BTZ}}] &= \lim_{r_c\to\infty}\frac{1}{16\pi G}\left(\int_{r_+}^{r_c}\mathrm{d}r\int_0^\beta \mathrm{d}t_{\text{E}}\int_0^{2\pi}\mathrm{d}\varphi\, r\,\frac{4}{\ell^2} - 2\int_0^\beta \mathrm{d}t_{\text{E}}\int_0^{2\pi}\mathrm{d}\varphi\sqrt{r_c^2 g_{tt}(r_c)}\,\frac{1}{\ell}\right)\\ &= \frac{\beta}{4G\ell^2}\lim_{r_c\to\infty}\left(r_c^2 - r_+^2 - \sqrt{(r_c^2 - r_+^2)(r_c^2 - r_-^2)}\right) = -\beta\,\frac{r_+^2 - r_-^2}{8G\ell^2}\end{aligned} \tag{36}$$

whereas inserting thermal AdS yields (there is no horizon, so we integrate from the center $r = 0$)

$$\begin{aligned}\Gamma_{\text{E}}[g_{\text{thAdS}}] &= \lim_{r_c\to\infty}\frac{1}{16\pi G}\left(\int_0^{r_c}\mathrm{d}r\int_0^\beta \mathrm{d}t_{\text{E}}\int_0^{2\pi}\mathrm{d}\varphi\, r\,\frac{4}{\ell^2} - 2\int_0^\beta \mathrm{d}t_{\text{E}}\int_0^{2\pi}\mathrm{d}\varphi\sqrt{r_c^2 g_{tt}(r_c)}\,\frac{1}{\ell}\right)\\ &= \frac{\beta}{4G\ell^2}\lim_{r_c\to\infty}\left(r_c^2 - r_c\sqrt{r_c^2+1}\right) = -\beta\,\frac{1}{8G\ell^2}\,.\end{aligned} \tag{37}$$

From (33) we obtain the following results for the respective free energies.

$$F_{\text{BTZ}} = -\frac{r_+^2 - r_-^2}{8G\ell^2} = -\frac{1}{2}\,TS = -\frac{\pi^2}{2G\ell^2}\,\frac{T^2}{1-\Omega^2} \qquad\qquad F_{\text{thAdS}} = -\frac{1}{8G\ell^2} \tag{38}$$

Thus, BTZ black holes are thermodynamically stable against tunneling into thermal AdS_3 if temperature is sufficiently large:

$$T \gg 1/S \qquad \leftrightarrow \qquad F_{\text{BTZ}} \ll F_{\text{thAdS}} \tag{39}$$

This means that thermodynamically stable BTZ black holes have to be sufficiently away from extremality.

Perturbative thermodynamical stability is easy to check by calculating all second derivatives of free energy $F_{\text{BTZ}}(T, \Omega)$ and verifying that the Hessian has only negative eigenvalues. Since the determinant of the Hessian is positive,

$$\det\frac{\partial^2 F_{\text{BTZ}}}{\partial(T,\Omega)} = \frac{\pi^4 T^2}{G^2\ell^4(1-\Omega^2)^3} > 0 \qquad \forall T \in (0,\infty),\ \Omega \in (-1,1)$$

it is sufficient to verify that the eigenvalues are real and their sum is negative for all positive temperatures and all angular velocities with absolute value smaller than unity. This is indeed the case. Thus, all BTZ black holes are thermodynamically stable perturbatively.

In higher dimensions the non-perturbative situation is exactly as above, i.e., there is a high temperature phase where black holes are the dominant saddle and a low temperature phase where thermal AdS is the dominant saddle. As discussed at the end of section 8.2, the perturbative situation differs in higher dimensions: while higher temperature black holes remain stable perturbatively, low temperature black holes have negative specific heat and are thus unstable both perturbatively and non-perturbatively. The phase transition between thermal AdS at low temperatures and black holes at high temperatures is known as Hawking–Page phase transition.

In an AdS/CFT context the Hawking–Page phase transition between thermal AdS at low temperature and black holes at high temperature is interpreted as confinement/deconfinement phase transition.

11.6 Correlation functions and the AdS/CFT correspondence

After Maldacena's seminal paper it was spelled out more explicitly how to obtain CFT correlation functions from a gravity calculation, namely by Gubser, Klebanov, Polaykov and by Witten. In general the AdS/CFT duality relates string theory (on $\text{AdS}_5 \times S^5$) with a specific CFT (namely four-dimensional maximally supersymmetric Yang–Mills). However, in a certain limit (large number of colors and strong coupling on the CFT side) the (super-)gravity approximation is sufficient. Thus, in this limit CFT correlation functions can be calculated using classical gravity. We close these lectures with a glimpse on how this is possible and some concrete examples in an $\text{AdS}_3/\text{CFT}_2$ context.

The formal relationship between CFT observables and string- (or gravity-) observables is captured by the proposed relation

$$\boxed{\langle e^{\int j\,\mathcal{O}}\rangle_{\text{CFT}} = Z_{\text{string}}\big[\phi\big|_{z=0} = j\big]} \tag{40}$$

where the left hand side is the generating function of CFT correlation functions for some operator $\mathcal{O}$ sourced by j and the right hand side is the string theory partition function evaluated with boundary conditions for the corresponding bulk field ϕ determined by setting its boundary value ($z = 0$ in the Feffermann–Graham expansion (1)) to the source j. Since discussing the implications and verifications of (40) is far beyond the scope of these lecture (see this review for more details) we focus now on one specific operator $\mathcal{O}$ that exists in any CFT, namely the stress-tensor $T^{\mu\nu}$ and consider only the gravity limit. Moreover, we shall restrict ourselves to $\text{AdS}_3/\text{CFT}_2$.

For this specific observable the statement (40) implies that correlation functions of the CFT stress tensor can be calculated on the gravity side by taking functional derivatives of the on-shell action with respect to the metric, which is the source of the boundary stress tensor. This means that the quantity $\mathcal{O}$ in the AdS/CFT dictionary (40) is the (boundary) stress tensor $T^{\mu\nu}$, j is its source and ϕ is the metric. If we set $j = 0$ in (40) we are calculating the 0-point function (or partition function), which on the right hand side is the partition function of gravity — which is precisely what we calculated in the previous subsection. If we functionally differentiate once with respect to the source and then set it to zero,

$$\frac{\delta}{\delta j_{\mu\nu}(x)}\,\langle e^{\int \mathrm{d}x' j_{\alpha\beta}(x')T^{\alpha\beta}(x')}\rangle_{\text{CFT}}\Big|_{j=0} = \langle T^{\mu\nu}(x)\rangle = T^{\mu\nu}_{\text{BY-ren}} \sim \frac{\delta\Gamma}{\delta\gamma^{(0)}_{\mu\nu}} \tag{41}$$

then we obtain the 1-point function (or vacuum expectation value) of the stress tensor on the CFT side and the first variation of the action with respect to the metric on the gravity side. Note that we have to vary the boundary metric $\gamma^{(0)}$ to obtain the 1-point function; this is clear, since varying $\gamma^{(2)}$ or any subleading component will lead to a vanishing result due to the well-defined variational principle. Thus, the role of the sources j is played by non-normalizable fluctuations of the metric, i.e., by "fluctuations" that violate the asymptotically AdS boundary conditions. In section 11.3 we showed that the response to such a non-normalizable variation is given by the holographically renormalized Brown–York stress tensor (14), which is thus the stress tensor of the dual CFT_2.

The same logic as above applies to higher-n point functions of the stress tensor (or to correlation functions of other gauge invariant operators in the CFT). Thus, calculating, say, the 42^{nd} functional derivative with respect to the metric of the action (12) should yield all (connected) 42-point functions[1]

$$\langle T^{\alpha\beta}(x^1)T^{\gamma\delta}(x^2)\dots T^{\psi\omega}(x^{42})\rangle_{\text{CFT}} \sim \frac{\delta^{42}\Gamma}{\delta g_{\alpha\beta}(x^1)\delta g_{\gamma\delta}(x^2)\dots\delta g_{\psi\omega}(x^{42})}\,. \tag{42}$$

If you read this claim for the first time I hope you are adequately surprised by it!

There is a number of reasons why the proposed relation (40) is of interest:

- conceptually, it is remarkable and a rather concrete implementation of the holographic principle that string theory (or its gravity limit) is equivalent to an ordinary quantum field theory in one dimension lower
- theoretically, AdS/CFT can be used to define quantum gravity in AdS, thus providing tools for quantum gravity calculations and the resolution of semi-classical puzzles such as the information paradox
- pragmatically, AdS/CFT can be employed as technical trick to convert calculations in strongly interacting CFTs (very hard) into calculations in weakly coupled gravity (rather simple); one example is the modeling of non-abelian plasma formation in strongly coupled quantum field theory as toy model for relativistic heavy ion collisions

We conclude these notes for the lectures "Black Holes II" at TU Wien with a short list of review articles and lecture notes for further reading.

References

[1] P. Townsend, "Black holes" *Lecture notes*; gr-qc/9707012.

[2] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri and Y. Oz, "Large N field theories, string theory and gravity," *Phys. Rept.* **323** (2000) 183; hep-th/9905111.

[3] D. Grumiller, W. Kummer and D. V. Vassilevich, "Dilaton gravity in two-dimensions," *Phys. Rept.* **369** (2002) 327; hep-th/0204253.

[4] T. Jacobson, "Introduction to quantum fields in curved space-time and the Hawking effect" *Lecture notes*; gr-qc/0308048

[5] J. McGreevy, "Holographic duality with a view toward many-body physics," *Adv. High Energy Phys.* **2010** (2010) 723105; 0909.0518.

[6] D. Grumiller, R. McNees and J. Salzer, "Black holes and thermodynamics - The first half century," *Fundam. Theor. Phys.* **178** (2015) 27; 1402.5127.

[7] D. Harlow, "Jerusalem Lectures on Black Holes and Quantum Information," *Rev. Mod. Phys.* **88** (2016) 015002; 1409.1231.

[8] T. Hartman, "Quantum Gravity and Black Holes" *Lecture notes*; http://www.hartmanhep.net/topics2015/

[9] R. McNees, "Conventions, Definitions, Identities, and Other Useful Formulae" *Formula collection*; http://jacobi.luc.edu/useful.pdf

Black Holes II, Daniel Grumiller, June 2018

[1] The explicit check that all stress tensor correlation functions from a CFT_2 match with corresponding functional variations of the action with respect to the metric can be found for AdS_3 as well as for its flat space limit in 1507.05620.