Springer Proceedings in Physics 176

Renata Kallosh
Emanuele Orazi *Editors*

Theoretical Frontiers in Black Holes and Cosmology

Theoretical Perspective in High Energy Physics

Springer

Springer Proceedings in Physics

Volume 176

The series Springer Proceedings in Physics, founded in 1984, is devoted to timely reports of state-of-the-art developments in physics and related sciences. Typically based on material presented at conferences, workshops and similar scientific meetings, volumes published in this series will constitute a comprehensive up-to-date source of reference on a field or subfield of relevance in contemporary physics. Proposals must include the following:

- name, place and date of the scientific meeting
- a link to the committees (local organization, international advisors etc.)
- scientific description of the meeting
- list of invited/plenary speakers
- an estimate of the planned proceedings book parameters (number of pages/articles, requested number of bulk copies, submission deadline).

More information about this series at http://www.springer.com/series/361

Renata Kallosh · Emanuele Orazi
Editors

Theoretical Frontiers in Black Holes and Cosmology

Theoretical Perspective in High Energy Physics

Springer

Editors
Renata Kallosh
Department of Physics
Stanford University
Stanford, CA
USA

Emanuele Orazi
Escola de Ciências e Tecnologia
Universidade Federal Rio Grande do Norte
Natal, RN
Brazil

ISSN 0930-8989 ISSN 1867-4941 (electronic)
Springer Proceedings in Physics
ISBN 978-3-319-31351-1 ISBN 978-3-319-31352-8 (eBook)
DOI 10.1007/978-3-319-31352-8

Library of Congress Control Number: 2016937523

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

This volume aims at providing a pedagogical review on recent developments and applications of black hole physics in the context of high energy physics and cosmology. The contributions are based on lectures delivered at the school "Theoretical Frontiers in Black Holes and Cosmology", held at the "International Institute of Physics" (IIP) in Natal, Brazil, in June 2015. The lectures give a panoramic view of mainstream research lines sharing black hole solutions to gravity and supergravity as common denominator. Starting with accessible and introductory concepts, the newcomer to the field will be brought to a level suitable to face cutting-edge research in the various topics considered in this book.

The only prerequisite for the reader is a working knowledge in field theory and group theory, and the knowledge of general relativity and supersymmetry is desirable. The primary audience is intended to be postgraduate students but the well-established techniques presented in this volume forms a useful review for any scientist working in the field. The selection of authors has been based on worldwide recognized contributions on geometric approaches to fundamental problems in the field of black hole physics.

The book is organized as follows: Chapter "Three Lectures on the FGK Formalism and Beyond" introduces the key role of dualities and the attractor mechanism in the context of singular solutions in ungauged supergravities. These concepts are further developed in Chap. "Introductory Lectures on Extended Supergravities and Gaugings", which is a review of the present methods to build up a gauged supergravity. A basic knowledge on how to gauge a supergravity is the necessary ingredient for Chap. "Supersymmetric Black Holes and Attractors in Gauged Supergravity" that deals with the construction of black hole solutions in a gauged supergravity. The relevance of these solutions is due to applications to gauge/gravity duality, where black hole backgrounds in the bulk are used to model finite temperature condensed matter systems on the boundary. In this framework, the asymptotical AdS space, generated by the gauging procedure, provides the right symmetries to describe a conformal system on the boundary. These first three contributions are intended to be a primer for the community of scientists working in the field of gauge/gravity duality that want to embed more complicated bulk backgrounds in the holographic settings. In Chap. "Lectures on Holographic Renormalization", we selected the holographic renormalization among the many topics in gauge/gravity duality, due to the strong overlapping with techniques used to find the scalar flows for black holes backgrounds in supergravity. Chapter "Nonsingular Black Holes in Palatini Extensions of General Relativity" introduces the reader to a different formulation of gravity based on metric-affine spaces. This approach allows to remove the singularity of general relativity giving rise to a wormhole structure. Finally, Chap. "Inflation: Observations and Attractors" is an introduction to inflation both from theoretical and experimental points of view, aimed at describing the role of cosmological attractors for inflationary model building.

We acknowledge the staff at the IIP for the support in organizing the school "Theoretical Frontiers in Black Holes and Cosmology" where these lectures have been delivered.

Stanford Renata Kallosh
Natal Emanuele Orazi

Contents

Contributors

Pedro F. Ramírez Instituto de Física Teórica UAM/CSIC C/ Nicolás Cabrera, Madrid, Spain

Antonio Gallerati Department DISAT, Politecnico di Torino, Torino, Italy

Dietmar Klemm Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano, Milano, Italy

Gonzalo J. Olmo Departamento de Física Teórica and IFIC, Centro Mixto Universidad de Valencia—CSIC, Paterna, Spain; Universidad de Valencia, Valencia, Spain; Departamento de Física, Universidade Federal da Paraíba, João Pessoa, Paraíba, Brazil

Tomás Ortín Instituto de Física Teórica UAM/CSIC C/ Nicolás Cabrera, Madrid, Spain

Ioannis Papadimitriou SISSA and INFN—Sezione di Trieste, Trieste, Italy

Diederik Roest Van Swinderen Institute for Particle Physics and Gravity, University of Groningen, Groningen, The Netherlands

Marco Scalisi Van Swinderen Institute for Particle Physics and Gravity, University of Groningen, Groningen, The Netherlands

Mario Trigiante Department DISAT, Politecnico di Torino, Torino, Italy

Three Lectures on the FGK Formalism and Beyond

Tomás Ortín and Pedro F. Ramírez

Abstract We review the formalism proposed by Ferrara, Gibbons and Kallosh to study charged, static, spherically symmetric black-hole solutions of $d = 4$ supergravity-like theories and its extension to objects of higher worldvolume dimensions in higher spacetime dimensions and the so-called H-FGK formalism based on variables transforming linearly under duality in the effective action. We also review applications of these formalisms to 4- and 5-dimensional supergravity theories.

1 The FGK Formalism for $d = 4$ Black Holes

Many results in black-hole physics[1] have been derived from the study of families of solutions, that is, solutions whose fields depend on a number of independent physical parameters (mass, electric and magnetic charges, angular momentum and moduli). Obtaining these families of solutions requires, typically, a great deal of effort. The FGK formalism [2] that we are going to review in this lecture dramatically simplifies this task for the static case in supergravity-like field theories. But it does much more than that, since it allows us to derive generic results about entire families of solutions without having to find them explicitly. One of these results is the general form of the celebrated *attractor mechanism* [3–6] that controls the behaviour of scalar fields in the near-horizon limit for extremal black holes and leads to the conclusion that their entropy is moduli-independent and a function of quantized charges only, which strongly suggest a microscopic explanation.

[1] Most of the material covered in these lectures, with additional complementary material and references can be found in the recent book [1].

T. Ortín (✉) · P. F. Ramírez
Instituto de Física Teórica UAM/CSIC C/ Nicolás Cabrera, 13–15, C.U.
Cantoblanco, 28049 Madrid, Spain
e-mail: Tomas.Ortin@csic.es

P. F. Ramírez
e-mail: p.f.ramirez@csic.es

R. Kallosh and E. Orazi (eds.), *Theoretical Frontiers in Black Holes and Cosmology*, Springer Proceedings in Physics 176,
DOI 10.1007/978-3-319-31352-8_1

The formalism relies heavily on the control over the global symmetries of the equations of motion (dualities) of the theory under consideration. Gaillard and Zumino showed in [7] that the symmetries that act on the vector fields are necessarily a subgroup of $\mathrm{Sp}(2\bar{n}, \mathbb{R})$ for theories containing $\bar{n}$ Abelian vector fields. We are going to start by reviewing this general result, taking the opportunity to introduce basic concepts and notation.

1.1 Generic Symmetries of 4-Dimensional Field Theories

In this section we are going to investigate which are the most general symmetries of the equations of motion of supergravity-like theories in 4 dimensions. These are theories defined in a curved space with metric $g_{\mu\nu}$, containing $\bar{n}$ the Abelian 1-form fields A^{Λ}[2] with field strengths $F^{\Lambda} = dA^{\Lambda}$ and a number of scalar fields φ^i parametrizing a space with metric $\mathcal{G}_{ij}(\varphi)$. The action contains an Einstein–Hilbert term for the metric and takes the general form

$$S[F, \varphi] = \int d^4x \sqrt{|g|}\, \big\{ R + \mathcal{G}_{ij}\partial_\mu\varphi^i\partial^\mu\varphi^j + 2\Im\mathfrak{m}\mathcal{N}_{\Lambda\Sigma}F^{\Lambda\,\mu\nu}F^{\Sigma}{}_{\mu\nu} - 2\Re\mathfrak{e}\mathcal{N}_{\Lambda\Sigma}F^{\Lambda\,\mu\nu}\star F^{\Sigma}{}_{\mu\nu}\big\}\,. \quad (1)$$

The $\bar{n}\times\bar{n}$ matrices that describe the coupling of the scalar fields to the vector fields are combined into $\mathcal{N}_{\Lambda\Sigma}(\varphi)$, the complex, symmetric, scalar-dependent *period matrix*. Its imaginary part must be negative-definite.

The bosonic sectors of all the 4-dimensional ungauged $N > 1$ supergravities have this form.[3] The addition of a scalar potential to this action will not change our main conclusions.

On top of standard global symmetries, this kind of theories can have the so-called electric-magnetic dualities[4] which do not leave the action invariant at all, but do leave invariant the complete set of equations of motion extended with the Bianchi identities of the vector field strengths. Gaillard and Zumino showed that there is an associated conserved current for each possible electric-magnetic duality, but it is not the standard Noether current and has to be computed in a different way. We will call it Noether-Gaillard-Zumino (NGZ) current.

---

[2] Capital Greek indices $\Lambda, \Sigma, \Delta, \Gamma$ etc. are used to label 4-dimensional vector fields.

[3] Ungauged $\mathcal{N} = 1, d = 4$ can have a scalar potential derived from a superpotential.

[4] Schrödinger was the first to consider electromagnetic duality transformations, which he introduced in the context of the Born-Infeld theory of non-linear electrodynamics [8]. These transformations were studied in curved spacetime in [9] and in the context of supergravity theories in [10, 11] for $\mathcal{N} = 1$ Maxwell-Einstein and pure $\mathcal{N} = 2$ supergravity, respectively. In [12, 13] it was first observed that, in 4 dimensional supergravity theories, electric-magnetic dualities can be extended to $\mathrm{U}(N)$. However, they were not studied in general field theories until the publication of [7] by Gaillard and Zumino, which we are going to review.

Gaillard and Zumino also showed that the largest possible group of symmetries of the equations of motion of a 4-dimensional theory of the kind we are considering is $\mathrm{Sp}(2\bar{n}, \mathbb{R})$ for theories containing $\bar{n}$ Abelian 1-forms A^Λ. The symmetry group of the equations of motion extended with the Bianchi identities of the Abelian 1-form fields

$$dF^\Lambda = 0, \tag{2}$$

will always be a subgroup of $\mathrm{Sp}(2\bar{n}, \mathbb{R})$.[5] In higher dimensions and for higher-rank form fields the group may be different. We will study this generalization in Lecture 2. Here we are going to review the original 4-dimensional result.

Let us start by defining a dual (or "magnetic") vector field strength $G_\Lambda(F, \varphi)$ for each of the fundamental (or "electric") vector field strengths F^Λ:

$$\star G_\Lambda{}^{\mu\nu} \equiv \frac{1}{4\sqrt{|g|}} \frac{\delta S}{\delta F^\Lambda{}_{\mu\nu}}, \quad \Rightarrow \quad G_\Lambda = \Re\mathfrak{e}\mathcal{N}_{\Lambda\Sigma} F^\Sigma + \Im\mathfrak{m}\mathcal{N}_{\Lambda\Sigma} \star F^\Sigma, \tag{3}$$

which implies

$$G_\Lambda{}^+ = \mathcal{N}^*_{\Lambda\Sigma} F^{\Sigma\,+}. \tag{4}$$

Now the Maxwell equations for each fundamental vector A^Λ can be written as Bianchi identity for the dual vector field strength

$$\nabla_\mu \star G_\Lambda{}^{\mu\nu} = 0, \quad \text{or} \quad dG_\Lambda = 0, \tag{5}$$

which can be locally solved by

$$G_\Lambda = dA_\Lambda, \tag{6}$$

for some 1-forms which are the dual (or "magnetic") 1-form fields.

The Bianchi identities (2) and the Maxwell equations (5) can now be combined linearly. To this end it is useful to define $2\bar{n}$-component vectors of the fundamental and dual 2-form field strengths and consider the linear transformations with a real constant matrix S

$$\begin{pmatrix} F' \\ G' \end{pmatrix} = S \begin{pmatrix} F \\ G \end{pmatrix}, \quad \text{with} \quad S \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{7}$$

Some of the transformations included in the general matrix S are conventional rotations between the 2-form fields but other transformations (involving the off-diagonal blocks B and C) are electric-magnetic duality rotations between the fundamental, electric, 2-form field strengths F^Λ and the dual, magnetic, 2-form field strengths G_Λ.

---

[5] Strictly speaking, this is the part of the symmetry group that acts on the vector fields. The symmetry group of a sector of the scalar fields that does not couple to the vector fields is not restricted at all.

The duality transformations (7) cannot be completely arbitrary: they have to respect the defining relation (4): G'_{Λ} is related to F'^{Λ} in the same form, that is

$$G'_{\Lambda}{}^{+} = \mathcal{N}'^{*}_{\Lambda\Sigma}\, F'^{\Sigma\,+}. \tag{8}$$

Using the definition of the transformations and the relation between the untransformed 2-form field strengths we get the condition

$$\left\{(C + D\mathcal{N}^{*}) - \mathcal{N}'^{*}(A + B\mathcal{N}^{*})\right\} F^{+} = 0, \tag{9}$$

which can only be satisfied if the the period matrix transforms as

$$\mathcal{N}' = (C + D\mathcal{N})(A + B\mathcal{N})^{-1}, \tag{10}$$

The transformed period matrix: $\mathcal{N}'$ must be symmetric and its imaginary part must remain negative-definite. The first condition is

$$C^{T}A - (A^{T}D - C^{T}B)\mathcal{N} + \mathcal{N}(D^{T}B)\mathcal{N} - \text{transposed} = 0, \tag{11}$$

and leads to

$$C^{T}A = A^{T}C, \quad B^{T}D = D^{T}B, \quad A^{T}D - C^{T}B = \kappa\mathbb{1}_{\bar{n}\times\bar{n}}, \tag{12}$$

for an arbitrary $\kappa \in \mathbb{R}$. Later on we will see that the invariance of the energy-momentum tensor (required by the duality-invariance of the metric) requires $\kappa = +1$.

The above properties of the matrices A, B, C and D allow us to write the transformation of the imaginary part of the period matrix in this form

$$\Im\mathfrak{m}\mathcal{N}' = \kappa(A^{T} + \mathcal{N}^{\dagger}B^{T})^{-1}\Im\mathfrak{m}\mathcal{N}(A + B\mathcal{N})^{-1}, \tag{13}$$

from which it follows that it will remain negative-definiteness only if $\kappa > 0$. This is consistent with the value $\kappa = +1$ which we have advanced and which we will use from now onwards.

The conclusion is that $S \in \mathrm{Sp}(2\bar{n}, \mathbb{R})$, which we can define as the the group of transformations S that preserve the symplectic *metric* Ω,

$$S^{T}\Omega S = \Omega, \qquad \Omega \equiv \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}. \tag{14}$$

Observe that we have not proven that the whole $\mathrm{Sp}(2\bar{n}, \mathbb{R})$ group leaves invariant the Maxwell and Bianchi identities. There are more conditions that we still have not considered which will restrict the actual symmetry group to a subgroup of $\mathrm{Sp}(2\bar{n}, \mathbb{R})$.

It is convenient to use a manifestly symplectic-covariant notation, introducing symplectic indices $M, N, \ldots$, equivalent to one upper index-lower index pair $\Lambda, \Sigma, \ldots$ to label the components of $2\bar{n}$-dimensional vectors transforming in the fundamental representation of $\mathrm{Sp}(2\bar{n}, \mathbb{R})$. For instance

$$(\mathcal{F}^M) \equiv \begin{pmatrix} F^\Lambda \\ G_\Lambda \end{pmatrix}, \qquad \mathcal{F}'^M = S^M{}_N \mathcal{F}^N. \tag{15}$$

Ω is used to raise and lower symplectic indices according to the following convention

$$\mathcal{F}_N \equiv \Omega_{NM}\mathcal{F}^M, \quad \mathcal{F}^N = \mathcal{F}_M \Omega^{MN}, \quad \Omega^{MN} = -(\Omega^{-1})^{MN} = \Omega_{MN}, \tag{16}$$

so that

$$(\mathcal{F}_M) = (G_\Lambda, -F^\Lambda). \tag{17}$$

Many objects in the theories that we are considering can be written using these symplectic vectors. For instance, the energy-momentum tensor for the 1-form fields is

$$T^{\text{vect}}_{\mu\nu} \equiv 2\frac{\delta S_{\text{vectors}}}{\delta g_{\mu\nu}} = -8\Im\mathfrak{m}\mathcal{N}_{\Lambda\Sigma}\left[F^\Lambda{}_\mu{}^\rho F^\Sigma{}_{\nu\rho} - \tfrac{1}{4}g_{\mu\nu}F^{\Lambda\,\rho\sigma}F^\Sigma{}_{\rho\sigma}\right], \tag{18}$$

where S_{vectors} corresponds to the last two terms in the generic supergravity-like action (1). This tensor can be rewritten in the two equivalent forms

$$T^{\text{vect}}_{\mu\nu} = -4\mathcal{M}_{MN}(\mathcal{N})\mathcal{F}^M{}_\mu{}^\rho \mathcal{F}^N{}_{\nu\rho} = -4\Omega_{MN} \star \mathcal{F}^M{}_\mu{}^\rho \mathcal{F}^N{}_{\nu\rho}, \tag{19}$$

where we have introduced the symmetric $2\bar{n} \times 2\bar{n}$ matrix $\mathcal{M}(\mathcal{N})$, which is defined in terms of the components of the period matrix by

$$(\mathcal{M}_{MN}(\mathcal{N})) \equiv \begin{pmatrix} I_{\Lambda\Sigma} + R_{\Lambda\Gamma}I^{\Gamma\Omega}R_{\Omega\Sigma} & -R_{\Lambda\Gamma}I^{\Gamma\Sigma} \\ -I^{\Lambda\Omega}R_{\Omega\Sigma} & I^{\Lambda\Sigma} \end{pmatrix}, \tag{20}$$

where we are using the shorthand notation

$$I_{\Lambda\Sigma} \equiv \Im\mathfrak{m}\mathcal{N}_{\Lambda\Sigma}, \quad R_{\Lambda\Sigma} \equiv \Re\mathfrak{e}\mathcal{N}_{\Lambda\Sigma}, \quad I_{\Lambda\Omega}I^{\Omega\Sigma} = \delta_\Lambda{}^\Sigma. \tag{21}$$

$\mathcal{M}_{MN}(\mathcal{N})$ itself is a symmetric symplectic matrix, that is, it satisfies

$$\mathcal{M}_{MP}(\mathcal{N})\Omega^{PQ}\mathcal{M}_{QN}(\mathcal{N}) = \Omega_{MN}, \;\Rightarrow\; (\mathcal{M}^{-1}(\mathcal{N}))^{MN} = \Omega^{MP}\mathcal{M}_{PQ}(\mathcal{N})\Omega^{QN}. \tag{22}$$

When we transform the period matrix as in (10), the matrix $\mathcal{M}_{MN}(\mathcal{N})$ transforms according to

$$\mathcal{M}_{MN}(\mathcal{N}') = (S^{-1})^P{}_M \mathcal{M}_{PQ}(\mathcal{N})(S^{-1})^Q{}_N \equiv \mathcal{M}'_{MN}(\mathcal{N}), \tag{23}$$

and, therefore, the energy-momentum tensor will be invariant under duality transformations. Notice that $\mathcal{M}$ will remain a symplectic matrix only if $\kappa = +1$, which is the restriction that identifies the symplectic group mentioned above.

We can also write the constraint (4) in a symplectic-covariant form using $\mathcal{F}^M$, $\mathcal{M}_{MN}(\mathcal{N})$ and the symplectic metric Ω_{MN}:

$$\star\mathcal{F}^M = \Omega^{MN}\mathcal{M}_{NP}(\mathcal{N})\mathcal{F}^P. \tag{24}$$

In the preceding discussion we have derived the transformation rule for the period matrix (10) but we have not yet discussed under which conditions it remains invariant (otherwise, we are not dealing with a symmetry). The invariance of the period matrix does not need to be absolute: it can be invariant up to transformations of the scalar fields. In other words: it is enough to demand that functional form of the period matrix remains the same in terms of transformed scalars φ'^i. Or, yet in another form: it is enough to demand that the linear transformation rule (10) be equivalent to a reparametrization of the scalars. This condition can be expressed in this form:

$$\mathcal{N}'(\varphi) = [C + D\mathcal{N}(\varphi)][A + B\mathcal{N}(\varphi)]^{-1} = \mathcal{N}(\varphi'). \tag{25}$$

Depending on the functional form of the period matrix (which is part of the definition of the theory), this condition will be satisfied for a different subgroup of $\mathrm{Sp}(2\bar{n}, \mathbb{R})$. It is clear that, in general, it will not be possible to satisfy it for the whole symplectic group.

But this is not the whole story: in this discussion we have only dealt with the contribution to the equations of motion of the last two terms in the action, but we are interested in the global symmetries of the complete set of equations of motion plus Bianchi identities. Besides the scalar fields also occur in their own kinetic term. Therefore, if the transformation of the period matrix has to be equivalent to a transformation of the scalars, this transformation must leave that kinetic term invariant. This can only happen if the transformation of the scalars induced by the duality transformations is an isometry of the metric $\mathcal{G}_{ij}(\varphi)$. If we write the infinitesimal transformations in the form

$$\delta\varphi^i = \alpha^A \xi_A^i(\varphi), \tag{26}$$

(the scalar transformations may be non-linear) where α^A are a set of global parameters, then the $\xi_A^i(\varphi)$ must be Killing vectors of the metric. Their Lie algebra

$$[\xi_A, \xi_B] = -f_{AB}{}^C \xi_C, \tag{27}$$

will be the Lie algebra of the duality group of the theory.[6]

---

[6]Up to the scalars which do not occur in the period matrix and, therefore, do not couple to the 1-form fields, whose global symmetry group is not restricted by any of the previous considerations. Examples of this kind of scalars are provided by the scalars in hypermultiplets of $\mathcal{N} = 2, d = 4$ and $d = 5$ supergravity theories.

1.2 The $d = 4$ FGK Formalism

Following Ferrara et al. [2], let us consider the static, spherically symmetric black-hole solutions of the 4-dimensional supergravity-like theory (1). Since there is no scalar potential nor cosmological constant, the black holes that we will be interested in are also asymptotically flat.

In order to find this kind of solutions we must impose the symmetry conditions on the equations of motion. This is usually done by making an *Ansatz* for all the fields of the theory. We do not want to study each theory case by case and, therefore, we will make an *Ansatz* general enough so the solutions of all the theories of the form (1) fit into it.

A somewhat surprising result of [14] is that the metrics of all the single, static, spherically-symmetric, asymptotically-flat black holes of these theories have the general form

$$\begin{aligned} ds^2 &= e^{2U}dt^2 - e^{-2U}\gamma_{\underline{mn}}dx^{\underline{m}}dx^{\underline{n}}, \\ \gamma_{\underline{mn}}dx^{\underline{m}}dx^{\underline{n}} &= \frac{r_0^4}{\sinh^4 r_0\rho}d\rho^2 + \frac{r_0^2}{\sinh^2 r_0\rho}d\Omega^2_{(2)}. \end{aligned} \tag{28}$$

where e^U, which we will call "metric function" (but it is sometimes called "warp factor"), is a function of the radial coordinate ρ which is different for each solution, $d\Omega^2_{(2)}$ is he metric of the round 2-sphere of unit radius

$$d\Omega^2_{(2)} = d\theta^2 + \sin^2\theta d\phi^2, \tag{29}$$

and r_0, the so-called *non-extremality parameter*, a function of the physical parameters of the solution to be determined, measures how far from the *extremal limit* a regular black-hole solution is. The extremal limit can be defined as the limit in which the Hawking temperature T vanishes and, as we are going to show, it corresponds to $r_0 = 0$ if in that limit the black hole horizon remains regular (otherwise it makes no sense to talk about extremal black hole and temperature at all).

The radial coordinate used to write the general *Ansatz* for the metric, ρ, is meant to go to minus infinity on the horizon and vanish at spatial infinity. In other words: the near-horizon limit is $\rho \to -\infty$ and the spatial infinity limit is at $\rho \to 0^-$.

Let us now proceed to prove the above statement. For $r_0 \neq 0$, in the near-horizon limit $\rho \to -\infty$, the 3-dimensional spatial metric $\gamma_{\underline{mn}}$ behaves as

$$\gamma_{\underline{mn}}dx^{\underline{m}}dx^{\underline{n}} \sim r_0^4 e^{4r_0\rho}d\rho^2 + r_0^2 e^{2r_0\rho}d\Omega^2_{(2)}. \tag{30}$$

This implies that only if the metric function behaves as

$$e^U \sim e^{C + r_0\rho}. \tag{31}$$

in the same limit, the full metric can have a regular horizon (g_{tt} vanishing while the coefficient of $d\Omega^2_{(2)}$ remains finite). The behavior of the full metric in that limit is, in that case

$$ds^2 \sim e^{2C+2r_0\rho}[dt^2 - r_0^4 e^{-4C} d\rho^2] - e^{-2C} r_0^2 d\Omega^2_{(2)}, \tag{32}$$

and the Bekenstein–Hawking (BH) entropy S (one quarter of the area of the horizon in our units) will be given by

$$S = \pi e^{-2C} r_0^2. \tag{33}$$

Changing the radial coordinate to $\rho = (e^{2C}\rho/r_0)^2 - C/r_0$ the time-radial part of the metric of the generic non-extremal black hole we are studying always takes the form of a *Rindler metric*

$$e^{2e^{2C}\rho/r_0}[dt^2 - d\rho^2], \tag{34}$$

and the Hawking temperature can be read from it by comparing it with

$$e^{\frac{4\pi\rho}{\beta}}[dt^2 - d\rho^2], \tag{35}$$

where $\beta = 1/T$. Thus,

$$T = \frac{e^{2C}}{2\pi r_0}, \tag{36}$$

which, together with the general value of the BH entropy obtained above, lead to the general relation derived in [15]

$$r_0 = 2ST, \tag{37}$$

which implies what we wanted to show. This formula is a generalization of the Smarr formula for Reissner–Nordström black holes of mass M and electric charge q which is usually written in the form

$$M = 2TS + q\phi^{\mathrm{h}}, \quad \text{with} \quad \phi^h = \frac{q}{M + r_0}, \quad \text{and} \quad r_0^2 = M^2 - q^2, \tag{38}$$

where ϕ^h is the electrostatic potential evaluated on the outer horizon at $r = M + r_0$. The Smarr formula looks like an integrated first law of black-hole thermodynamics and clearly contains a great deal of information about it.

In the extremal limit, the generic black-hole metric becomes

$$ds^2 = e^{2U}dt^2 - e^{-2U}\frac{1}{\rho^2}\left[\frac{1}{\rho^2}d\rho^2 + d\Omega^2_{(2)}\right], \tag{39}$$

which can be cast in a more common form by changing the radial coordinate ρ to $r = -1/\rho$:

$$ds^2 = e^{2U}dt^2 - e^{-2U}\left[dr^2 + r^2 d\Omega^2_{(2)}\right] = e^{2U}dt^2 - e^{-2U}d\mathbf{x}^2. \tag{40}$$

One of the surprising things about the generic black-hole metric (28) is that it only contains a function to be determined by using the equations of motion of the theory, namely e^U while a generic static and spherically-symmetric metrics depend on two different functions. In a sense, the *Ansatz* (28) has already solved the equation for one of them.[7] This will simplify dramatically the equations of motion. To get some intuition about the metric function e^U and the non-extremality parameter r_0, let us see what they look like in the simplest black-hole solutions.

For Schwarzschild black holes

$$e^{-2U} = e^{-2r_0\rho}, \quad r_0 = M, \tag{41}$$

while for the Reissner–Nordström black holes

$$e^{-2U} = \left(\frac{r_+}{2r_0}e^{-r_0\rho} - \frac{r_-}{2r_0}e^{r_0\rho}\right)^2, \quad \text{with} \quad r_0 = \sqrt{M^2 - q^2}, \quad \text{and} \quad r_\pm = M \pm r_0. \tag{42}$$

We must also make compatible *Ansatzë* for the 1-form and scalar fields, which will also be static and spherically symmetric. For the scalars, it is enough to assume that they are functions of ρ only.

For the vector fields the situation is more complicated: the 2-form field strength of a magnetic monopole is spherically symmetric but depends on the angular coordinates. However, its Hodge dual only depends on ρ and so does the dual 1-form field. Our *Ansatz* must make judicious use of both the dual 1-form fields and the electric ones in order to have simple radial dependence. Thus, we are going to assume the time component of each fundamental 1-form, $A^\Lambda{}_t$, is a function of ρ that we call $\psi^\Lambda(\rho)$ and that the time component of each magnetic 1-form field $A_{\Lambda\, t}$ is another function of ρ, that we call $\chi_\Lambda(\rho)$:

$$A^\Lambda{}_t = \psi^\Lambda(\rho), \quad \Rightarrow \quad F^\Lambda{}_{\underline{m}t} = \partial_{\underline{m}}\psi^\Lambda, \qquad A_{\Lambda\, t} = \chi_\Lambda(\rho), \quad \Rightarrow \quad G_{\Lambda\,\underline{m}t} = \partial_{\underline{m}}\chi_\Lambda, \tag{43}$$

where $\partial_{\underline{m}}$ are the partial derivatives with respect to the three spatial Cartesian coordinates x^m to which the metric $\gamma_{\underline{mn}}$ refers. Using the relations

$$\begin{aligned} F^\Lambda &= I^{-1\,\Lambda\Gamma} R_{\Gamma\Sigma} \star F^\Sigma - I^{-1\,\Lambda\Sigma} \star G_\Sigma, \\ G_\Lambda &= (I + RI^{-1}I)_{\Lambda\Sigma} \star F^\Sigma - R_{\Lambda\Gamma} I^{-1\,\Gamma\Sigma} \star G_\Sigma, \end{aligned} \tag{44}$$

The $G_{\Lambda\,\underline{m}t}$ components of the magnetic 2-form field strengths will determine the angular components of the fundamental 2-form field strengths $F^\Lambda{}_{\theta\phi}$ and vice-versa. As a result, the whole 2-form field strengths (both fundamental and dual) will be determined by the functions ψ^Λ and χ_Λ.

Having defined completely our *Ansatz*, it is time to substitute it into the equations of motion. We will first use the metric (28) with an unspecified time-independent

[7]We will see in more detail in Lecture 2 that this is exactly the case.

spatial metric $\gamma_{\underline{mn}}$ allowing for a general spatial dependence for the fields. In other words, we will not assume spherical symmetry in a first stage. We will do it in second stage, specifying the metric $\gamma_{\underline{mn}}$ as done in (28).

Only the time components of the Maxwell equations and Bianchi identities are non-trivial (the spatial components are automatically solved by our *Ansatz*) and they can be written as the following symplectic-covariant differential equations in the 3-dimensional space with metric $\gamma_{\underline{mn}}$:

$$\nabla_{\underline{m}}\left[e^{-2U}\mathcal{M}_{MN}\partial^{\underline{m}}\Psi^{N}\right]=0, \quad \text{where} \quad (\Psi^{M})\equiv\begin{pmatrix}\psi^{\Lambda}\\ \chi_{\Lambda}\end{pmatrix}. \tag{45}$$

We see that the symplectic matrix $\mathcal{M}_{MN}=\mathcal{M}_{MN}(\mathcal{N})$ defined in (20) arises naturally in this problem.

The scalar equations of motion take the 3-dimensional form

$$\nabla_{\underline{m}}(\mathcal{G}_{ij}\partial^{\underline{m}}\varphi^{j})-\tfrac{1}{2}\partial_{i}\mathcal{G}_{jk}\partial_{\underline{m}}\varphi^{j}\partial^{\underline{m}}\varphi^{k}-\tfrac{1}{2}\partial_{i}\left(4e^{-2U}\mathcal{M}_{MN}\right)\partial_{\underline{m}}\Psi^{M}\partial^{\underline{m}}\Psi^{N}=0. \tag{46}$$

As for the Einstein equations, which must be conveniently written using (19)

$$G_{\mu\nu}+\mathcal{G}_{ij}\left[\partial_{\mu}\varphi^{i}\partial_{\nu}\varphi^{j}-\tfrac{1}{2}g_{\mu\nu}\partial_{\rho}\varphi^{i}\partial^{\rho}\varphi^{j}\right]+4\mathcal{M}_{MN}(\mathcal{N})\mathcal{F}^{M}{}_{\mu}{}^{\rho}\mathcal{F}^{N}{}_{\nu\rho}=0, \tag{47}$$

the flat 00, $0m$ and mn components take the form

$$R+2(\partial U)^{2}-4\nabla^{2}U+\mathcal{G}_{ij}\partial_{\underline{m}}\varphi^{i}\partial_{\underline{n}}\varphi^{j}-4e^{-2U}\mathcal{M}_{MN}\partial_{\underline{m}}\Psi^{M}\partial^{\underline{m}}\Psi^{N}=0, \tag{48}$$

$$\partial_{[\underline{m}}\psi^{\Lambda}\partial_{\underline{n}]}\chi_{\Lambda}=0, \tag{49}$$

$$\begin{aligned}G_{mn}+2\left[\partial_{m}U\partial_{n}U-\tfrac{1}{2}\delta_{mn}(\partial U)^{2}\right]+\mathcal{G}_{ij}\left[\partial_{m}\varphi^{i}\partial_{n}\varphi^{j}-\tfrac{1}{2}\delta_{mn}\partial_{q}\varphi^{i}\partial^{q}\varphi^{j}\right]\\ +4e^{-2U}\mathcal{M}_{MN}\left[\partial_{m}\Psi^{M}\partial_{n}\Psi^{N}-\tfrac{1}{2}\delta_{mn}\partial_{q}\Psi^{M}\partial^{q}\Psi^{N}\right]=0.\end{aligned} \tag{50}$$

This completes the first stage of our calculation, but we still have to massage the result to cast it in a more convenient form.

First, we eliminate R from the first of these last equations using the trace of the third and now all the 3-dimensional equations that we have obtained (except for the next to last one, which is a constraint which will be solved by requiring spherical symmetry) are nothing but the equations of a set of scalar fields $(\phi^{A})\equiv(U,\varphi^{i},\Psi^{M})$ coupled to 3-dimensional gravity which can be derived from the effective 3-dimensional action [16]

$$S[\gamma,\phi]=\int d^{3}x\sqrt{\gamma}\left[R(\gamma)+\mathcal{G}_{AB}(\phi)\gamma^{\underline{mn}}\partial_{\underline{m}}\phi^{A}\partial_{\underline{n}}\phi^{B}\right], \tag{51}$$

where we have defined the metric of indefinite signature $\mathcal{G}_{AB}$

$$(\mathcal{G}_{AB}) \equiv \left(\begin{array}{c|c|c} 2 & & \\ \hline & \mathcal{G}_{ij} & \\ \hline & & 4e^{-2U}\mathcal{M}_{MN} \end{array}\right). \tag{52}$$

The constraint (49) has to be added to the equations of motion derived from the effective action.

In the second stage of this calculation we specify the form of $\gamma_{\underline{mn}}$ for which the Ricci tensor has as only non-vanishing component $R_{\rho\rho} = -2r_0^2$ and we restrict the scalar fields to be functions of ρ only. This solves the constraint (49) while the rest of the equations of motion reduce to[8]

$$\frac{d}{d\rho}(\mathcal{G}_{AB}\dot{\phi}^B) - \tfrac{1}{2}\partial_A \mathcal{G}_{BC}\dot{\phi}^B\dot{\phi}^C = 0, \tag{53}$$

$$\mathcal{G}_{AB}\dot{\phi}^A\dot{\phi}^B - 2r_0^2 = 0, \tag{54}$$

where an overdot indicates a (ordinary) derivative with respect to ρ.

The first equation, which is the geodesic equation in the space with metric $\mathcal{G}_{AB}$ parametrized by the scalars ϕ^A can be derived from the effective action

$$S[\phi] = \int d\rho\, \mathcal{G}_{AB}\dot{\phi}^A\dot{\phi}^B, \tag{55}$$

which has the form of the action of a point particle moving in a space with metric $\mathcal{G}_{AB}$ and coordinates ϕ^A, ρ being the particle's proper time.[9]

The second equation is a constraint. The first term is just the "Hamiltonian" of the system, which is conserved because there is no explicit dependence on the evolution parameter ρ. The constraint relates the value of the Hamiltonian to the non-extremality parameter of the black-hole metric.

This almost completes our calculation. We have reduced the problem of finding static, spherically symmetric, asymptotically-flat black-hole solutions of the supergravity-like action (1) to that of finding solutions of a mechanical system and the solutions are just geodesics in a space with metric $\mathcal{G}_{AB}$.

---

[8]Needless to say, we always have to substitute our *Ansatzë* in the equations of motion and *not* in the action as it is sometimes done in certain literature. Sometimes the final result (the equations of motion obtained from that action) is equivalent, but, often, it is not. In this case, it is clearly not equivalent: we get a constraint that cannot be obtained from the action.

[9]Similar actions arise in the search of other types of solutions of our supergravity-like action which depend effectively on only one direction: cosmologies, instantons, domain walls, etc. See, for instance, [17] and references therein.

Often, the metric $\mathcal{G}_{AB}$ is that of a Riemannian symmetric space[10] and there are many group-theoretical methods to find the geodesics. See, for instance, [16–33].

However, even in non-symmetric spaces, there is a subset of equations of this system that can be integrated immediately[11]: $\mathcal{G}_{AB}$ does not depend on the scalars Ψ^M and the corresponding conserved quantities, $\mathcal{Q}^M$[12] used to integrate the corresponding equations

$$\frac{d}{d\rho}(\mathcal{G}_{MN}\dot{\Psi}^N) = 0, \quad \Rightarrow \quad \mathcal{G}_{MN}\dot{\Psi}^N = 4e^{-2U}\mathcal{M}_{MN}\dot{\Psi}^N = \mathcal{Q}_M/\alpha. \tag{56}$$

This relation can be inverted to eliminate $\dot{\Psi}^M$ from the rest of the equations of motion, which, upon the definition of the *black-hole potential* $V_{\text{bh}} = V_{\text{bh}}(\varphi, \mathcal{Q})$[13]

$$- V_{\text{bh}}(\varphi, \mathcal{Q}) \equiv -\tfrac{1}{2}\mathcal{Q}^M \mathcal{M}_{MN}\mathcal{Q}^N. \tag{57}$$

take the final form [2]

$$\ddot{U} + e^{2U} V_{\text{bh}} = 0, \tag{58}$$

$$\frac{d}{d\rho}(\mathcal{G}_{ij}\dot{\varphi}^j) - \tfrac{1}{2}\partial_i \mathcal{G}_{jk}\dot{\varphi}^j\dot{\varphi}^k + e^{2U}\partial_i V_{\text{bh}} = 0, \tag{59}$$

$$\dot{U}^2 + \tfrac{1}{2}\mathcal{G}_{ij}\dot{\varphi}^i\dot{\varphi}^j + e^{2U} V_{\text{bh}} = r_0^2. \tag{60}$$

Yet again, the first two equations can be obtained from an effective action which now takes the form

$$S_{\text{eff}}[U, \varphi^i] = \int d\rho \, \left[\dot{U}^2 + \tfrac{1}{2}\mathcal{G}_{ij}\dot{\varphi}^i\dot{\varphi}^j - e^{2U} V_{\text{bh}}\right]. \tag{61}$$

which we will call FGK effective action. This is our final result: an effective, mechanical, action which, supplemented by a constraint, gives the equations of motion corresponding to the static, spherically symmetric, asymptotically flat black-hole solutions of any theory of the form (1).

---

[10] This is always the case in $\mathcal{N} \geq 3$, $d = 4$ supergravities.

[11] Observe that integrating these equations of motion will break most of the symmetries of action (55) and no longer we will be able to use group-theoretical methods to solve the equations of motion. We will, nevertheless, obtain very powerful results.

[12] These conserved quantities can be identified up to a normalization constant α to be determined, with the electric q_Λ and magnetic p^Λ:$(\mathcal{Q}^M) \equiv \begin{pmatrix} p^\Lambda \\ q_\Lambda \end{pmatrix}$.

[13] From now on we will set the normalization constant $\alpha = 1/2$ for convenience.

This result is so general that it will allow us to study very general properties of the black-hole solutions (specially for the extremal ones, supersymmetric or not) without having to know them explicitly. We do that in the next section.

1.3 FGK Theorems and the Attractor Mechanism

Let us first consider regular extreme black holes, whose metric has the form (39) or (40). In the near-horizon limit $\rho \to -\infty$ of a regular extremal black hole the metric function e^{-2U} must diverge as

$$e^{-2U} \sim \frac{A}{4\pi}\rho^2, \tag{62}$$

where A is the area of the event horizon and, therefore, the metric will always take the form of the metric of a Robinson–Bertotti solution which is that of $AdS_2 \times S^2$, both with radii equal to $\sqrt{A/(4\pi)}$

$$ds^2 \sim \frac{4\pi}{A}\frac{dt^2}{\rho^2} - \frac{A}{4\pi}\frac{d\rho^2}{\rho^2} - \frac{A}{4\pi}d\Omega^2_{(2)}. \tag{63}$$

We are going to assume as in [2] that the the scalar fields are finite on the horizon of a regular black-hole solution and satisfy the near-horizon condition

$$\lim_{\rho\to-\infty} \mathcal{G}_{ij}\dot{\varphi}^i\dot{\varphi}^j e^{2U}\rho^4 = \lim_{\rho\to-\infty} \frac{4\pi}{A}\mathcal{G}_{ij}\dot{\varphi}^i\dot{\varphi}^j\rho^2 \equiv \xi^2 < \infty\,. \tag{64}$$

Multiplying the Hamiltonian constraint (60) by $e^{2U}\rho^4$ and then by $A^2/(4\pi)$ and using the above assumptions we get a bound for the area of the horizon in relation with the value of the black-hole potential on the horizon:

$$A + \frac{A^2}{8\pi}\xi^2 + 4\pi V_{\rm bh}(\varphi_{\rm h}, \mathcal{Q}) = 0, \quad \Rightarrow \quad A \leq -4\pi V_{\rm bh}(\varphi_{\rm h}, \mathcal{Q}). \tag{65}$$

In terms of a new coordinate $\varpi \equiv -\log(-\rho)$, the definition of the parameter ξ becomes

$$\xi^2 = \lim_{\varpi\to-\infty} \frac{4\pi}{A}\mathcal{G}_{ij}\frac{d\varphi^i}{d\varpi}\frac{d\varphi^j}{d\varpi}, \tag{66}$$

but the r.h.s. is nothing but the kinetic term of the scalar fields in the original action. This identity implies[14]

[14] If the limit was any non-vanishing constant (the only possibility if the scalar metric is going to be regular on the horizon) then φ would be linear in ϖ and would diverge on the near-horizon-limit.

$$\lim_{\varpi \to -\infty} \frac{d\varphi^j}{d\varpi} = \lim_{\rho \to -\infty} \rho \frac{d\varphi^j}{d\rho} = 0. \tag{67}$$

We conclude that, as matter of fact, $\xi^2 = 0$, and the above bound for the area is an identity:

$$S/\pi = -V_{\rm bh}(\varphi_{\rm h}, \mathcal{Q}). \tag{68}$$

But, what is the value of the scalars on the horizon $\varphi_{\rm h}$? Let us analyze the near-horizon limit of their equations of motion (59). Multiplying them by ρ^2 and taking into account (62) and (67) we find

$$\lim_{\rho \to -\infty} \rho^2 \ddot{\varphi}^i = -\frac{4\pi}{A} \mathcal{G}^{ij} \partial_j V_{\rm bh}\Big|_{\varphi=\varphi_{\rm h}}, \tag{69}$$

which provides us with the necessary information to expand the scalars as a power series around the horizon

$$\varphi^i \sim \frac{4\pi}{A} \mathcal{G}^{ij} \partial_j V_{\rm bh}\Big|_{\varphi=\varphi_{\rm h}} \log(-\rho) + \alpha\rho + \varphi^i_{\rm h} + \mathcal{O}(1/\rho). \tag{70}$$

We have assumed that the scalars should take a finite value over a regular horizon. Then, the first two coefficients in the above expansion must vanish. That is $\alpha = 0$ and

$$\partial_i V_{\rm bh}|_{\varphi=\varphi_{\rm h}} = 0. \tag{71}$$

The regularity of the horizon in the extremal limit implies that the possible values of the scalars on the horizon (whose popular name is *attractors*) are the critical points of the black-hole potential and these values determine the entropy through (68).

If the attractors $\varphi_{\rm h}$ depend only on the charges, tat is $\varphi_{\rm h}(\mathcal{Q})$, the values of the scalars on the horizon will be entirely independent of the values of the scalars at spatial infinity φ^i_∞ (known as *moduli*). This is the basic *attractor mechanism* [3–6]. In this case it is evident that the entropy will only depend on the quantized charges

$$S/\pi = -V_{\rm bh}(\varphi_{\rm h}(\mathcal{Q}), \mathcal{Q}). \tag{72}$$

However, in general, $V_{\rm bh}$ may have *flat directions* around a given attractor and some of the equations (71) may not be independent. As a result, the attractor depends on the parameters of the flat directions. Since the only independent parameters of an extremal black-hole solution are the charges $\mathcal{Q}^M$ and the moduli φ^i_∞,[15] those parameters must be (functions of) the moduli and $\varphi^i_{\rm h} = \varphi^i_{\rm h}(\mathcal{Q}, \varphi_\infty)$. The values of the scalars on the horizon are not attractors in the standard sense.

Nevertheless, as point out by Sen in [34], even in that case the entropy (the black-hole potential at the attractor) is a function of the quantized charges only.

[15]The mass M depends on these through the equation $r_0 = 0$.

The independence of the BH entropy of extremal black holes on the moduli (the only continuous parameters the solutions depend on) is the most important consequence of the attractor mechanism as it strongly suggests the existence of an interpretation of the entropy based on microscopic state counting.

We can show explicitly that there is at least one extremal black hole for each attractor: the so-called *double extremal black hole* whose scalars are constant for all values of ρ, the constant being equal to the attractor, according to the above theorem.

The metric function of any non-extremal black hole with $\varphi^i_\infty = \varphi^i_{\rm h}$ takes the form

$$e^{-U} = \cosh r_0\rho - M\frac{\sinh r_0\rho}{r_0}, \quad \text{with} \quad r_0^2 = M^2 + V_{\rm bh}(\varphi_{\rm h}, \mathcal{Q}) \geq 0. \tag{73}$$

This is identical to the metric of the Reissner–Nordström black hole (42). The entropy is just

$$S/\pi = (M + r_0)^2, \quad \Rightarrow \quad T = \frac{r_0}{2S} = \frac{r_0}{2\pi(M + r_0)^2}. \tag{74}$$

Taking in the above formulae the extremal limit $r_0 = 0$ we immediately find the double-extremal solutions and their entropies.

On the other hand, in all $\mathcal{N} > 1, d = 4$ supergravities there are supersymmetric black holes whose metric is that of an extremal black hole. This means that the corresponding the black-hole potential of the supergravity theory must admit at least a *supersymmetric attractor*, which is unique.

Let us study the spatial-infinity limit ($\rho \to 0^-$) in the non-extremal case To $\mathcal{O}(\rho^2)$ we must have the following behaviour

$$U \sim M\rho, \quad \varphi^i \sim \varphi^i_\infty + \Sigma^i\rho, \tag{75}$$

where M is the black-hole mass and the constants Σ^i are, by definition, the *scalar charges*. Taking into account the above behaviors, the same limit in (60) gives

$$M^2 + \tfrac{1}{2}\mathcal{G}_{ij}(\varphi_\infty)\Sigma^i\Sigma^j + V_{\rm bh}(\varphi_\infty, \mathcal{Q}) = r_0^2 \geq 0, \tag{76}$$

which can be read as a non-extremality bound.

The scalar charges are not independent quantities characterizing regular black holes, according to the no-hair theorem. Therefore, they must be functions $\Sigma^i = \Sigma^i(\varphi_\infty, \mathcal{Q}, M)$. Knowing them we could turn the above identity into a formula for the non-extremality parameter $r_0^2 = r_0^2(\varphi_\infty, \mathcal{Q}, M)$, but they are not known in general.

For double extremal black holes, $\Sigma^i = 0$ by definition, which leads to the relation

$$M^2_{\text{double extremal}} = -V_{\rm bh}(\varphi_{\rm h}, \mathcal{Q}) = S/\pi, \tag{77}$$

which we could have obtained from the explicit solution above as well.

1.4 The FGK Formalism for $\mathcal{N} = 2$, $d = 4$ Supergravity

Ungauged $\mathcal{N} = 2$, $d = 4$ supergravity theories with n vector multiplets are specially well suited for putting the FGK formalism to use. We are only interested in the bosonic sector, and we will not consider hyperscalars (the scalar in hypermultiplets) because they do not couple to the vector fields and they can only lead to singular solutions because their charges would be independent and would constitute primary hair. They can be consistently truncated in the bosonic action, which takes the form

$$S = \int d^4x \sqrt{|g|} \left[R + 2\mathcal{G}_{ij^*} \partial_\mu Z^i \partial^\mu Z^{*\, j^*} \right.$$
$$\left. + 2\Im\mathfrak{m}\mathcal{N}_{\Lambda\Sigma} F^{\Lambda\,\mu\nu} F^{\Sigma}{}_{\mu\nu} - 2\Re\mathfrak{e}\mathcal{N}_{\Lambda\Sigma} F^{\Lambda\,\mu\nu} \star F^{\Sigma}{}_{\mu\nu} \right], \tag{78}$$

where Z^i $i = 1, \ldots, n$ are the complex scalars in the vector multiplets and $(A^\Lambda) = (A^0, A^i)$ are the vector fields (A^0 belongs to the supergravity multiplet). The metric $\mathcal{G}_{ij^*}$ is a Kähler metric and it is related to the period matrix by a structure called *Special Geometry* (see, for instance [1] and references therein). In Special Geometry, all the scalar functions that appear in the theory (Kähler potential, connection and metric, period matrix etc.) can be derived from the so-called *canonical, covariantly holomorphic symplectic section* $\mathcal{V}^M(Z, Z^*)$ that defines the theory. An alternative characterization of the theory is through the so-called *prepotential*, but, sometimes, it cannot be defined in certain frames.

The action is of the general form of (1), although the scalar fields are complex. The FGK action and the Hamiltonian constraint take the form

$$S[U, Z^i] = \int d\rho \left[\dot{U}^2 + \mathcal{G}_{ij^*} \dot{Z}^i \dot{Z}^{*\, j^*} - e^{2U} V_{\rm bh} \right],$$
$$r_0^2 = \dot{U}^2 + \mathcal{G}_{ij^*} \dot{Z}^i \dot{Z}^{*\, j^*} + e^{2U} V_{\rm bh}. \tag{79}$$

Using the relations of Special Geometry, it can be seen that the black-hole potential can be written in terms of an object called *central charge* $\mathcal{Z}$

$$- V_{\rm bh}(Z, Z^*, \mathcal{Q}) = |\mathcal{Z}|^2 + 4\mathcal{G}^{ij^*} \partial_i |\mathcal{Z}| \partial_{j^*} |\mathcal{Z}|. \tag{80}$$

where the central charge is defined in terms of the charges and of the symplectic section by

$$\mathcal{Z}(Z, Z^*, \mathcal{Q}) \equiv \mathcal{V}_M \mathcal{Q}^M. \tag{81}$$

The supersymmetric black holes (SBHs) of these theories are always extremal and saturate the supersymmetric (or BPS) bound:

$$M = |\mathcal{Z}_\infty|. \tag{82}$$

Combining this bound with the bound (77) that holds for double extremal black holes we get two relations which are valid for all SBHs:

$$S/\pi = |\mathcal{Z}_{\rm h}|^2, \quad \partial_i |\mathcal{Z}||_{\rm h} = 0, \tag{83}$$

where $\mathcal{Z}_{\rm h}$ is only a function of the charges, and which say that the supersymmetric attractors are also critical points of $|\mathcal{Z}|$ and that the entropy is also determined by (the square of) the absolute value of the central charge on the horizon [2].

The special form of the black-hole potential for these theories, (80), allows us to rewrite the action as a sum of non-negative terms, à *la* Bogomol'nyi [35]

$$S[U, Z^i] = \int d\rho \left[\left(\dot{U} \pm e^U |\mathcal{Z}|\right)^2 + \mathcal{G}_{ij^*} \left(\dot{Z}^i \pm 2e^U \partial^i |\mathcal{Z}|\right) \left(\dot{Z}^{*\, j^*} \pm 2e^U \partial^{j^*} |\mathcal{Z}|\right) \right], \tag{84}$$

up to the boundary term $\mp 2e^U |\mathcal{Z}|$. Then, the action is extremized by the configurations that make all these terms vanish and these configurations must solve the equations of motion derived from the action. The terms in the action vanish if

$$\dot{U} = \mp e^U |\mathcal{Z}|, \qquad \dot{Z}^i = \mp 2e^U \partial^i |\mathcal{Z}|. \tag{85}$$

These first-order (*flow*) imply the second-order Euler-Lagrange equations of motion and should be easier to solve or, at least, to analyze, than those. However, one has to take into account that the Hamiltonian constraint, is only implied by these equations for the extremal case $r_0 = 0$.

These BPS equations can also be derived from the condition of unbroken supersymmetry and are clearly associated to the extremal, supersymmetric solutions and attractors.[16] We know that, in general, there are more, non-supersymmetric, extremal solutions. They are associated to generalizations of the central charge sometimes called *fake central charges* and they satisfy similar flow equations. Let us see how this comes about.

1.4.1 Flow Equations

It is a fact that the black-hole potential can be written in the form (80) for other functions of the scalars and charges $W(Z, Z^*, \mathcal{Q})$ different from the central charge and which receive different names in the literature. We will call them *fake central charges* and the black-hole potential reads in terms of them

$$-V_{\rm bh}(Z, Z^*, \mathcal{Q}) = W^2 + 4\mathcal{G}^{ij^*} \partial_i W \partial_{j^*} W. \tag{86}$$

[16] In the near-horizon limit, these equations give the attractor mechanism for the supersymmetric case.

Then we can rewrite the FGK action again *à la Bogomol'nyi* as in (84) with $|\mathcal{Z}|$ replaced by W and, following the same reasoning, we get the flow equations

$$\dot{U} = \mp e^{U} W, \qquad \dot{Z}^i = \mp 2 e^{U} \partial^i W. \tag{87}$$

These fake central charges are associated to extremal non-supersymmetric black-hole solutions of $\mathcal{N} = 2, d = 4$ supergravity theories [36–42].

The flow equations are not easy to solve, but their analysis leads to the following important and general conclusions:

1. The extrema of the fake central charge are always extrema of the black-hole potential
$$\partial_i W|_{Z_\text{h}} = 0 \quad \Rightarrow \partial_i V_\text{bh}|_{Z_\text{h}} = 0. \tag{88}$$

2. The mass and the scalar charges of the solutions are given by the values of the fake central charge and its derivatives at spatial infinity:
$$M = W_\infty, \quad \Sigma^i = - \lim_{\rho \to 0^-} \mathcal{G}^{ij} \partial_j W\,. \tag{89}$$

2 The General FGK Formalism

Black holes are not the only interesting solutions that supergravity-like theories can have, specially in higher dimensions where higher-rank $(p+1)$ differential-form fields that couple to p-branes through Wess–Zumino terms of the form

$$q \int A_{(p+1)\,\mu_1 \ldots \mu_{p+1}} dX^{\mu_1} \wedge \cdots \wedge dX^{\mu_{p+1}}, \tag{90}$$

can occur. It is natural to try to generalize the FGK formalism to handle those cases and try to use the power of the formalism to derive general properties of p-brane solutions in d dimensions. This generalization was worked out in [43] and we will follow it in this second lecture. The plan of this second lecture will be very similar to that of the first lecture: first, we will define the form of the supergravity-like theories we want to work with and study the possible global symmetries. Then, we will define appropriate *Ansatzë* for the different fields and cases and will substitute it into the equations of motion, reducing their number and dimensionality. In the end we will have a number of equations in a single variable most of which can be derived from an effective FGK-like action. Then we will study general properties of the solutions, deriving theorems similar to those studied in the first lecture. We will finish this lecture with the application of the formalism to some simple theories.

Although most higher-dimensional supergravities include potentials of different ranks we will restrict ourselves to the potentials $A^{\Lambda}_{(p+1)}$ of a single rank $(p+1)$ to study charged p-brane solutions. Our action will contain couplings to scalar fields to scalar fields ϕ^i. similar to those of the $d = 4$ $p = 0$ case we studied in the first lecture.

In general the dual fields $A_{\Lambda\,(\tilde{p}+1)}$ ($\tilde{p} \equiv d - p - 4$) have different rank and p-branes cannot couple to them. Thus, we cannot consider magnetic charges in general. For the same reason, terms of the form $F_{(p+2)} \star F_{(p+2)}$ where

$$F_{(p+2)} \equiv dA_{(p+1)}, \tag{91}$$

are the $(p+2)$-form field strengths, make no sense in the action. However, for some values of d and p we can have $p = \tilde{p}$ and p-branes can carry magnetic charges (electric with respect to the dual potentials $A_{\Lambda\,(p+1)}$) and $F_{(p+2)} \star F_{(p+2)}$ terms do make sense.

In order to save time and energy, we will treat all cases simultaneously introducing always magnetic charges and $F_{(p+2)} \star F_{(p+2)}$ terms even when they do not make sense with the convention that we must ignore them except when they do.

Therefore, the generalization of the action in (1) that we are going to study is

$$\begin{aligned}\mathcal{I}[g, A^{\Lambda}_{(p+1)}, \phi^i] = \int d^d x \sqrt{|g|}\,\big\{R &+ \mathcal{G}_{ij}(\phi)\partial_\mu\phi^i\partial^\mu\phi^j \\ &+ 4\tfrac{(-1)^p}{(p+2)!} I_{\Lambda\Sigma}(\phi) F^{\Lambda}_{(p+2)} \cdot F^{\Sigma}_{(p+2)} \\ &+ 4\xi^2 \tfrac{(-1)^p}{(p+2)!} R_{\Lambda\Sigma}(\phi) F^{\Lambda}_{(p+2)} \cdot \star F^{\Sigma}_{(p+2)}\Big\}\,. \end{aligned} \tag{92}$$

where

$$F^{\Lambda}{}_{(p+2)} \equiv dA^{\Lambda}_{(p+1)}, \quad \text{or} \quad F^{\Lambda}_{(p+2)\,\mu_1\ldots\mu_{p+2}} = (p+2)\partial_{[\mu_1|}A^{\Lambda}{}_{(p+1)\,|\mu_2\ldots\mu_{p+2}]}, \tag{93}$$

are the $(p+2)$-form field strengths, and we are using the notation

$$F^{\Lambda}_{(p+2)} \cdot F^{\Sigma}_{(p+2)} \equiv F^{\Lambda}_{(p+2)\,\mu_1\ldots\mu_{p+2}} F^{\Sigma}_{(p+2)}{}^{\mu_1\ldots\mu_{p+2}}. \tag{94}$$

The scalar-dependent matrix $I_{\Lambda\Sigma}(\phi)$ is symmetric and negative-definite and the scalar-dependent matrix $R_{\Lambda\Sigma}$ will have the same symmetry as the $F^{\Lambda}_{(p+2)} \cdot \star F^{\Sigma}_{(p+2)}$ term:

$$R_{\Lambda\Sigma} = -\xi^2 R_{\Sigma\Lambda}, \quad \xi^2 = -(-1)^{d/2} = (-1)^{p+1}. \tag{95}$$

We have added a ξ^2 factor for convenience to the action. The value of ξ ($+1$ or $+i$) will determine the duality group. It is understood that we must set in the results $R_{\Lambda\Sigma} = 0$ whenever $p \neq \tilde{p} = (d-4)/2$.

This is all, but there is still a possibility that we have not discussed: in the special case $d = 4n+2$, $p = \tilde{p}$-branes can also be self- or anti-self-dual (and, yet, real, as different from the $d = 4$, $p = 0$ case) with the $(p+2)$-field strengths satisfying the corresponding constraint. In our framework we can take this into account by electric and magnetic charges up to a sign.

Thus, we can consider all the possible cases at once working with the above action, taking into account the particular properties of given p and d afterwards.

2.1 Duality Rotations in Higher Dimensions and Ranks

The next step will be to study the general dualities of the (d, p) supergravity-like theories defined by actions similar to the $(d = 4, p = 0)$ one in (1), following the same steps as we (following Gaillard and Zumino) took in the first lecture for the $(d = 4, p = 0)$ case.

First, we define the dual (*magnetic*) $(\tilde{p} + 2)$-form field strengths $G_{(\tilde{p}+2)\,\Lambda}$ by

$$G_{(\tilde{p}+2)\,\Lambda} \equiv R_{\Lambda\Sigma} F^{\Sigma}_{(p+2)} + I_{\Lambda\Sigma} \star F^{\Sigma}_{(p+2)}, \tag{96}$$

in terms of which the equations of motion ("Maxwell equations") of the $(p+1)$-form potentials take the form

$$dG_{(\tilde{p}+2)\,\Lambda} = 0, \tag{97}$$

which is identical to that of the Bianchi identities of the electric $(p+2)$-form potentials

$$dF^{\Lambda}_{(p+2)} = 0. \tag{98}$$

We can rotate into each other the last two equations only if $p = \tilde{p}$, but we are going to construct a $2n$ vector with these field strengths anyway with the understanding that only in $p = \tilde{p}$ case we can mix the upper and lower components:

$$\left(\mathcal{F}^M\right) \equiv \begin{pmatrix} F^{\Lambda}_{(p+2)} \\ G_{(\tilde{p}+2)\,\Lambda} \end{pmatrix}. \tag{99}$$

Then, the Maxwell equations (97) and Bianchi identities (98) can be written as

$$d\mathcal{F}^M = 0. \tag{100}$$

These equations will be preserved by non-singular linear transformations[17]

$$\mathcal{F}'^M = S^M{}_N \mathcal{F}^N, \quad S \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{101}$$

After this transformation, the new magnetic field strengths must be given in terms of the transformed electric one by (96) and this is only possible if we also transform

[17] When $p \neq \tilde{p}$ we have to set $B = C = 0$.

the matrices R, I. It is very convenient to express these transformations in terms of the *generalized period matrix*

$$\mathcal{N} \equiv R + \xi I, \tag{102}$$

which will be real or complex in different dimensions, depending on ξ. Using this matrix the relation between the magnetic and electric field strengths takes a form similar to (4)

$$G_{(p+2)\Lambda}{}^{+} = \mathcal{N}^{*}_{\Lambda\Sigma}\, F^{\Sigma\,+}_{(p+2)}, \tag{103}$$

if we define

$$G_{(p+2)\Lambda}{}^{\pm} = \tfrac{1}{2}\left(G_{(p+2)\Lambda} \pm \xi\right), \quad \Rightarrow \quad \star G_{(p+2)\Lambda}{}^{\pm} = \pm\xi^{3} G_{(p+2)\Lambda}{}^{\pm}. \tag{104}$$

The transformations are, then

$$\mathcal{N}' = (C + D\mathcal{N})\,(A + B\mathcal{N})^{-1}\,, \tag{105}$$

and, when $p \neq \tilde{p}$, setting $R = B = C = 0$ we get

$$I' = DIA^{-1}. \tag{106}$$

Next, let us consider the contribution of the $(p+1)$-form potentials to the energy-momentum tensor. It can be written in the following convenient form:

$$T^{A_{(p+1)}}_{\mu\nu} = \frac{4(-1)^{p+1}}{(p+1)(p+1)!}\mathcal{M}_{MN}(\mathcal{N})\mathcal{F}^{M}{}_{\mu}{}^{\rho_1\ldots\rho_{p+1}}\mathcal{F}^{N}{}_{\nu\rho_1\ldots\rho_{p+1}}, \tag{107}$$

where we have introduced the symmetric matrix

$$(\mathcal{M}_{MN}(\mathcal{N})) \equiv \begin{pmatrix} I - \xi^2 R I^{-1} R & \xi^2 R I^{-1} \\ -I^{-1}R & \xi^4 I^{-1} \end{pmatrix}, \tag{108}$$

When $p = \tilde{p}$ all forms have the same rank and the above expression for the energy-momentum tensor is always consistent. When $p \neq \tilde{p}$ $R = 0$ and $\mathcal{M}_{MN}$ is diagonal and only the indices forms of the same rank are contracted in each term and the expression is, with this understanding, consistent as well.

Using this matrix we can express the self-duality constraint of the field strengths (103) in the form

$$\mathcal{M}_{MN}(\mathcal{N})\mathcal{F}^{N} = \xi^2 \Omega_{MN} \star \mathcal{F}^{M}, \quad \text{where} \quad (\Omega_{MN}) \equiv \begin{pmatrix} 0 & \mathbb{I} \\ \xi^2\mathbb{I} & 0 \end{pmatrix}. \tag{109}$$

We will use the matrix Ω_{MN} as a metric to raise and lower indices as in the symplectic case of the first lecture.

Using the self-duality relation we can finally express the energy-momentum tensor in the form in which it will be easier to determine its symmetry group

$$T^{A_{(p+1)}}_{\mu\nu} = \frac{4(-1)^{p+1}\xi^4}{p+1}\Omega_{MN} \star \mathcal{F}^M{}_\mu{}^{\rho_1\cdots\rho_{p+1}}\mathcal{F}^N{}_{\nu\rho_1\cdots\rho_{p+1}}. \tag{110}$$

It is now very easy to see that the only linear transformations of the field strengths $\mathcal{F}^M$ that will leave the energy-momentum tensor are those that leave the matrix Ω_{MN} invariant: for $p = \tilde{p}$ $\mathrm{O}(n,n)$ when $\xi^2 = +1$ and $\mathrm{Sp}(2n,\mathbb{R})$ when $\xi^2 = -1$. For $p \neq \tilde{p}$ there is not constraint and we can have $\mathrm{GL}(n)$ rotating among each other the electric field strengths.

2.2 The Generalized FGK Effective Action

The next step is to make an *Ansatz* adequate to describe single, charged, static, flat,[18] black p-branes solutions of the action (92) in $d = p + \tilde{p} + 4$ dimensions. We will use a transverse radial coordinate ρ such that the event horizon is at $\rho \to \infty$ instead of $-\infty$ since this option presents problems in $d \neq 4$.

An educated *Ansatz* for the metric based on the known solutions (such as the original solutions of [44] or those in the general [1]) is [43, 45][19]

$$ds^2_{(d)} = e^{\frac{2}{p+1}\tilde{U}}\left[W^{\frac{p}{p+1}}dt^2 - W^{-\frac{1}{p+1}}d\mathbf{y}^2_p\right] - e^{-\frac{2}{\tilde{p}+1}\tilde{U}}d\sigma^2_{\tilde{p}+3}, \tag{111}$$

$$d\sigma^2_{\tilde{p}+3} = \left(\frac{\omega/2}{\sinh\left(\frac{\omega}{2}\rho\right)}\right)^{\frac{2}{\tilde{p}+1}}\left[\left(\frac{\omega/2}{\sinh\left(\frac{\omega}{2}\rho\right)}\right)^2\frac{d\rho^2}{(\tilde{p}+1)^2} + d\Omega^2_{(\tilde{p}+2)}\right], \tag{112}$$

where ρ is the radial coordinate in the $(\tilde{p}+3)$-dimensional transverse space, $d\Omega^2_{(\tilde{p}+2)}$ is the metric of the round $(\tilde{p}+2)$-sphere of unit radius, $\mathbf{y}_p = (y^1,\ldots,y^p)$ are the spatial worldvolume coordinates, $\tilde{U}$ and W are two functions of ρ to be found and ω is the *non-extremality parameter*.

This metric reduces to that in (28) for $d = 4$, $p = 0$ with the identification $\omega = -2r_0$ and to the d-dimensional black-hole ($p = 0$) metric used in [45] in which W disappears and $\tilde{U}$ is just the U used there.

On the other hand, this metric has one undetermined function more than we might have expected. We will see that the equation of motion of W can always be integrated, leaving $\tilde{U}$ as the only function to be found by solving the equations of motion.

---

[18] Flat in the spatial directions of its worldvolume where the metric should be Euclidean.

[19] This metric has also been obtained in [46].

The *Ansatz* for the $(p+1)$-form fields is a direct generalization of that of the $d=4$, $p=0$ case:

$$A^{\Lambda}_{(p+1)\,t\underline{y}^1\ldots\underline{y}^p} = \psi^{\Lambda}(\rho),$$
$$A_{(p+1)\,\Lambda\,t\underline{y}^1\ldots\underline{y}^p} = \chi_{\Lambda}(\rho), \tag{113}$$

but now we will ignore the dual $(p+1)$-form fields $A_{(p+1)\,\Lambda}$, setting $\chi_\Lambda = 0$ unless $p = \tilde{p}$.

Finally, we will assume that all the scalars depend only on ρ.

Substituting the Ansatz into the Maxwell equations and Bianchi identities (100) we get

$$\frac{d}{d\rho}\left[e^{-2\tilde{U}}\,\mathcal{M}_{MN}\,\dot{\Psi}^N\right] = 0, \tag{114}$$

where $\left(\Psi^M\right) \equiv \left(\begin{smallmatrix}\psi^\Lambda \\ \chi_\Lambda\end{smallmatrix}\right)$ and the overdots indicate derivation w.r.t. ρ.

These equations can be integrated right away and they give

$$\dot{\Psi}^M = \alpha e^{2\tilde{U}}\mathcal{M}^{MN}\mathcal{Q}_N, \tag{115}$$

where the integration constants $\mathcal{Q}_M$ are the charges with respect to the electric and magnetic potentials and α is a normalization constant. We will replace $\dot{\Psi}^M$ by the above value in all the equations, which never depend on Ψ^M.

Plugging the *Ansatz* into the Einstein equations

$$G_{\mu\nu} + \mathcal{G}_{ij}\left[\partial_\mu\varphi^i\partial_\nu\varphi^j - \tfrac{1}{2}g_{\mu\nu}\partial_\rho\varphi^i\partial^\rho\varphi^j\right] + \frac{4\xi^2}{(p+1)!}\mathcal{M}_{MN}\mathcal{F}^M{}_\mu{}^{\rho_1\ldots\rho_{p+1}}\mathcal{F}^N{}_{\nu\rho_1\ldots\rho_{p+1}} = 0, \tag{116}$$

we get three equations. The first equation is

$$\frac{d^2\ln W}{d\rho^2} = 0, \quad \Rightarrow \quad W = e^{\gamma\rho}, \tag{117}$$

where we have normalized $W = 1$ at spatial infinity. As we advanced, this leaves us with just one function to be determined by solving the remaining, model-dependent, equations of motion.

Before finding these equations it is worth studying the implications of this form of W for the p-brane metric (111), which now takes the form

$$ds^2_{(d)} = e^{\frac{2}{p+1}\tilde{U}}\left[e^{\frac{p}{p+1}\gamma\rho}dt^2 - e^{-\frac{1}{p+1}\gamma\rho}d\mathbf{y}_p^2\right] - e^{-\frac{2}{\tilde{p}+1}\tilde{U}}d\sigma^2_{(\tilde{p}+3)}, \tag{118}$$

with $d\sigma^2_{\tilde{p}+3}$ still given by (112). This metric now depends on two different constants ω and γ while we expect it to depend on just one: the non-extremality parameter ω.

Actually, γ is a function of ω in branes with regular horizon: In the near-horizon ($\rho \to +\infty$) limit, for $\omega \neq 0$ the angular part of $d\sigma^2_{(\tilde{p}+3)}$ behaves as

$$\sim e^{\frac{1}{\tilde{p}+1}\omega\rho}(-\omega)^{\frac{2}{\tilde{p}+1}} d\Omega^2_{(\tilde{p}+2)}, \tag{119}$$

and the angular part of the whole metric will be regular only if, in the same limit, $\tilde{U}$ behaves as

$$\tilde{U} \sim C + (\omega/2)\rho. \tag{120}$$

Requiring the regularity of the worldvolume components of the metric in this limit and using the above behaviour of $\tilde{U}$ we conclude that

$$\gamma = \omega. \tag{121}$$

The constant C determines the *entropy density by unit (world-) volume* $\tilde{S}$: the constant-time sections of the event horizons of the branes described by the metric (118), whose worldvolume is not compact, have the topology $\mathbb{R}^p \times S^{\tilde{p}+2}$ and have an infinite volume. Only the entropy per unit worldvolume is finite.

It is convenient to further normalize it by dividing by the volume of the $S^{\tilde{p}+2}$ of unit radius which we denote by $\omega_{(\tilde{p}+2)}$ (for 4-dimensional black holes $\tilde{S} = S/\pi$) which leads to the definition

$$\tilde{S} \equiv \frac{A_{\mathrm{h}\,(\tilde{p}+2)}}{\omega_{(\tilde{p}+2)}}, \tag{122}$$

where $A_{\mathrm{h}\,(\tilde{p}+2)}$ is the volume of the $(\tilde{p}+2)$-dimensional constant worldvolume sections of the horizon.

Now, using (120) we get

$$\tilde{S} = \left(-e^{-C}\omega\right)^{\frac{\tilde{p}+2}{\tilde{p}+1}}, \quad \Rightarrow \quad e^C = -\omega\tilde{S}^{-\frac{\tilde{p}+1}{\tilde{p}+2}}. \tag{123}$$

From this discussion we conclude that the metric of a black, non-extremal ($\omega \neq 0$) p-brane with regular horizon in d dimensions always has the form

$$\begin{aligned} ds^2_{(d)} &= e^{\frac{2}{\tilde{p}+1}\tilde{U}}\left[e^{\frac{p}{p+1}\omega\rho}dt^2 - e^{-\frac{1}{p+1}\omega\rho}d\mathbf{y}^2_p\right] - e^{-\frac{2}{\tilde{p}+1}\tilde{U}}d\sigma^2_{(\tilde{p}+3)}, \\ d\sigma^2_{\tilde{p}+3} &= \left(\frac{\omega/2}{\sinh\left(\frac{\omega}{2}\rho\right)}\right)^{\frac{2}{\tilde{p}+1}}\left[\left(\frac{\omega/2}{\sinh\left(\frac{\omega}{2}\rho\right)}\right)^2\frac{d\rho^2}{(\tilde{p}+1)^2} + d\Omega^2_{(\tilde{p}+2)}\right]. \end{aligned} \tag{124}$$

and the near-horizon limit of the metric function is

$$e^{\tilde{U}} \sim (-\omega)\tilde{S}^{-\frac{\tilde{p}+1}{\tilde{p}+2}}e^{\frac{\omega}{2}\rho}. \tag{125}$$

Now, the near-horizon limit of the time-radial part of the metric (124) can be recast in the form

$$\sim e^{\frac{2}{p+1}C} \exp\left(-\frac{(\tilde{p}+1)e^{Cc}}{(-\omega)^{\frac{1}{\tilde{p}+1}}}\rho\right)\left[dt^2 - d\rho^2\right], \quad \text{where} \quad c \equiv \frac{d-2}{(p+1)(\tilde{p}+1)}, \tag{126}$$

and comparing it with the Rindler metric (35) we find that the inverse Hawking temperature is

$$\beta = \frac{4\pi(-\omega)^{\frac{1}{\tilde{p}+1}}}{(\tilde{p}+1)e^{Cc}}, \tag{127}$$

and the following generalization of (37) holds

$$(-\omega)^{\frac{1}{\tilde{p}+1}} = \tfrac{4\pi}{\tilde{p}+1} T\, \tilde{S}^{\frac{(d-2)}{(p+1)(\tilde{p}+2)}}, \tag{128}$$

justifying our calling ω the non-extremality parameter for all d and p.

To finish our study of the black p-brane metric, let us compute the tension following [47, 48]. Let us expand the metric around Minkowski's, far from the brane where the field is weak $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ with

$$h_{\mu\nu} = \frac{c_{\mu\nu}}{r^{\tilde{p}+1}}, \tag{129}$$

where $c_{\mu\nu}$ is a constant tensor and r is a radial coordinate such that the angular part of the metric is, asymptotically, $r^2 d\Omega_{\tilde{p}+3}$. Then, the p-brane's energy-momentum tensor t_{ab} (where the indices ab cover the worldvolume directions) is given by

$$t_{ab} = -\frac{\omega_{\tilde{p}+2}}{16\pi G_{\rm N}^{(d)}}\left[(\tilde{p}+1)c_{ab} + \eta_{ab}\eta^{cd}c_{cd}\right], \tag{130}$$

where ${G_N}^{(d)}$ is the d-dimensional Newton constant. The brane tension T_p is just the t_{00} component which, for the above p-brane metric (124), is given in units such that

$$\omega_{\tilde{p}+2}(\tilde{p}+2) = 8\pi G_{\rm N}^{(d)} = 1, \tag{131}$$

($G_{\rm N}^{(4)} = 1$ for $p = \tilde{p} = 0$) by

$$T_p = -\frac{1}{(p+1)(\tilde{p}+2)}\left[(d-2)\tilde{u} + p(\tilde{p}+1)\omega/2\right], \tag{132}$$

where we have defined the constant

$$\tilde{u} \equiv -\left.\dot{\tilde{U}}\right|_{\rho\to 0^+}. \tag{133}$$

For $d = 4$, $p = \tilde{p} = 0$ $T_p = M$, the black-hole mass and the above formula gives

$$M = -\tilde{u}, \Rightarrow e^{\tilde{U}} \sim e^{-M\rho}. \tag{134}$$

Let us go back to the substitution of our *Ansatz* into the Einstein equations of motion. Substituting our results for W and Ψ^M into them we find two equations for $\tilde{U}, \varphi^i$

$$\ddot{\tilde{U}} + e^{2\tilde{U}} V_{\rm bb} = 0, \tag{135}$$

$$\dot{\tilde{U}}^2 + \tfrac{(p+1)(\tilde{p}+1)}{d-2} \mathcal{G}_{ij} \dot{\varphi}^i \dot{\varphi}^j + e^{2\tilde{U}} V_{\rm bb} = (\omega/2)^2, \tag{136}$$

where

$$V_{\rm bb}(\varphi, \mathcal{Q}) \equiv 2\alpha^2 \tfrac{(p+1)(\tilde{p}+1)}{(d-2)} \mathcal{M}_{MN} \mathcal{Q}^M \mathcal{Q}^N, \tag{137}$$

is the *black-brane potential*.

Finally, from the equations of motion of the scalars

$$\nabla^2 \varphi^i + \Gamma_{jk}{}^i \dot{\varphi}^j \dot{\varphi}^k + \frac{2}{(p+2)!} \partial^i (F^\Lambda \star G_\Lambda) = 0, \tag{138}$$

we get the equation

$$\ddot{\varphi}^i + \Gamma_{jk}{}^i \dot{\varphi}^j \dot{\varphi}^k + \tfrac{d-2}{2(\tilde{p}+1)(p+1)} e^{2\tilde{U}} \partial^i V_{\rm bb} = 0. \tag{139}$$

Equations (135) and (139) can be derived from a mechanical effective action

$$S[\tilde{U}, \varphi^i] = \int d\rho \left\{ \dot{\tilde{U}}^2 + \tfrac{(p+1)(\tilde{p}+1)}{d-2} \mathcal{G}_{ij} \dot{\varphi}^i \dot{\varphi}^j - e^{2\tilde{U}} V_{\rm bb} \right\}, \tag{140}$$

and (136) is the Hamiltonian constraint which must be imposed on the solutions of the above effective action.

2.3 FGK Theorems for Static Flat Branes

We will just state the results, since they are obtained in exactly the same way as in the $d = 4$, $p = 0$ case studied in full detail in the first lecture.

The extremal metric ($\omega \longrightarrow 0$) is given by

$$ds^2_{(d)} = e^{\frac{2\tilde{U}}{p+1}} \left[dt^2 - d\mathbf{y}^2_p \right] - \frac{e^{-\frac{2\tilde{U}}{\tilde{p}+1}}}{\rho^{\frac{2}{\tilde{p}+1}}} \left[\frac{1}{\rho^2} \frac{d\rho^2}{(\tilde{p}+1)^2} + d\Omega^2_{(\tilde{p}+2)} \right], \tag{141}$$

and the transverse part can be seen to be the Euclidean metric in $\mathbb{R}^{\tilde{p}+3}$ by making the coordinate change $\rho = 1/r^{\tilde{p}+1}$.

In the near-horizon limit it always takes the form

$$ds^2_{(d)} = \rho^{\frac{-2}{p+1}} \tilde{S}^{-\frac{2(\tilde{p}+1)}{(p+1)(\tilde{p}+2)}} \left[dt^2 - d\mathbf{y}^2_p\right] - \tilde{S}^{\frac{2}{\tilde{p}+2}} \left[\frac{1}{\rho^2}\frac{d\rho^2}{(\tilde{p}+1)^2} + d\Omega^2_{(\tilde{p}+2)}\right], \quad (142)$$

which is the metric of $AdS_{p+2} \times S^{\tilde{p}+2}$ with radii equal to $\tilde{S}^{\frac{1}{\tilde{p}+2}}$.

The regularity of the fields in the near-horizon limit leads to the following relation between the entropy density and black-brane potential

$$\tilde{S} = [-V_{\text{bb}}(\varphi_{\text{h}}, \mathcal{Q})]^{\frac{\tilde{p}+2}{2(\tilde{p}+1)}}\,, \quad (143)$$

and to the conclusion that the attractors ϕ^i_{h} are the critical points of the black-brane potential on the horizon

$$\partial_i V_{\text{bb}}|_{\varphi=\varphi_{\text{h}}} = 0. \quad (144)$$

The attractor mechanism also works in this general context and the entropy density of an extremal black p-brane will only depend on the quantized charges.

The generalization of the bound (76) for non-extremal p-branes is

$$\tilde{u}^2 + \frac{(p+1)(\tilde{p}+1)}{d-2}\mathcal{G}_{ij}(\varphi_\infty)\Sigma^i\Sigma^j + V_{\text{bb}}(\varphi_\infty, \mathcal{Q}) = (\omega/2)^2, \quad (145)$$

where $\tilde{u}$ is not the p-brane tension T_p but is related to it by (132). For uncharged (*Schwarzschild*) branes $\tilde{u} = \omega/2$ and $T_p = -\omega/2$.

2.4 FGK Formalism for the Black Holes of $\mathcal{N} = 1, d = 5$ Theories

The simplest higher-dimensional theory to which we can apply the generalized FGK formalism is that of ungauged $\mathcal{N} = 1, d = 5$ supergravity coupled to n vector supermultiplets.[20] The 1-forms can couple to black holes and their dual 2-form potentials, can couple to black strings. We have to consider both cases separately and we start by the black-hole case.

[20]We ignore the hypermultiplets for exactly the same reasons as in the $\mathcal{N} = 2, d = 4$ case.

The bosonic sector of these theories is controlled by the action

$$S[F,\varphi] = \int d^5x\sqrt{|g|}\left\{R + \tfrac{1}{2}g_{xy}\partial_\mu\phi^x\partial^\mu\phi^y - \tfrac{1}{4}a_{IJ}F^I{}_{\mu\nu}F^{J\mu\nu} \right. \\ \left. + \frac{C_{IJK}\varepsilon^{\mu\nu\rho\sigma\lambda}}{12\sqrt{3}\sqrt{|g|}}F^I{}_{\mu\nu}F^J{}_{\rho\sigma}A^K{}_\lambda\right\}, \tag{146}$$

where ϕ^x $x = 1,\ldots,n$ are the real scalars in the vector multiplets and $(A^I) = (A^0, A^x)$ are the vector fields (A^0 belongs to the supergravity multiplet). The real metric g_{xy} is related to the scalar-dependent kinetic matrix a_{IJ} and to the symmetric, constant tensor C_{IJK} that defines the theory by a structure called *Real Special Geometry* (see, for instance [1] and references therein).

Since our solutions will be either black holes or black strings, always static and non-intersecting, we can safely ignore the last term and, then, we have a theory which is of the general form (92) replacing $p = 0$, $\tilde{p} = 1$, $\mathcal{G}_{ij}$ by $\frac{1}{2}g_{xy}$ and $I_{\Lambda\Sigma}$ by a_{IJ} ($R_{\Lambda\Sigma} = 0$ here and there are no magnetic charges). The effective action is obtained by making the same replacements in (140) and Hamiltonian constraint (136). They take the simple form (writing U instead of $\tilde{U}$)

$$S[U,\phi^x] = \int d\rho\,\left\{\dot{U}^2 + \tfrac{1}{3}g_{xy}\dot{\phi}^x\dot{\phi}^y - e^{2U}V_{\text{bh}}\right\}, \tag{147}$$

$$(\omega/2)^2 = \dot{U}^2 + \tfrac{1}{3}g_{xy}\dot{\phi}^x\dot{\phi}^y + e^{2U}V_{\text{bh}}. \tag{148}$$

In these equations the black-hole potential $V_{\text{bh}}(\phi,q)$ is given, with the normalization $\alpha^2 = 3/32$, by these two equivalent expressions

$$-V_{\text{bh}}(\phi,q) = a^{IJ}q_Iq_J = \mathcal{Z}_{\text{e}}^2 + 3g^{xy}\partial_x\mathcal{Z}_{\text{e}}\partial_y\mathcal{Z}_{\text{e}}, \tag{149}$$

where $\mathcal{Z}_{\text{e}}(\phi,q)$ is the *(electric) black-hole central charge* defined, in its turn, by

$$\mathcal{Z}_{\text{e}}(\phi,q) \equiv h^I(\phi)q_I. \tag{150}$$

The solutions for the metric function e^U must be substituted in the general metric of regular 5-dimensional black holes, which is always of the form

$$ds^2 = e^{2U}dt^2 - e^{-U}\left(\frac{\omega/2}{\sinh\left(\frac{\omega}{2}\rho\right)}\right)\left[\left(\frac{\omega/2}{\sinh\left(\frac{\omega}{2}\rho\right)}\right)^2\frac{d\rho^2}{4} + d\Omega^2_{(3)}\right]. \tag{151}$$

The SBHs of these theories saturate the supersymmetric (BPS) bound

$$M = \mathcal{Z}_{\text{e}}(\phi_\infty,q). \tag{152}$$

Furthermore, the values of the scalars on the horizon are critical points of the black-hole central charge $\mathcal{Z}_{\rm e}$ and of the black-hole potential:

$$\partial_x \mathcal{Z}_{\rm e}|_{\phi_{\rm h}} = 0, \quad \text{and} \quad \partial_x V_{\rm bh}|_{\phi_{\rm h}} = 0. \tag{153}$$

However, as in the $\mathcal{N} = 2, d = 4$ case there are non-supersymmetric attractors which extremize the black-hole potential but not the back-hole central charge.

Equation (147) can be rewritten *à la Bogomol'nyi*:

$$S[U, \phi^x] = \int d\rho \left\{ (\dot{U} \pm e^U \mathcal{Z}_{\rm e})^2 + \tfrac{1}{3} g_{xy} (\dot{\phi}^x \pm 3 e^U \partial^x \mathcal{Z}_{\rm e})(\dot{\phi}^y \pm 3 e^U \partial^y \mathcal{Z}_{\rm e}) \mp \frac{d}{d\rho}(2 e^U \mathcal{Z}_{\rm e}) \right\}, \tag{154}$$

and the standard arguments lead to the flow equations

$$\dot{U} = \mp e^U \mathcal{Z}_{\rm e} \qquad \dot{\phi}^x = \mp 3 e^U \partial^x \mathcal{Z}_{\rm e}, \tag{155}$$

which can be handled as in the $\mathcal{N} = 2, d = 4$ case. In particular, the first equation leads precisely to (152), which characterizes supersymmetric black holes. Non-supersymmetric black holes are associated to fake central charges different from $\mathcal{Z}_{\rm e}$, as in $\mathcal{N} = 2, d = 4$.

There is another way of rewriting the action *à la Bogomol'nyi* using the scalar functions $h^I(\phi)$, which are constrained to satisfy

$$C_{IJK} h^I h^J h^K = 1, \tag{156}$$

a constraint which is precisely solved by the physical scalars. Using simple identities of Real Special Geometry we arrive to

$$S[U, \phi^x] = \int d\rho \left\{ e^{2U} a^{IJ} \left[\frac{d}{d\rho}(e^{-U} h_I) \pm q_I \right] \left[\frac{d}{d\rho}(e^{-U} h_J) \pm q_J \right] \pm \frac{d}{d\rho}(2 e^U \mathcal{Z}_{\rm e}) \right\}, \tag{157}$$

whose associated flow equations are

$$\frac{d}{d\rho}(e^{-U} h_I) = \mp q_I. \tag{158}$$

These can be solved immediately, giving

$$e^{-U} h_I = A_I \mp q_I \rho, \tag{159}$$

for some integration constants A_I. These are harmonic functions in the 4-dimensional spatial, transverse space. It is a well-known result that the static, timelike supersymmetric solutions of these theories can be constructed in terms of harmonic functions (which can have more poles that the ones we have obtained, which must be consistent with spherical symmetry) [49]. We have just recovered this result in a very simple way.

2.5 FGK Formalism for the Black Strings of $\mathcal{N} = 1, d = 5$ Theories

The 1-form fields A^I, can be dualized into 2-form fields, B_I associated to black-string solutions as follows: the Maxwell equations for the 1-forms are (ignoring the contribution of the Chern–Simons term)

$$d\left(a_{IJ} \star F^J\right) = 0, \tag{160}$$

which we can solve locally by

$$a_{IJ} \star F^J = dB_I \equiv H_I, \quad \Rightarrow \quad F^I = a^{IJ} \star H_I, \tag{161}$$

where and a^{IJ} is the inverse of a_{IJ}. The Bianchi identity for the 2-form field strengths becomes the Maxwell equation for the dual 2-forms

$$dF^I = 0, \quad \longrightarrow \quad d\left(a^{IJ} \star H_I\right) = 0, \tag{162}$$

and this equation (and the Einstein equation, conveniently dualized) can be derived from an action of the form

$$S = \int \sqrt{g}\left\{R + \tfrac{1}{2} g_{xy} \partial_\mu \phi^x \partial^\mu \phi^y + \tfrac{1}{2 \cdot 3!} a^{IJ} H_I H_J\right\}. \tag{163}$$

The effective action and Hamiltonian constraint can be immediately written using the data in the above action and take the explicit form

$$S[\tilde{U}, \phi^x] = \int d\rho \left\{\dot{\tilde{U}}^2 + \tfrac{1}{3} g_{xy} \dot{\phi}^x \dot{\phi}^y - e^{2U} V_{\text{bs}}\right\}, \tag{164}$$

$$(\omega/2)^2 = \dot{\tilde{U}}^2 + \tfrac{1}{3} g_{xy} \dot{\phi}^x \dot{\phi}^y + e^{2U} V_{\text{bs}}. \tag{165}$$

In these two equations $V_{\text{bs}}(\phi, p)$ is the *black-string potential*. It is again given by two equivalent expressions

$$-V_{\text{bs}}(\phi, p) \equiv a_{IJ}\, p^I p^J = \mathcal{Z}_{\text{m}}^2 + 3 \partial_x \mathcal{Z}_{\text{m}} \partial^x \mathcal{Z}_{\text{m}}, \tag{166}$$

where $\mathcal{Z}_{\mathrm{m}}$ is *(magnetic) string central charge*, defined by

$$\mathcal{Z}_{\mathrm{m}}(\phi, p) = h_I(\phi) p^I, \tag{167}$$

and where we have denoted by p^I the string charges.

$\mathcal{Z}_{\mathrm{m}}(\phi, p)$ plays for black strings exactly the same role played by $\mathcal{Z}_{\mathrm{e}}(\phi, q)$ for black holes: it allows us to rewrite the effective action *à la Bogomol'nyi* and find flow equations, the string tension of supersymmetric strings is determined by its value at infinity and the entropy density by its near-horizon behavior. The supersymmetric attractors are also critical values of $\mathcal{Z}_{\mathrm{m}}$.

Once we have solved the equations and we have found the metric function e^U we just have to replace it in the metric

$$ds^2 = e^{\tilde{U}}\left[e^{\frac{\omega}{2}\rho}dt^2 - e^{-\frac{\omega}{2}\rho}dy^2\right] - e^{-2\tilde{U}}\left(\frac{\omega/2}{\sinh\left(\frac{\omega}{2}\rho\right)}\right)^2\left[\left(\frac{\omega/2}{\sinh\left(\frac{\omega}{2}\rho\right)}\right)^2 d\rho^2 + d\Omega^2_{(2)}\right], \tag{168}$$

which describes a black string in a 5-dimensional spacetime lying along the direction parametrized by the coordinate y.

3 The H-FGK Formalism

The FGK equations are still hard to solve if we want to construct explicitly the black-hole solutions and we are not happy enough with just determining the attractors. Even the first-order equations are difficult to solve and that requires the determination of the corresponding fake central charge in advance.

In contrast, the supersymmetric solutions of supergravity theories are easy to construct, probably because of the choice of building blocks (the basic functions whose equations need to be solved) which have the property that they transform linearly under duality, unlike the scalar fields.

Could we use these building blocks in non-supersymmetric solutions? In other words: can we use the supersymmetry-inspired variables transforming linearly under duality in the FGK action?

In [45, 50] it was shown that in some models of $\mathcal{N} = 2, d = 4$ and $\mathcal{N} = 1, d = 5$ supergravity coupled to vector supermultiplets, the non-extremal solutions can be written in terms of the same building blocks as the supersymmetric ones, the difference being the functional form of the building blocks, which are always harmonic functions in the supersymmetric cases.

This is not accident: these supergravity theories can be formulated in terms of those building blocks, as shown in [51–53]. It is natural to try to combine this fact and the FGK formalism, as suggested above. In order to do this, we are going to make a quick review of the form that the supersymmetric, static, spherically symmetric, asymptotically black-hole solutions of these solutions have for the case of $\mathcal{N} = 2, d = 4$ supergravity. We start by reviewing what a supersymmetric solution is.

3.1 Supersymmetric Solutions

A configuration of a supersymmetric theory (such as a supergravity theory) is said to be *supersymmetric* if it is invariant under some supersymmetry transformations. A supergravity theory is invariant under an infinite number of these transformations, because they are local, but a supersymmetric solutions will only be invariant under some.

The supersymmetry transformations are generated by the spinor ε and their action on the bosonic fields ϕ^b and fermionic field ϕ^f takes the generic form[21]

$$\begin{aligned} \delta_\varepsilon \phi^b &\sim \bar{\varepsilon}\phi^f, \\ \delta_\varepsilon \phi^f &\sim \partial\varepsilon + (\phi^b + \bar{\phi}^f \phi^f)\varepsilon, \end{aligned} \tag{169}$$

We are usually interested in purely bosonic configurations only. These have $\phi^f = 0$ which is always a consistent truncation. Then, the supersymmetry variations of the bosonic fields vanish automatically. A bosonic configuration will be supersymmetric if the second equation, the supersymmetry variation of the fermionic fields of the theory, vanishes for some parameter $\varepsilon(x)$

$$\delta_\varepsilon \phi^f \sim \partial\varepsilon + \phi^b \varepsilon = 0. \tag{170}$$

This condition is known as *Killing Spinor Equations (KSEs)*.

Supersymmetric solutions enjoy remarkable properties. They can be identified, characterized, classified and constructed by analyzing the KSEs under the assumption that a solution with $\varepsilon \neq 0$ exists.[22]

We are going to review the results for supersymmetric black-hole solutions of ungauged $\mathcal{N} = 2, d = 4$ supergravity coupled to vector supermultiplets in the next section.

3.2 Supersymmetric, Static $\mathcal{N} = 2, d = 4$ Black Holes

All the supersymmetric solutions of ungauged $\mathcal{N} = 2, d = 4$ supergravity coupled to vector supermultiplets were classified in [54] and a recipe to construct all of them

[21] The term $\partial\varepsilon$ only occurs in theories with local supersymmetry: supergravities.

[22] For more information on supersymmetric solutions see, for instance, [1] and references therein.

was found. For the black-hole solutions of a model characterized by the canonical, covariantly holomorphic symplectic section $\mathcal{V}^M$ we can proceed as follows:

1. Introduce the auxiliary function X with the same Kähler weight as $\mathcal{V}^M$ to define define real symplectic vectors $\mathcal{R}^M$ and $\mathcal{I}^M$, which have vanishing Kähler weight

$$\mathcal{R}^M + i\mathcal{I}^M \equiv \mathcal{V}^M / X. \tag{171}$$

2. The components of $\mathcal{R}^M$ can be expressed in terms of the components of $\mathcal{I}^M$ only. Finding the expressions $\mathcal{R}^M(\mathcal{I})$ is equivalent to solving the so-called *stabilization* or *Freudenthal duality* equations [55].
3. The map $\mathcal{I}^M \longrightarrow \mathcal{R}^M(\mathcal{I})$ defines an operation called *Freudenthal duality* [56–58] that can be generalized to any symplectic vector of the same theory. We will denote the Freudenthal dual of $\mathcal{I}^M$ by $\tilde{\mathcal{I}}^M \equiv \mathcal{R}^M(\mathcal{I})$.
   This operation is an antiinvolution

$$\tilde{\tilde{\mathcal{I}}}^M = -\mathcal{I}^M. \tag{172}$$

4. We define the *Hesse potential* $W(\mathcal{I})$ as the symplectic product of $\mathcal{I}^M$ and its Freudenthal dual

$$W(\mathcal{I}) = \tilde{\mathcal{I}}_M \mathcal{I}^M = \mathcal{R}_M(\mathcal{I})\mathcal{I}^M. \tag{173}$$

   Its most fundamental property is that it is homogenous of second degree on the $\mathcal{I}^M$.
5. In the static, spherically symmetric black-hole solutions, the components of the symplectic vector $\mathcal{I}^M$ are harmonic functions in $\mathbb{E}^3$, H^M with a single pole, satisfying the constraint

$$H_M dH^M = 0. \tag{174}$$

   Using the FGK coordinate ρ of the first lecture, these functions must take the form

$$\mathcal{I}^M = H^M \equiv A^M - B^M \rho, \quad \text{with} \quad A_M B^M = 0. \tag{175}$$

   It can be shown that the integration constants B^M can be identified with the electric and magnetic charges

$$B^M = \mathcal{Q}^M / \sqrt{2}. \tag{176}$$

6. The choice of harmonic functions determines completely all the fields of the supersymmetric solution. We just have to give the recipe to reconstruct them in terms of the harmonic functions. First of all, the metric function is given by the Hesse potential

$$e^{-2U} = \frac{1}{2|X|^2} = W(H). \tag{177}$$

Thus, in the near-horizon limit

$$e^{-2U} \sim \frac{1}{2r^2} W(\mathcal{Q}), \tag{178}$$

where $W(\mathcal{Q})$ is the Hesse potential evaluated on the charges and the black-hole entropy is completely determined by the Hesse potential which, being symplectic-invariant is duality invariant[23]

$$S/\pi = W(\mathcal{Q})/2. \tag{179}$$

The vector field strengths and the complex scalar fields are given in terms of the harmonic functions by

$$\mathcal{F}^M = -\tfrac{1}{\sqrt{2}} d(\tilde{H}^M e^{2U}) \wedge dt - \tfrac{1}{\sqrt{2}} e^{2U} \star (dt \wedge dH^M),$$

$$Z^i = \frac{\tilde{H}^i + iH^i}{\tilde{H}^0 + iH^0}.$$

Observe that the auxiliary variable X can be written in terms of the metric function and a phase α

$$X = \tfrac{1}{\sqrt{2}} e^{U+i\alpha}, \tag{180}$$

which does not occur in any of the bosonic fields and, therefore, does not occur in the FGK action.

Also, observe that the scalar fields and the metric (the Hesse potential) are invariant under Freudenthal duality, but not the vector fields: their transformation is equivalent to the replacement of the charge symplectic vector $\mathcal{Q}^M$ by its Freudenthal dual $\tilde{\mathcal{Q}}^M$. Freudenthal duality will not respect supersymmetry but, will it transform solutions into solutions? To investigate this and other questions we want to replace the variables used in the original FGK formalism U, Z^i by the symplectic vector $\mathcal{I}^M$ which we will denote by H^M to follow the literature.

3.3 The H-FGK Formalism

The details of the change of variables from U, Z^i to H^M are rather technical and can be found in the original reference [59]. We will just quote the result (effective action and Hamiltonian constraint)

[23] All symplectic vectors transform linearly under the duality transformations, just as the vector fields, according to the Gaillard and Zumino results reviewed in the first lecture.

$$- S_{\text{H-FGK}}[H] = \int d\rho \left\{ \tfrac{1}{2} g_{MN} \dot{H}^M \dot{H}^N - V \right\}, \tag{181}$$

$$-r_0^2 = \tfrac{1}{2} g_{MN} \dot{H}^M \dot{H}^N + V, \tag{182}$$

where the metric $g_{MN}(H)$ and the potential $V(H)$ are given in terms of $W(H)$ by

$$g_{MN}(H) \equiv \partial_M \partial_N \log W - 2 \frac{H_M H_N}{W^2}, \tag{183}$$

$$V(H) \equiv \left\{ -\tfrac{1}{4} \partial_M \partial_N \log W + \frac{H_M H_N}{W^2} \right\} \mathcal{Q}^M \mathcal{Q}^N = -V_{\text{bh}}/W, \tag{184}$$

and we will notice that this change of variables introduces one extra variable: from $2\bar{n}+1$ to $2\bar{n}+2$. This means that the H-FGK effective action must be invariant under a local symmetry. It is not difficult to realize that the symmetry must be associated to local shifts of α, the phase of X which, as mentioned above, does not occur in the variables of the FGK formalism but enters in the definition of the variables of the H-FGK formalism.[24]

A sign of the existence of a local symmetry that would allow us to eliminate one variable is that the metric $g_{MN}(H)$ always admits a null eigenvector [53, 58]

$$\tilde{H}^M g_{MN} = 0, \tag{185}$$

and it is singular. This has to be taken into account when deriving the equations of motion, which take the form

$$g_{MN} \ddot{H}^N + (\partial_N g_{PM} - \tfrac{1}{2} \partial_M g_{NP}) \dot{H}^N \dot{H}^P + \partial_M V = 0. \tag{186}$$

It is not difficult to show that $\dot{H}^M = \mathcal{Q}^M/\sqrt{2}$ is always a solution, which is the general SBH of the theory. This is not so easy to prove in the FGK formalism.

Multiply these equations with H^M and using the homogeneity properties of the Hesse potential and the Hamiltonian constraint we get

$$\tilde{H}_M \left(\ddot{H}^M - r_0^2 H^M \right) + \frac{(\dot{H}^M H_M)^2}{W} = 0. \tag{187}$$

If we impose the condition

$$H_M \dot{H}^M = 0, \tag{188}$$

[24] Let us stress that this local shift of the phase of X cannot be interpreted as a Kähler transformation because such a transformation would act on all the fields with non-vanishing Kähler weight, which is not the case.

which we have not used at all in the definition of the H-FGK effective action but that arises naturally in the recipe for constructing static SBHs (in particular with no NUT charge [60]) the above equation takes the form

$$\tilde{H}_M \left(\ddot{H}^M - r_0^2 H^M\right) = 0, \tag{189}$$

which are solved by harmonic functions in the extremal case and by hyperbolic functions in the non-extremal one. See [61] for an exhaustive study of these solutions.

Not all solutions are of this form, though. The most general non-supersymmetric ones of the t^3 and STU models, for instance, have non-harmonic H^Ms [37, 62–64]. We have called these solutions, which do not satisfy the above constraint (188), *unconventional solutions* [58] and there is still much to learn about them. The most general non-extremal solution of the STU model has been proposed in [65] but not in terms of the H^M variables.

All the FGK theorems and, in particular, the attractor mechanism, can be recast in these variables:

1. The values of the H-variables on the horizon of an extremal black hole, H_{h}^M, extremize the black-hole potential

$$\partial_M V_{\text{bh}}|_{H_{\text{h}}} = 0. \tag{190}$$

The H_{h}^M are the attractors in this language and are defined up to a global factor because $V_{\text{bh}}(H)$ is homogeneous of degree zero on the H-variables. The values of the scalars on the horizon (the usual attractors) are completely determined by these:

$$Z_{\text{h}}^i = \frac{\tilde{H}_{\text{h}}^i + iH_{\text{h}}^i}{\tilde{H}_{\text{h}}^0 + iH_{\text{h}}^0}. \tag{191}$$

For SBHs the attractors are just $H_{\text{h}}^M = -\mathcal{Q}^M/\sqrt{2}$ and

$$Z_{\text{h}}^i = \frac{\tilde{\mathcal{Q}}^i + ip^i}{\tilde{\mathcal{Q}}^0 + ip^0}. \tag{192}$$

2. The entropy is completely determined by the attractors:

$$S/\pi = W(H_{\text{h}}). \tag{193}$$

For SBHs

$$S/\pi = W(\mathcal{Q})/2. \tag{194}$$

3.4 Freudenthal Duality

In this formalism, Freudenthal duality appears in two ways:

1. If $H_{\mathrm{h}}^{M} = B^{M}$ is an attractor extremizing the black-hole potential, then its Freudenthal dual $\tilde{B}^{M}$ is also an attractor that extremizes the same black-hole potential and the entropy of the corresponding extremal black holes is the same, because its expression (the Hesse potential evaluated on B^{M}) is manifestly Freudenthal-duality invariant:

$$S(B)/\pi = \tfrac{1}{2}W(B) = \tfrac{1}{2}W(\tilde{B}) = S(\tilde{B})/\pi. \tag{195}$$

   This fact was first observed in a more restricted case in [56].
2. There is a local symmetry in the H-FGK action, as we have discussed above. We are going to see that the discrete Freudenthal duality is nothing but one of of these local transformations for a particular (constant) choice of the gauge parameter.

The existence of a null eigenvector of the metric, $\tilde{H}^{M}$ can be used to prove the following identity that relates the equations of motion of the H-FGK action

$$\tilde{H}^{M}\frac{\delta S_{\text{H-FGK}}}{\delta H^{M}} = 0, \tag{196}$$

and which can be seen as the Noether identity associated to a local symmetry of the theory. Multiplying this identity by an infinitesimal arbitrary function $f(\rho)$ and integrating the expression over ρ we get an expression that we can rewrite as the transformation of the action under a local symmetry with parameter f:

$$\delta_f S_{\text{H-FGK}} = \int d\rho\, \delta_f H^{M}\frac{\delta S_{\text{H-FGK}}}{\delta H^{M}} = 0, \quad \text{where} \quad \delta_f H^{M} \equiv f(\rho)\tilde{H}^{M}. \tag{197}$$

The above transformations have been explicitly checked to leave invariant the complete H-FGK action. Their finite form is

$$(\tilde{H}'^{M} + iH'^{M}) = e^{if(\rho)}(\tilde{H}^{M} + iH^{M}), \quad \Rightarrow \mathcal{V}^{M}/X' = e^{if(\rho)}\mathcal{V}^{M}/X, \tag{198}$$

which corresponds to a transformation of the phase of X, as we advanced

$$\delta_f \alpha = -f. \tag{199}$$

For $f = -\pi/2$ we recover the discrete Freudenthal duality transformations.

The reason for the existence of this local symmetry in the H-FGK action is clear, but its physical meaning is unknown. There are no higher-dimensional analogues of this purely 4-dimensional symmetry that preserves the black-hole entropy. More work is necessary to understand this mysterious symmetry.

Acknowledgments TO would like to thank the organizers of the *School on Theoretical Frontiers in Black Holes and Cosmology*, the International Institute of Physics of Natal and, specially, Emanuele Orazi for the opportunity to participate in an excellent school in a wonderful setting, for the financial support and, last, but not least, for their kindness during the school and workshop. This work has been supported in part by the Spanish Ministry of Science and Education grant FPA2012-35043-C02-01, the *Centro de Excelencia Severo Ochoa* Program grant SEV-2012-0249. The work of PFR was supported by *Severo Ochoa* pre-doctoral grant SVP-2013-067903 TO wishes to thank M.M. Fernández for her permanent support.

References

1. T. Ortín, *Gravity and Strings*, 2nd edn. (Cambridge University Press, Cambridge, 2015)
2. S. Ferrara, G.W. Gibbons, R. Kallosh, Black holes and critical points in moduli space. Nucl. Phys. B **500**, 75–93 (1997)
3. S. Ferrara, R. Kallosh, A. Strominger, $\mathcal{N} = 2$ extremal black holes. Phys. Rev. D **52**, 5412 (1995)
4. A. Strominger, Macroscopic entropy of $\mathcal{N} = 2$ extremal black holes. Phys. Lett. B **383**, 39 (1996)
5. S. Ferrara, R. Kallosh, Supersymmetry and attractors. Phys. Rev. D **54**, 1514 (1996)
6. S. Ferrara, R. Kallosh, Universality of supersymmetric attractors. Phys. Rev. D **54**, 1525 (1996)
7. M.K. Gaillard, B. Zumino, Duality rotations for interacting fields. Nucl. Phys. B **193**, 221 (1981)
8. E. Schrödinger, Contributions to Born's new theory of the electromagnetic field. Proc. Roy. Soc. Lond. **A150**, 465–477 (1935)
9. C.W. Misner, J.A. Wheeler, Classical physics as geometry: gravitation, electromagnetism, unquantized charge, and mass as properties of curved empty space. Ann. Phys. **2**, 525 (1957)
10. P. van Nieuwenhuizen, J.A.M. Vermaseren, One loop divergences in the quantum theory of supergravity. Phys. Lett. B **65**, 263 (1976)
11. M.T. Grisaru, P. van Nieuwenhuizen, J.A.M. Vermaseren, One loop renormalizability of pure supergravity and of Maxwell-Einstein theory in extended supergravity. Phys. Rev. Lett. **37**, 1662 (1976)
12. S. Ferrara, J. Scherk, B. Zumino, Algebraic properties of extended supergravity theories. Nucl. Phys. B **121**, 393 (1977)
13. E. Cremmer, J. Scherk, S. Ferrara, U(n) invariance in extended supergravity. Phys. Lett. B **68**, 234 (1977)
14. G.W. Gibbons, Antigravitating black hole solitons with scalar hair in $N = 4$ supergravity. Nucl. Phys. B **207**, 337 (1982)
15. G.W. Gibbons, R. Kallosh, B. Kol, Moduli, scalar charges and the first law of black hole thermodynamics. Phys. Rev. Lett. **77**, 4992–4995 (1996)
16. P. Breitenlohner, D. Maison, G.W. Gibbons, Four-dimensional black holes from Kaluza-Klein theories. Commun. Math. Phys. **120**, 295 (1988)
17. E. Bergshoeff, W. Chemissany, A. Ploegh, M. Trigiante, T. Van Riet, Generating geodesic flows and supergravity solutions. Nucl. Phys. B **812**, 343 (2009)
18. P. Dobiasch, D. Maison, Stationary, spherically symmetric solutions of Jordan's unified theory of gravity and electromagnetism gen. Rel. Grav. **14**, 231 (1982)
19. D.V. Gal'tsov, O.A. Rytchkov, Generating branes via sigma models. Phys. Rev. D **58**, 122001 (1998)
20. P. Fré, V. Gili, F. Gargiulo, A.S. Sorin, K. Rulik, M. Trigiante, Cosmological backgrounds of superstring theory and solvable algebras: oxidation and branes. Nucl. Phys. B **685**, 3 (2004)
21. P. Fré, A.S. Sorin, Integrability of supergravity billiards and the generalized Toda lattice equation. Nucl. Phys. B **733**, 334 (2006)

22. M. Gunaydin, A. Neitzke, B. Pioline, A. Waldron, Quantum attractor flows. JHEP **0709**, 056 (2007)
23. D. Gaiotto, W. Li, M. Padi, Non-supersymmetric attractor flow in symmetric spaces. JHEP **0712**, 093 (2007)
24. W. Chemissany, A. Ploegh, T. Van Riet, A Note on scaling cosmologies, geodesic motion and pseudo-susy. Class. Quantum Grav. **24**, 4679 (2007)
25. M. Berkooz, B. Pioline, 5D black holes and non-linear sigma models. JHEP **0805**, 045 (2008)
26. P. Fré, A.S. Sorin, The Weyl group and asymptotics: all supergravity billiards have a closed form general integral. Nucl. Phys. B **815**, 430 (2009)
27. G. Bossard, H. Nicolai, K.S. Stelle, Universal BPS structure of stationary supergravity solutions. JHEP **0907**, 003 (2009)
28. G. Bossard, The extremal black holes of N=4 supergravity from so(8,2+n) nilpotent orbits. Gen. Rel. Grav. **42**, 539 (2010)
29. W. Chemissany, P. Fré, A.S. Sorin, The integration algorithm of Lax equation for both generic Lax matrices and generic initial conditions. Nucl. Phys. B **833**, 220 (2010)
30. W. Chemissany, J. Rosseel, M. Trigiante, T. Van Riet, The full integration of black hole solutions to symmetric supergravity theories. Nucl. Phys. B **830**, 391 (2010)
31. M. Günaydin, *Lectures on Spectrum Generating Symmetries and U-duality in Supergravity, Extremal Black Holes, Quantum Attractors and Harmonic Superspace*,
32. P. Figueras, E. Jamsin, J.V. Rocha, A. Virmani, Integrability of five dimensional minimal supergravity and charged rotating black holes. Class. Quant. Grav. **27**, 135011 (2010)
33. W. Chemissany, P. Fre, J. Rosseel, A.S. Sorin, M. Trigiante, T. Van Riet, Black holes in supergravity and integrability. JHEP **1009**, 080 (2010)
34. A. Sen, Entropy function for heterotic black holes. JHEP **0603**, 008 (2006)
35. E.B. Bogomol'nyi, Stability of Classical Solutions. Sov. J. Nucl. Phys. **24** (1976) 449 [*Yad. Fiz.* **24** (1976) 861]
36. A. Ceresole, G. Dall'Agata, Flow equations for non-BPS extremal black holes. JHEP **0703**, 110 (2007)
37. G. Lopes Cardoso, A. Ceresole, G. Dall'Agata, J.M. Oberreuter and J. Perz, First-order flow equations for extremal black holes in very special geometry, JHEP **0710**, 063 (2007)
38. L. Andrianopoli, R. D'Auria, E. Orazi, M. Trigiante, First order description of black holes in moduli space. JHEP **0711**, 032 (2007)
39. S. Ferrara, A. Marrani, E. Orazi, Maurer-Cartan equations and black hole superpotentials in $N = 8$ supergravity. Phys. Rev. D **81**, 085013 (2010)
40. G. Bossard, Y. Michel, B. Pioline, Extremal black holes, nilpotent orbits and the true fake superpotential. JHEP **1001**, 038 (2010)
41. A. Ceresole, G. Dall'Agata, S. Ferrara, A. Yeranyan, Universality of the superpotential for $d = 4$ extremal black holes. Nucl. Phys. B **832**, 358 (2010)
42. A. Ceresole, G. Dall'Agata, S. Ferrara, A. Yeranyan, First order flows for $\mathcal{N} = 2$ extremal black holes and duality invariants. Nucl. Phys. B **824**, 239 (2010)
43. A. de Antonio Martín, T. Ortín, C.S. Shahbazi, The FGK formalism for black p-branes in d dimensions, JHEP **1205**, 045 (2012)
44. G.T. Horowitz, A. Strominger, Black strings and p-Branes. Nucl. Phys. B **360**, 197 (1991)
45. P. Meessen, T. Ortín, Non-extremal black holes of $\mathcal{N} = 2$, d=5 supergravity. Phys. Lett. B **707**, 178–183 (2012)
46. B. Janssen, P. Smyth, T. Van Riet, B. Vercnocke, A first-order formalism for timelike and spacelike brane solutions. JHEP **0804**, 007 (2008)
47. J.X. Lu, ADM masses for black strings and p-branes. Phys. Lett. B **313**, 29 (1993)
48. R.C. Myers, Stress tensors and Casimir energies in the AdS/CFT correspondence. Phys. Rev. D **60**, 046002 (1999)
49. W.A. Sabra, General BPS black holes in five-dimensions. Mod. Phys. Lett. A **13**, 239 (1998)
50. P. Galli, T. Ortín, J. Perz, C.S. Shahbazi, Non-extremal black holes of $\mathcal{N} = 2$, d=4 supergravity. JHEP **1107**, 041 (2011)
51. T. Mohaupt, K. Waite, Instantons, black holes and harmonic functions. JHEP **0910**, 058 (2009)

52. T. Mohaupt, O. Vaughan, Non-extremal black holes, harmonic functions, and attractor equations. Class. Quantum Grav. **27**, 235008 (2010)
53. T. Mohaupt, O. Vaughan, The Hesse potential, the c-map and black hole solutions. JHEP **1207**, 163 (2012)
54. P. Meessen, T. Ortín, The supersymmetric configurations of N=2, D=4 supergravity coupled to vector supermultiplets. Nucl. Phys. B **749**, 291 (2006)
55. P. Bueno, P. Galli, P. Meessen, T. Ortín, Black holes and equivariant charge vectors in $\mathcal{N} = 2$, d=4 supergravity. JHEP **1309**, 010 (2013)
56. L. Borsten, D. Dahanayake, M.J. Duff, W. Rubens, Black holes admitting a Freudenthal dual. Phys. Rev. D **80**, 026003 (2009)
57. S. Ferrara, A. Marrani, A. Yeranyan, Freudenthal duality and generalized special geometry. Phys. Lett. B **701**, 640 (2011)
58. P. Galli, P. Meessen, T. Ortín, The Freudenthal gauge symmetry of the black holes of $\mathcal{N} = 2$, d=4 supergravity. JHEP **1305**, 011 (2013)
59. P. Meessen, T. Ortín, J. Perz, C.S. Shahbazi, H-FGK-formalism for black-hole solutions of $\mathcal{N} = 2$, d=4 and d=5 supergravity. Phys. Lett. B **709**, 260 (2012)
60. J. Bellorín, P. Meessen, T. Ortín, Supersymmetry, attractors and cosmic censorship. Nucl. Phys. B **762**, 229–255 (2007)
61. P. Galli, T. Ortín, J. Perz, C.S. Shahbazi, Black hole solutions of $\mathcal{N} = 2$, d=4 supergravity with a quantum correction, in the H-FGK formalism
62. E.G. Gimon, F. Larsen, J. Simon, Constituent model of extremal non-BPS black holes. JHEP **0907**, 052 (2009)
63. P. Galli, K. Goldstein, S. Katmadas, J. Perz, First-order flows and stabilisation equations for non-BPS extremal black holes. JHEP **1106**, 070 (2011)
64. G. Bossard, S. Katmadas, Duality covariant non-BPS first order systems. JHEP **1209**, 100 (2012)
65. D.D.K. Chow, G. Compère, Seed for general rotating non-extremal black holes of $\mathcal{N} = 8$ supergravity. Class. Quantum Grav. **31**, 022001 (2014)

Introductory Lectures on Extended Supergravities and Gaugings

Antonio Gallerati and Mario Trigiante

Abstract In an ungauged supergravity theory, the presence of a *scalar potential* is allowed only for the minimal $N = 1$ case. In extended supergravities, a non-trivial scalar potential can be introduced without explicitly breaking supersymmetry only through the so-called *gauging procedure*. The latter consists in promoting a suitable global symmetry group to local symmetry to be gauged by the vector fields of the theory. Gauged supergravities provide a valuable approach to the study of superstring flux-compactifications and the construction of phenomenologically viable, string-inspired models. The aim of these lectures is to give a pedagogical introduction to the subject of gauged supergravities, covering just selected issues and discussing some of their applications.

1 Introduction

A long-standing problem of high energy theoretical physics is the formulation of a fundamental theory unifying the four interactions. Superstring theory in ten dimensions and M-theory in eleven seem to provide a promising theoretical framework where this unification could be achieved. However, there are many shortcomings originating from this theoretical formulation.

First of all, these kinds of theories are defined in dimensions $D > 4$, and, since we live in a four-dimensional universe, a fundamental requirement for any predictable model is the presence of a mechanism of dimensional reduction from ten or eleven dimensions to four. Moreover, the non-perturbative dynamics of the theory is far from being understood, and there is no mechanism to select a vacuum state for our universe (i.e. it is not clear how to formulate a phenomenological viable description for the model). Finally, there are more symmetries than those observed experimentally. These models, in fact, encode Supersymmetry (SUSY), but our universe is not

A. Gallerati · M. Trigiante (✉)
Department DISAT, Politecnico di Torino, Corso Duca Degli Abruzzi 24, 10129 Torino, Italy
e-mail: mario.trigiante@gmail.com

R. Kallosh and E. Orazi (eds.), *Theoretical Frontiers in Black Holes and Cosmology*, Springer Proceedings in Physics 176,
DOI 10.1007/978-3-319-31352-8_2

supersymmetric and its gauge interactions are well described, at our energy scales, by the Standard Model (SM). Therefore deriving a phenomenologically viable model from string/M-theory also requires the definition of suitable mechanisms of supersymmetry breaking.

Spontaneous compactification. The simplest way for deriving a four-dimensional theory from a higher dimensional one is through *spontaneous compactification* which generalizes the original Kaluza–Klein (KK) compactification of five-dimensional general relativity on a circle. We consider the low-energy dynamics of superstring/M-theory on space-time solutions with geometry of the form

$$\mathbf{M_4}^{(1,3)} \times \mathscr{M}_{int}, \tag{1}$$

where $\mathbf{M_4}^{(1,3)}$ is the maximally symmetric four dimensional space-time with Lorentzian signature and $\mathscr{M}_{int}$ is a compact internal manifold. The $D = 10$ or $D = 11$ fields, excitations of the microscopic fundamental theory, are expanded in normal modes ($Y_{(n)}$) on the internal manifold

$$\Phi(x^\mu,\, y^\alpha) = \sum_{(n)} \Phi_{(n)}(x^\mu)\; Y_{(n)}(y^\alpha), \tag{2}$$

the coefficients $\Phi_{(n)}$ of this expansion describing massive fields in $\mathbf{M_4}^{(1,3)}$ with mass of the order of $\frac{1}{R}$, where R is the "size" of the internal manifold $\mathscr{M}_{int}$. These are the Kaluza–Klein states, forming an infinite tower.

In many cases, a consistent truncation of the massless modes $\Phi_{(0)}$ is well described by a $D = 4$ Supergravity theory (SUGRA), an effective field theory consistently describing superstring dynamics on the chosen background at energies Λ, where

$$\Lambda \ll \frac{1}{R} \ll \text{string scale}. \tag{3}$$

The effective supergravity has $\mathbf{M_4}^{(1,3)}$ as vacuum solution, and its general features depend on the original microscopic theory and on the chosen compactification. In fact, the geometry of $\mathscr{M}_{int}$ affects the amount of supersymmetry of the low-energy SUGRA, as well as its internal symmetries.

Internal manifold, compactification and dualities. According to the Kaluza–Klein procedure, the isometries of $\mathscr{M}_{int}$ induce gauge symmetries in the lower-dimensional theory gauged by the vectors originating from the metric in the reduction mechanism (KK vectors). The internal manifold $\mathscr{M}_{int}$ also affects the field content of the $D = 4$ theory, which arrange in supermultiplets according to the residual (super)symmetry of the vacuum solution $\mathbf{M_4}^{(1,3)}$.

The compactification of superstring/M-theory on a *Ricci-flat* internal manifold (like a torus or a Calabi Yau space) in the absence of fluxes of higher-order form field-strengths, yields, in the low-energy limit, an effective four-dimensional SUGRA, which involves the massless modes on $\mathbf{M_4}^{(1,3)}$. The latter is an ungauged theory, namely the vector fields are not minimally coupled to any other field of the theory. At the classical level, ungauged supergravity models feature an on-shell global symmetry group, which was conjectured to encode the known superstring/M-theory dualities [3]. The idea behind these dualities is that superstring/M-theory provide a redundant description for the same microscopic degrees of freedom: different compactifications of the theory turns out to define distinct descriptions of the same quantum physics. These descriptions are connected by dualities, which also map the correspondent low-energy description into one another. The global symmetry group G of the classical $D = 4$ supergravity is in part remnant of the symmetry of the original higher dimensional theory, i.e. invariance under reparametrizations in $\mathscr{M}_{int}$.[1]

Ungauged versus Gauged models. From a phenomenological point of view, extended supergravity models on four dimensional Minkowski vacua, obtained through ordinary Kaluza–Klein reduction on a Ricci-flat manifold, are not consistent with experimental observations. These models typically contain a certain number of massless scalar fields—which are associated with the geometry of the internal manifold $\mathscr{M}_{int}$—whose vacuum expectation values (vevs) define a continuum of degenerate vacua. In fact, there is no scalar potential that encodes any scalar dynamics, so we cannot avoid the degeneracy. This turns into an intrinsic lack of predictiveness for the model, in addition to a field-content of the theory which comprises massless scalar fields coupled to gravity, whose large scale effects are not observed in our universe.

Another feature of these models, as we said above, is the absence of a internal local-symmetry gauged by the vector fields. This means that no matter field is charged under a gauge group, hence the name *ungauged supergravity*.

Realistic quantum field theory models in four dimensions, therefore, require the presence of a non-trivial scalar potential, which could solve (in part or completely) moduli-degeneracy problem and, on the other hand, select a vacuum state for our universe featuring desirable physical properties like, for instance

- introduce mass terms for the scalars;
- support the presence of some effective cosmological constant;
- etc.

The phenomenologically uninteresting ungauged SUGRAs can provide a general framework for the construction of realistic model. In a $D = 4$ extended supergravity model (i.e. having $N > 1$ susy), it is possible to introduce a scalar potential, without explicitly breaking supersymmetry, through the so-called *gauging procedure* [4–12]. The latter can be seen as a *deformation* of an ungauged theory and consists in promoting some suitable subgroup G_g of the global symmetry group of the Lagrangian to *local* symmetry. This can be done by introducing minimal couplings for the vector fields, mass deformation terms and the scalar potential itself. The coupling of the

[1] In part they originate from gauge symmetries associated with the higher dimensional antisymmetric tensor fields.

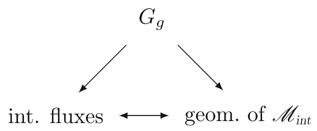
(formerly abelian) vector fields to the new local gauge group gives us matter fields that are charged under this new local gauge symmetry.

In particular, in the presence of fluxes of higher-order form field-strengths across cycles of the internal manifold

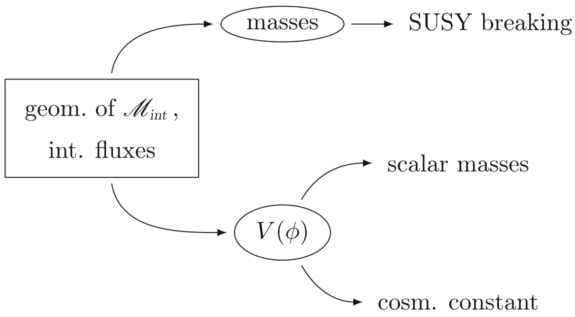
$$\langle \int_{\Sigma_p} F_{(p)} \rangle \neq 0, \tag{4}$$

the non-linear dynamics of the low lying modes (or of a consistent truncation thereof) is, in most cases, captured by a $D = 4$ theory which is gauged.

The gauge group G_g of the lower dimensional SUGRA depends on the geometry of the internal manifold and on the possible internal fluxes



The fluxes and the structure of the internal manifold, aside from the gauge symmetry, also induce masses and a scalar potential $V(\phi)$ (for reviews on flux-compactifications see [13–15]). These mass terms produce, in general, supersymmetry breaking already at the classical level (which is phenomenologically desirable) and the presence of a scalar potential lift the moduli degeneracy (already at the tree level) and may produce an effective cosmological constant term

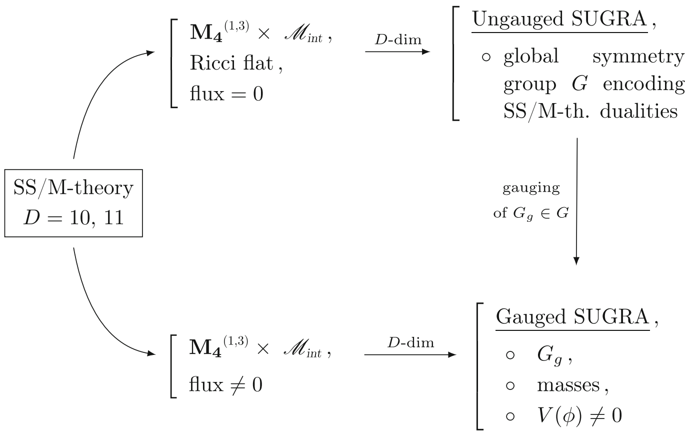


Supergravity theories in D dimensions are consistently defined independently of their higher-dimensional origin, and are totally defined by

- amount of supersymmetry;
- field content;
- local symmetry, gauged by the vector fields (feature of gauged SUGRAs).

When originating from superstring/M-theory compactifications, gauged SUGRAs offer a unique window on the perturbative low-energy dynamics of these theories, since they describe the full non-linear dynamics of the low lying modes. In general, there is a correspondence between vacua of the microscopic fundamental theory and vacua of the low-energy supergravity. However, there are several gauged SUGRAs whose superstring/M-theory origin is not known.

Gauged supergravities are obtained from ungauged ones, with the same field content and amount of SUSY, through the gauging previously mentioned procedure, which is well-defined and works provided the gauge group G_g satisfies some stringent conditions originating from the requirement of gauge invariance and supersymmetry.



As mentioned above, gauging is the only known way to introduce a scalar potential in extended supergravities without an explicit breaking of the supersymmetry. However this procedure will in general break the global symmetry group of the ungauged theory. The latter indeed acts as a generalized electric-magnetic duality and is thus broken by the minimal couplings, which only involve the electric vector fields. As a consequence of this, in a gauged supergravity we loose track of the string/M-theory dualities, which were described by global symmetries of the original ungauged theories.

The drawback can be avoided using the *embedding tensor* formulation of the gauging procedure [5, 8, 16–18] in which all deformations involved by the gauging is encoded in a single object, the embedding tensor, which is itself covariant with respect to the global symmetries of the ungauged model. This allows to formally restore such symmetries at the level of the gauged field equations and Bianchi identities, provided the embedding tensor is transformed together with all the other fields. The global symmetries of the ungauged theory now act as equivalences between gauged supergravities. Since the embedding tensor encodes all background quantities in the compactification describing the fluxes and the structure of the internal manifold, the action of the global symmetry group on it allows to systematically study the effect of dualities on flux compactifications.

These lectures are organized as follows.

In Sect. 2 we briefly review the general structure of ungauged supergravities.

In Sect. 3 we discuss the gauging procedure in the electric symplectic frame and comment on the relation between the embedding tensor and the internal fluxes and the action on the latter of dualities. We end the section by discussing, as an example, the gauging of the maximal four dimensional theory.

In Sect. 4 we review a manifestly covariant formulation of the gauging procedure and introduce the notion of tensor hierarchy in higher dimensions.

2 Review of Ungauged Supergravities

Let us recall some basic aspects of the extended ungauged $D = 4$ supergravity.

Field content and bosonic action. The bosonic sector consists in the graviton $g_{\mu\nu}(x)$, n_v vector fields $A^\Lambda_\mu(x)$, n_s scalar fields $\phi^s(x)$ and is described by bosonic Lagrangian of the following general form[2]

$$\frac{1}{e}\mathscr{L}_{\rm B} = -\frac{R}{2} + \frac{1}{2}\,\mathcal{G}_{st}(\phi)\,\partial_\mu\phi^s\,\partial^\mu\phi^t + \frac{1}{4}\,\mathcal{I}_{\Lambda\Sigma}(\phi)\,F^\Lambda_{\mu\nu}\,F^{\Sigma\,\mu\nu} + \frac{1}{8\,e}\,\mathcal{R}_{\Lambda\Sigma}(\phi)\,\varepsilon^{\mu\nu\rho\sigma}\,F^\Lambda_{\mu\nu}\,F^\Sigma_{\rho\sigma}, \tag{5}$$

where $e = \sqrt{|\mathrm{Det}(g_{\mu\nu})|}$ and the n_v vector field strengths are defined as usual:

$$F^\Lambda_{\mu\nu} = \partial_\mu A^\Lambda_\nu - \partial_\nu A^\Lambda_\mu\ . \tag{6}$$

Let us comment on the general characteristics of the above action.

- The scalar fields ϕ^s are described by a non-linear σ-model, that is they are coordinates of a non-compact, *Riemannian* n_s*-dimensional* differentiable manifold (target space), named *scalar manifold* and to be denoted by $\mathscr{M}_{\rm scal}$. The positive definite metric on the manifold is $\mathcal{G}_{st}(\phi)$. The corresponding kinetic part of the Lagrangian density reads:

$$\mathscr{L}_{\rm scal} = \frac{e}{2}\,\mathcal{G}_{st}(\phi)\,\partial_\mu\phi^s\partial^\mu\phi^t. \tag{7}$$

  The σ-model action is clearly invariant under the action of global (i.e. space-time independent) isometries of the scalar manifold. As we shall discuss below, the group G can be promoted to a global symmetry group of the field equations and Bianchi identities (i.e. *on-shell global symmetry group*) provided its (non-linear) action on the scalar fields is combined with an electric-magnetic duality transformation on the vector field strengths and their magnetic duals.
- The two terms containing the vector field strengths will be called vector kinetic terms. A general feature of supergravity theories is that the scalar fields are non-minimally coupled to the vector fields as they enter these terms through symmetric

[2] Using the "mostly minus" convention and $8\pi\mathrm{G_N} = c = \hbar = 1$.

matrices $\mathcal{I}_{\Lambda\Sigma}(\phi)$, $\mathcal{R}_{\Lambda\Sigma}(\phi)$ which contract the vector field strengths. The former $\mathcal{I}_{\Lambda\Sigma}(\phi)$ is negative definite and generalizes the $-1/g^2$ factor in the Yang–Mills kinetic term. The latter $\mathcal{R}_{\Lambda\Sigma}(\phi)$ generalizes the θ-term.

- There is a $U(1)^{n_v}$ gauge invariance associated with the vector fields:

$$A^\Lambda_\mu \to A^\Lambda_\mu + \partial_\mu \zeta^\Lambda. \tag{8}$$

  All the fields are neutral with respect to this symmetry group.
- There is no scalar potential. In an ungauged supergravity a scalar potential is allowed only for $N = 1$ (called the *F-term potential*). In extended supergravities a non-trivial scalar potential can be introduced without explicitly breaking supersymmetry only through the *gauging procedure*, which implies the introduction of a local symmetry group to be gauged by the vector fields of the theory and which will be extensively dealt with in the following.

The fermion part of the action is totally determined by supersymmetry once the bosonic one is given. Let us discuss in some detail the scalar sector and its mathematical description.

2.1 Scalar Sector and Coset Geometry

As mentioned above the scalar fields ϕ^s are coordinates of a Riemannian scalar manifold $\mathscr{M}_{\rm scal}$, with metric $\mathcal{G}_{st}(\phi)$. The isotropy group H of $\mathscr{M}_{\rm scal}$ has the general form

$$H = H_{\rm R} \times H_{\rm matt}, \tag{9}$$

where $H_{\rm R}$ is the R–symmetry group and $H_{\rm matt}$ is a compact group acting on the matter fields. The gravitino and spin-$\frac{1}{2}$ fields will transform in representations of the H group. The maximal theory $N = 8$ describes the gravitational multiplet only and thus $H = H_{\rm R} = {\rm SU}(8)$. The isometry group G of $\mathscr{M}_{\rm scal}$ clearly defines the global symmetries of the scalar action.

In $N > 2$ theories the scalar manifold is constrained by supersymmetry to be homogeneous symmetric, namely to have the general form

$$\mathscr{M}_{\rm scal} = \frac{G}{H}, \tag{10}$$

where G is the semisimple non-compact Lie group of isometries and H its maximal compact subgroup (Table 1). Generic homogeneous spaces $\mathscr{M}_{\rm scal}$ can always be written in the above form though G need not be semisimple. The action of an isometry transformation $g \in G$ on the scalar fields ϕ^r parametrizing $\mathscr{M}_{\rm scal}$ is defined by means of a *coset representative* $L(\phi) \in G/H$ as follows:

$$g \cdot L(\phi^r) = L(g \star \phi^r) \cdot h(\phi^r, g), \tag{11}$$

Table 1 Examples of homogeneous symmetric scalar manifolds in extended supergravities and their real dimensions n_s. We have omitted in the list the homogeneous symmetric quaternionic Kaehler manifolds in the N=2 models

N	$\frac{G}{H}$	n_s
8	$\frac{E_{7(7)}}{SU(8)}$	70
6	$\frac{SO^*(12)}{U(6)}$	30
5	$\frac{SU(5,1)}{U(5)}$	10
4	$\frac{SL(2,\mathbb{R})}{SO(2)} \times \frac{SO(6,n)}{SO(6)\times SO(n)}$	6n + 2
3	$\frac{SU(3,n)}{S[U(3)\times U(n)]}$	6n
2	$\frac{SU(1,n+1)}{U(n+1)}$	2(n + 1)
	$\frac{SL(2,\mathbb{R})}{SO(2)} \times \frac{SO(2,n+2)}{SO(2)\times SO(n+2)}$	2(n + 2)+2
	$\frac{Sp(6)}{U(3)}$	12
	$\frac{SU(3,3)}{S[U(3)\times U(3)]}$	18
	$\frac{SO^*(12)}{U(6)}$	30
	$\frac{E_{7(-25)}}{U(1)\times E_6}$	54

where $g \star \phi^r$ denote the transformed scalar fields, non-linear functions of the original ones ϕ^r, and $h(\phi^r, g)$ is a *compensator* in H. The coset representative is defined modulo the right-action of H and is fixed by the chosen parametrization of the manifold. Of particular relevance in supergravity is the so-called *solvable parametrization*, which corresponds to fixing the action of H so that L belongs to a solvable Lie group[3] $G_S = \exp(\mathscr{S})$, generated by a solvable Lie algebra $\mathscr{S}$ and defined, in the symmetric case, by the Iwasawa decomposition of G with respect to H. The scalar fields are then parameters of the solvable Lie algebra $\mathscr{S}$:

$$L(\phi^r) = e^{\phi^r T_r} \in \exp(\mathscr{S}), \tag{12}$$

where $\{T_r\}$ is a basis of $\mathscr{S}$ $(r = 1, \ldots, n_s)$. All homogeneous scalar manifolds occurring in supergravity theories admit this parametrization, which is useful when the four-dimensional supergravity originates from the Kaluza–Klein reduction of a higher-dimensional one on some internal compact manifold. The solvable coordinates directly describe dimensionally reduced fields and moreover this parametrization makes the shift symmetries of the metric manifest.

The Lie algebra $\mathfrak{g}$ of G can be decomposed into the Lie algebra $\mathfrak{H}$ generating H, and a coset space $\mathfrak{K}$:

[3] A solvable Lie group G_S can be described (locally) as a the Lie group generated by *solvable Lie algebra* $\mathscr{S}$: $G_S = \exp(\mathscr{S})$. A Lie algebra $\mathscr{S}$ is solvable iff, for some $k > 0$, $\mathbf{D}^k\mathscr{S} = 0$, where the *derivative* $\mathbf{D}$ of a Lie algebra $\mathfrak{g}$ is defined as follows: $\mathbf{D}\mathfrak{g} \equiv [\mathfrak{g}, \mathfrak{g}]$, $\mathbf{D}^n\mathfrak{g} \equiv [\mathbf{D}^{n-1}\mathfrak{g}, \mathbf{D}^{n-1}\mathfrak{g}]$. In a suitable basis of a given representation, elements of a solvable Lie group or a solvable Lie algebra are all described by upper (or lower) triangular matrices.

$$\mathfrak{g} = \mathfrak{H} \oplus \mathfrak{K}, \tag{13}$$

where in general we have:

$$[\mathfrak{H},\, \mathfrak{H}] \subset \mathfrak{H}\,; \qquad [\mathfrak{H},\, \mathfrak{K}] \subset \mathfrak{K}\,; \qquad [\mathfrak{K},\, \mathfrak{K}] \subset \mathfrak{H} \oplus \mathfrak{K}, \tag{14}$$

that is the space $\mathfrak{K}$ supports a representation $\mathcal{K}$ of H with respect to its adjoint action. An alternative choice of parametrization corresponds to defining the coset representative as an element of $\exp(\mathfrak{K})$:

$$L(\phi^r) = e^{\phi^r K_r} \ \in \ \exp(\mathfrak{K}), \tag{15}$$

where $\{K_r\}$ is a basis of $\mathfrak{K}$. As opposed to the solvable parametrization, the coset representative is no-longer a group element, since $\mathfrak{K}$ does not close an algebra, see last of (14). The main advantage of this parametrization is that the action of H on the scalar fields is *linear*:

$$\forall h \in H \ : \quad h\, L(\phi^r) = h\, e^{\phi^r K_r}\, h^{-1}\, h = e^{\phi^r\, h K_r\, h^{-1}}\, h = L(\phi'^r)\, h, \tag{16}$$

where $\phi'^r = (h^{-1})_s{}^r\, \phi^s$, and $h_s{}^r$ describes h in the representation $\mathcal{K}$. This is not the case for the solvable parametrization since $[\mathfrak{H},\, \mathscr{S}] \nsubseteq \mathscr{S}$.

In all parametrizations, the origin $\mathcal{O}$ is defined as the point in which the coset representative equals the identity element of G and thus the H-invariance of $\mathcal{O}$ is manifest: $L(\mathcal{O}) = \mathbf{I}$.

If the manifold, besides being homogeneous, is also *symmetric*, the space $\mathfrak{K}$ can be defined so that:

$$[\mathfrak{K},\, \mathfrak{K}] \subset \mathfrak{H}. \tag{17}$$

In this case the (13) defines the Cartan decomposition of $\mathfrak{g}$ into *compact* and *non-compact* generators, in $\mathfrak{H}$ and $\mathfrak{K}$, respectively. This means that, in a given matrix representation of $\mathfrak{g}$, a basis of the carrier vector space can be chosen so that the elements of $\mathfrak{H}$ and of $\mathfrak{K}$ are represented by anti-hermitian and hermitian matrices, respectively.

The geometry of $\mathscr{M}_{\text{scal}}$ is described by vielbein and an H-connection constructed out of the left-invariant one-form

$$\Omega \ = \ L^{-1}\, dL \ \in \ \mathfrak{g}, \tag{18}$$

satisfying the Maurer-Cartan equation:

$$d\Omega + \Omega \wedge \Omega = 0. \tag{19}$$

The Vielbein and H-connection are defined by decomposing Ω according to (13)

$$\Omega(\phi) = \Phi(\mathcal{P}) + w(\phi); \qquad w \in \mathfrak{H}, \quad \mathcal{P} \in \mathfrak{K}. \tag{20}$$

Let us see how these quantities transform under the action of G. For any $g \in G$, using (11), we can write $L(g \star \phi) = g\, L(\phi)\, h^{-1}$, so that:

$$\Omega(g \star \phi) = h\, L(\phi)^{-1}\, g^{-1} d(g\, L(\phi)\, h^{-1}) = h\, L(\phi)^{-1}\, dL(\phi)\; h^{-1} + h\; dh^{-1}. \quad (21)$$

From (20) we find:

$$\mathcal{P}(g \star \phi) + w(g \star \phi) = h\, \mathcal{P}(\phi)\, h^{-1} + h\, w(\phi) h^{-1} + h\; dh^{-1}. \quad (22)$$

Since $h\; dh^{-1}$ is the left-invariant 1-form on $\mathfrak{H}$, it has value in this algebra. Projecting the above equation over $\mathfrak{K}$ and $\mathfrak{H}$, we find:

$$\mathcal{P}(g \star \phi) = h\, \mathcal{P}(\phi)\, h^{-1}, \quad (23)$$
$$w(g \star \phi) = h\, w(\phi)\, h^{-1} + h\; dh^{-1}. \quad (24)$$

We see that w transforms as an H-connection while the matrix-valued one-form $\mathcal{P}$ transforms linearly under H. The vielbein of the scalar manifold are defined by expanding $\mathcal{P}$ in a basis $\{K_{\underline{s}}\}$ of $\mathfrak{K}$ (underlined indices $\underline{s}, \underline{r}, \underline{t}, \ldots$ are rigid tangent-space indices, as opposed to the curved coordinate indices $s, r, t, \ldots$):

$$\mathcal{P}(\phi) = V^{\underline{s}}(\phi)\, K_{\underline{s}}. \quad (25)$$

From (23) it follows that the vielbein 1-forms $V^{\underline{s}}(\phi) = V_s{}^{\underline{s}}(\phi) d\phi^s$ transform under the action of G as follows:

$$V^{\underline{s}}(g \star \phi) \;=\; V^{\underline{t}}(\phi)\, (h^{-1})_{\underline{t}}{}^{\underline{s}} \;=\; h^{\underline{s}}{}_{\underline{t}} V^{\underline{t}}(\phi). \quad (26)$$

For symmetric spaces, from (19) it follows that w and $\mathcal{P}$ satisfy the following conditions

$$\mathscr{D}\mathcal{P} \;\equiv\; d\mathcal{P} + w \wedge \mathcal{P} + \mathcal{P} \wedge w \;=\; 0, \quad (27)$$
$$R(w) \;\equiv\; dw + w \wedge w \;=\; -\mathcal{P} \wedge \mathcal{P}, \quad (28)$$

where we have defined the H-covariant derivative $\mathscr{D}\mathcal{P}$ of $\mathcal{P}$ and the $\mathfrak{H}$-valued curvature $R(w)$ of the manifold. The latter can be written in components:

$$R(w) = \frac{1}{2} R_{rs}\, d\phi^r \wedge d\phi^s \quad \Rightarrow \quad R_{rs} = -[\mathcal{P}_r,\, \mathcal{P}_s] \in \mathfrak{H}. \quad (29)$$

We define the metric at the origin $\mathcal{O}$ as the H-invariant matrix:

$$\eta_{\underline{st}} \equiv k\, \mathrm{Tr}(K_{\underline{s}}\, K_{\underline{t}}) > 0, \quad (30)$$

where k is a positive number depending on the representation, so that the metric in a generic point reads:

$$ds^2(\phi) \equiv \mathcal{G}_{st}(\phi)d\phi^s\, d\phi^t \equiv V_s{}^{\underline{s}}(\phi)V_t{}^{\underline{t}}(\phi)\eta_{\underline{st}}\, d\phi^s\, d\phi^t = k\,\mathrm{Tr}(\mathcal{P}_s\,\mathcal{P}_t). \tag{31}$$

As it follows from (23), (26), the above metric is manifestly invariant under global G-transformations acting on L to the left (as well as local H-transformations acting on L to the right):

$$ds^2(g\star\phi) = ds^2(\phi)\ . \tag{32}$$

The σ-model Lagrangian can be written in the form:

$$\mathscr{L}_{\mathrm{scal}} = \frac{e}{2}\,\mathcal{G}(\phi)_{st}\partial_\mu\phi^s\,\partial^\mu\phi^t = \frac{e}{2}\,k\,\mathrm{Tr}\big(\mathcal{P}_\mu(\phi)\,\mathcal{P}^\mu(\phi)\big), \qquad \mathcal{P}_\mu = \mathcal{P}_s\frac{\partial\phi^s}{\partial x^\mu}, \tag{33}$$

and, just as the metric ds^2, is manifestly invariant under global G and local H-transformations acting on L as in (11).

The bosonic part of the equations of motion for the scalar fields can be derived from the Lagrangian (5) and read:

$$\mathscr{D}_\mu(\partial^\mu\phi^s) = \frac{1}{4}\,\mathcal{G}^{st}\,\left[F^\Lambda_{\mu\nu}\,\partial_t\,\mathcal{I}_{\Lambda\Sigma}\,F^{\Sigma\,\mu\nu} + F^\Lambda_{\mu\nu}\partial_t\,\mathcal{R}_{\Lambda\Sigma}\,{}^*F^{\Sigma\,\mu\nu}\right], \tag{34}$$

where $\partial_s \equiv \frac{\partial}{\partial\phi^s}$, while $\mathscr{D}_\mu$ also contains the Levi-Civita connection $\tilde{\Gamma}$ on the scalar manifold:

$$\mathscr{D}_\mu(\partial_\nu\phi^s) \equiv \nabla_\mu(\partial_\nu\phi^s) + \tilde{\Gamma}^s_{t_1t_2}\partial_\mu\phi^{t_1}\,\partial_\nu\phi^{t_2}, \tag{35}$$

∇_μ being the covariant derivative containing the Levi-Civita connection on space-time.

Let us end this paragraph by introducing, in the coset geometry, the Killing vectors describing the infinitesimal action of isometries on the scalar fields. Let us denote by t_α the infinitesimal generators of G, defining a basis of its Lie algebra $\mathfrak{g}$ and satisfying the corresponding commutation relations

$$[t_\alpha,\, t_\beta] = \mathbf{f}_{\alpha\beta}{}^\gamma\, t_\gamma, \tag{36}$$

$\mathbf{f}_{\alpha\beta}{}^\gamma$ being the structure constants of $\mathfrak{g}$. Under an infinitesimal G-transformation generated by $\epsilon^\alpha\, t_\alpha$ ($\epsilon^\alpha \ll 1$):

$$g \approx \mathbf{I} + \epsilon^\alpha\, t_\alpha, \tag{37}$$

the scalars transform as:

$$\phi^s \to \phi^s + \epsilon^\alpha\, k^s_\alpha(\phi), \tag{38}$$

$k^s_\alpha(\phi)$ being the Killing vector associated with t_α. The action of g on the scalars is defined by (11), neglecting terms of order $O(\epsilon^2)$:

$$(\mathbf{I} + \varepsilon^\alpha\, t_\alpha)\, L(\phi) = L(\phi + \varepsilon^\alpha\, k_\alpha)\left(\mathbf{I} - \frac{1}{2}\,\varepsilon^\alpha W^I_\alpha\, J_I\right), \tag{39}$$

where $(\mathbf{I} - \frac{1}{2}\,\varepsilon^\alpha W^I_\alpha\, J_I)$ denotes, expanded to linear order in ε, the compensating transformation $h(\phi, g)$, $\{J_I\}$ being a basis of $\mathfrak{H}$. Equating the terms proportional to ε^α, multiplying to the left by L^{-1} and using the expansion (20) of the left-invariant 1-form, we end up with the following equation:

$$L^{-1} t_\alpha L \;=\; k^s_\alpha\,(\mathcal{P}_s + w_s) - \frac{1}{2}\, W^I_\alpha\, J_I \;=\; k^s_\alpha\, V_s{}^{\underline{s}}\, K_{\underline{s}} + \frac{1}{2}\,(k^s_\alpha \omega^I_s - W^I_\alpha)\, J_I, \tag{40}$$

where we have expanded the H-connection along J_I as follows:

$$w_s = \frac{1}{2}\,\omega^I_s\, J_I. \tag{41}$$

Equation (40) allows to compute k_α for homogeneous scalar manifolds by projecting $L^{-1} t_\alpha L$ along the directions of the coset space $\mathfrak{K}$. These Killing vectors satisfy the following algebraic relations (note the minus sign on the right hand side with respect to (36):

$$[k_\alpha,\, k_\beta] = -\mathbf{f}_{\alpha\beta}{}^\gamma\, k_\gamma, \tag{42}$$

We can split, according to the general structure (9), the H-generators J_I into $H_{\rm R}$-generators $J_{\mathbf{a}}$ $(\mathbf{a} = 1, \ldots, \dim(H_{\rm R}))$ and $H_{\rm matt}$-generators $J_{\mathbf{m}}$ $(\mathbf{m} = 1, \ldots, \dim(H_{\rm matt}))$, and rewrite (40) in the form:

$$L^{-1} t_\alpha L \;=\; k^s_\alpha\, V_s{}^{\underline{s}}\, K_{\underline{s}} - \frac{1}{2}\,\mathscr{P}^{\mathbf{a}}_\alpha\, J_{\mathbf{a}} - \frac{1}{2}\,\mathscr{P}^{\mathbf{m}}_\alpha\, J_{\mathbf{m}}. \tag{43}$$

The quantities

$$\mathscr{P}^{\mathbf{a}}_\alpha = -(k^s_\alpha \omega^{\mathbf{a}}_s - W^{\mathbf{a}}_\alpha), \tag{44}$$

generalize the so called *momentum maps* in N = 2 theories, which provide a Poissonian realization of the isometries t_α. One can verify the general property:

$$k^s_\alpha\, R^{\mathbf{a}}_{st} = \mathscr{D}_t \mathscr{P}^{\mathbf{a}}_\alpha, \tag{45}$$

where $\mathscr{D}_s$ denotes the H-covariant derivative and we have expanded the curvature $R[w]$ defined in (28) along J_I:

$$R[w] = \frac{1}{2}\, R^I_{st}\, d\phi^s \wedge d\phi^t\, J_I. \tag{46}$$

These objects are important in the gauging procedure since they enter the definition of the gauged connections for the fermion fields as well as gravitino-shift matrix $\mathbb{S}_{AB}$ (see Sect. 3). For all those isometries which do not produce compensating transformations in $H_{\rm R}$, $W_\alpha^{\bf a}=0$ and $\mathscr{P}_\alpha^{\bf a}$ are easily computed to be

$$\mathscr{P}_\alpha^{\bf a}=-k_\alpha^s\omega_s^{\bf a}.$$

This is the case, in the solvable parametrization, for all the isometries in $\mathscr{S}$, which include translations in the axionic fields.

In N = 2 models with non-homogeneous scalar geometries, though we cannot apply the above construction of k_α, $\mathscr{P}_\alpha^{\bf a}$, the momentum maps are constructed from the Killing vectors as solutions to the differential equations (45). In general, in these theories, with each isometry t_α of the scalar manifold, we can associate the quantities $\mathscr{P}_\alpha^{\bf a}$, $\mathscr{P}_\alpha^{\bf m}$ which are related to the corresponding Killing vectors k_α through general relations (see [19] for a comprehensive account of N = 2 theories).

2.2 Vector Sector

We can associate with the electric field strengths $F^\Lambda_{\mu\nu}$ their magnetic duals $\mathcal{G}_{\Lambda\,\mu\nu}$ defined as:

$$\mathcal{G}_{\Lambda\,\mu\nu}\equiv-\varepsilon_{\mu\nu\rho\sigma}\frac{\partial\mathscr{L}_4}{\partial F^\Lambda_{\rho\sigma}}=\mathcal{R}_{\Lambda\Sigma}\,F^\Sigma_{\mu\nu}-\mathcal{I}_{\Lambda\Sigma}\,{}^*F^\Sigma_{\mu\nu}, \tag{47}$$

where we have omitted fermion currents in the expression of $\mathcal{G}_\Lambda$ since we are only focussing for the time being on the bosonic sector of the theory. In ordinary Maxwell theory (no scalar fields), $\mathcal{I}_{\Lambda\Sigma}=-\delta_{\Lambda\Sigma}$ and $\mathcal{R}_{\Lambda\Sigma}=0$, so that $\mathcal{G}_{\Lambda\,\mu\nu}$ coincides with the Hodge-dual of $F^\Lambda_{\mu\nu}$: $\mathcal{G}_\Lambda={}^*F^\Lambda$.

In terms of F^Λ and $\mathcal{G}_\Lambda$ the bosonic part of the Maxwell equations read

$$\nabla^\mu({}^*F^\Lambda_{\mu\nu})=0;\qquad\nabla^\mu({}^*\mathcal{G}_{\Lambda\,\mu\nu})=0, \tag{48}$$

In order to set the stage for the discussion of global symmetries, it is useful to rewrite the scalar and vector field equations in a different form. Using (47) and the property that ${}^{**}F^\Lambda=-F^\Lambda$, we can express ${}^*F^\Lambda$ and ${}^*\mathcal{G}_\Lambda$ as linear functions of F^Λ and $\mathcal{G}_\Lambda$:

$${}^*F^\Lambda=\mathcal{I}^{-1\,\Lambda\Sigma}\,(\mathcal{R}_{\Sigma\Gamma}\,F^\Gamma-\mathcal{G}_\Sigma)\ ; \tag{49}$$

$${}^*\mathcal{G}_\Lambda=(\mathcal{R}\mathcal{I}^{-1}\mathcal{R}+\mathcal{I})_{\Lambda\Sigma}\,F^\Sigma-(\mathcal{R}\mathcal{I}^{-1})_\Lambda{}^\Sigma\,\mathcal{G}_\Sigma, \tag{50}$$

where, for the sake of simplicity, we have omitted the space-time indices. It is useful to arrange F^Λ and $\mathcal{G}_\Lambda$ in a single $2n_v$-dimensional vector $\mathbb{F}\equiv(\mathbb{F}^M)$ of two-forms:

$$\mathbb{F}=\left(\frac{1}{2}\,\mathbb{F}^M_{\mu\nu}\,dx^\mu\wedge dx^\nu\right)\equiv\begin{pmatrix}F^\Lambda_{\mu\nu}\\ \mathcal{G}_{\Lambda\mu\nu}\end{pmatrix}\frac{dx^\mu\wedge dx^\nu}{2}, \tag{51}$$

in terms of which the Maxwell equations read:

$$d\mathbb{F} = 0, \tag{52}$$

and (50) are easily rewritten in the following compact form:

$$^*\mathbb{F} = -\mathbb{C}\mathcal{M}(\phi^s)\,\mathbb{F}, \tag{53}$$

where

$$\mathbb{C} = (\mathbb{C}^{MN}) \equiv \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix}, \tag{54}$$

I, **0** being the $n_v \times n_v$ identity and zero-matrices, respectively, and

$$\mathcal{M}(\phi) = (\mathcal{M}(\phi)_{MN}) \equiv \begin{pmatrix} (\mathcal{R}\mathcal{I}^{-1}\mathcal{R} + \mathcal{I})_{\Lambda\Sigma} & -(\mathcal{R}\mathcal{I}^{-1})_{\Lambda}{}^{\Gamma} \\ -(\mathcal{I}^{-1}\mathcal{R})^{\Delta}{}_{\Sigma} & \mathcal{I}^{-1\,\Delta\Gamma} \end{pmatrix}, \tag{55}$$

is a symmetric, negative-definite matrix, function of the scalar fields. The reader can easily verify that this matrix is also symplectic, namely that:

$$\mathcal{M}(\phi)\mathbb{C}\mathcal{M}(\phi) = \mathbb{C}. \tag{56}$$

This matrix contains $\mathcal{I}_{\Lambda\Sigma}$ and $\mathcal{R}_{\Lambda\Sigma}$ as components, and therefore defines the non-minimal coupling of the scalars to the vector fields.

After some algebra, we can also rewrite (34) in a compact form as follows

$$\mathscr{D}_\mu(\partial^\mu\phi^s) = \frac{1}{8}\,\mathcal{G}^{st}\,\mathbb{F}^T_{\mu\nu}\partial_t\mathcal{M}(\phi)\,\mathbb{F}^{\mu\nu}, \tag{57}$$

2.3 Coupling to Gravity

We can now compute the Einstein equations:

$$R_{\mu\nu} - \frac{1}{2}\,g_{\mu\nu}\,R = T^{(S)}_{\mu\nu} + T^{(V)}_{\mu\nu} + T^{(F)}_{\mu\nu}, \tag{58}$$

where the three terms on the right hand side are the energy-momentum tensors of the scalars, vectors and fermionic fields, respectively. The first two can be cast in the following general form

$$T^{(S)}_{\mu\nu} = \mathcal{G}_{rs}(\phi)\,\partial_\mu\phi^r\partial_\nu\phi^s - \frac{1}{2}\,g_{\mu\nu}\,\mathcal{G}_{rs}(\phi)\,\partial_\rho\phi^r\partial^\rho\phi^s, \tag{59}$$

$$T^{(V)}_{\mu\nu} = \left(F^T_{\mu\rho}\,\mathcal{I}\,F_\nu{}^\rho - \frac{1}{4}\,g_{\mu\nu}\,(F^T_{\rho\sigma}\mathcal{I}F^{\rho\sigma})\right), \tag{60}$$

where in the last equation the vector indices Λ, Σ have been suppressed for the sake of notational simplicity. It is convenient for our next discussion, to rewrite, after some algebra, the right hand side of (60) as follows

$$T^{(V)}_{\mu\nu} = \frac{1}{2}\,\mathbb{F}^T_{\mu\rho}\,\mathcal{M}(\phi)\,\mathbb{F}_\nu{}^\rho\,, \tag{61}$$

so that (58) can be finally recast in the following form:

$$R_{\mu\nu} = \mathcal{G}_{rs}(\phi)\,\partial_\mu\phi^r\partial_\nu\phi^s + \frac{1}{2}\,\mathbb{F}^T_{\mu\rho}\,\mathcal{M}(\phi)\,\mathbb{F}_\nu{}^\rho + \cdots\,, \tag{62}$$

where the ellipses refer to fermionic terms.

The scalar fields enter the kinetic terms of the vector fields through the matrices $\mathcal{I}(\phi)$ and $\mathcal{R}(\phi)$. As a consequence of this, a symmetry transformation of the scalar part of the Lagrangian will not in general leave the vector field part invariant.

2.4 Global Symmetry Group

In extended supergravity models (N $>$ 1) the (identity sector of the) global symmetry group G of the scalar action can be promoted to a global invariance [20] of, at least, the field equations and the Bianchi identities, provided its (non-linear) action on the scalar fields is associated with a linear transformation on the vector field strengths $F^\Lambda_{\mu\nu}$ and their magnetic duals $\mathcal{G}_{\Lambda\,\mu\nu}$:

$$g \in G : \begin{cases} \phi^r \;\rightarrow\; g\star\phi^r & \text{(non–linear)},\\ \begin{pmatrix} F^\Lambda \\ \mathcal{G}_\Lambda \end{pmatrix} \;\rightarrow\; \mathscr{R}_v[g]\cdot\begin{pmatrix} F^\Lambda \\ \mathcal{G}_\Lambda \end{pmatrix} = \begin{pmatrix} A[g]^\Lambda{}_\Sigma & B[g]^{\Lambda\Sigma} \\ C[g]_{\Lambda\Sigma} & D[g]_\Lambda{}^\Sigma \end{pmatrix}\begin{pmatrix} F^\Sigma \\ \mathcal{G}_\Sigma \end{pmatrix} & \text{(linear)}. \end{cases} \tag{63}$$

The transformations (63) are clearly a symmetry of the scalar action and of the Maxwell equations ($d\mathbb{F} = 0$) if F^Λ and $\mathcal{G}_\Lambda$ were independent, since the latter are clearly invariant with respect to any linear transformation on $\mathbb{F}^M$. The definition $\mathcal{G}_\Lambda$ in (47) as a function of F^Λ, ${}^*F^\Lambda$ and the scalar fields, which is equivalently expressed by the twisted self-duality condition (53), however poses constraints on the $2n_v \times 2n_v$ matrix $\mathscr{R}_v[g] = (\mathscr{R}_v[g]^M{}_N)$. In order for (63) to be an invariance of the vector equations of motion (52) and (53) the following conditions have to be met:

(i) for each $g \in G$ (more precisely in the identity sector of G), the matrix $\mathscr{R}_v[g]$ should be *symplectic*, namely

$$\mathscr{R}_v[g]^T\,\mathbb{C}\,\mathscr{R}_v[g] = \mathbb{C}; \tag{64}$$

(ii) the symplectic, scalar dependent, matrix $\mathcal{M}(\phi)$ should transform as follows:

$$\mathcal{M}(g \star \phi) = \mathscr{R}_v[g]^{-T} \mathcal{M}(\phi)\, \mathscr{R}_v[g]^{-1}, \tag{65}$$

where we have used the short-hand notation $\mathscr{R}_v[g]^{-T} \equiv (\mathscr{R}_v[g]^{-1})^T$.

The reader can indeed verify that conditions (i) and (ii) are sufficient to guarantee invariance of (53) under (63). The symplectic transformation $\mathscr{R}_v[g]$, associated with each element g of G, mixes electric and magnetic field strengths, acting therefore as a generalized electric–magnetic duality and defines a *symplectic representation* $\mathscr{R}_v$ of G:

$$\forall g \in G \quad \stackrel{\mathscr{R}_v}{\longrightarrow} \quad \mathscr{R}_v[g] \in \mathrm{Sp}(2n_v,\, \mathbb{R}). \tag{66}$$

The field strengths and their magnetic duals transform therefore, under the duality action (63) of G in a $2n_v$-dimensional symplectic representation.

We denote by $\mathscr{R}_{v*} = \mathscr{R}_v^{-T}$ the representation dual to $\mathscr{R}_v$, acting on covariant symplectic vectors, so that, for any $\mathbf{g} \in G$:

$$\begin{aligned} \mathscr{R}_{v*}[\mathbf{g}] = (\mathscr{R}_{v*}[\mathbf{g}]_M{}^N) = \mathscr{R}_v[\mathbf{g}]^{-T} = -\mathbb{C}\mathscr{R}_v[\mathbf{g}]\mathbb{C} \\ \Rightarrow \quad \mathscr{R}_{v*}[\mathbf{g}]_M{}^N = \mathbb{C}_{MP}\, \mathscr{R}_v[\mathbf{g}]^P{}_Q\, \mathbb{C}^{NQ}, \end{aligned} \tag{67}$$

where we have used the property that $\mathscr{R}_v$ is a symplectic representation.[4]

From (64) and (65), it is straightforward to verify the manifest G-invariance of the scalar field equations and the Einstein equations written in the forms (57) and (62).

Conditions (i) and (ii) are verified in extended supergravities as a consequence of supersymmetry. In these theories indeed supersymmetry is large enough as to connect certain scalar fields to vector fields and, as a consequence of this, symmetry transformations on the former imply transformations on the latter (more precisely transformations on the vector field strengths F^Λ and their duals $\mathcal{G}_\Lambda$). The existence of a symplectic representation $\mathscr{R}_v$ of G, together with the definition of the matrix $\mathcal{M}$ and its transformation property (65), are built-in in the mathematical structure of the scalar manifold. More precisely they follow from the definition on $\mathscr{M}_{\mathrm{scal}}$ of a *flat symplectic structure*. Supersymmetry totally fixes $\mathcal{M}(\phi)$ and thus the coupling of the scalar fields to the vectors, aside from a freedom in the choice of the basis of the symplectic representation (*symplectic frame*) which amounts to a change in the definition of $\mathcal{M}(\phi)$ by a constant symplectic transformation E:

$$\mathcal{M}(\phi) \to \mathcal{M}'(\phi) = E\mathcal{M}(\phi)E^T. \tag{68}$$

Clearly if $E \in \mathscr{R}_{v*}[G] \subset \mathrm{Sp}(2n_v, \mathbb{R})$, its effect on $\mathcal{M}(\phi)$ can be offset be a redefinition of the scalar fields, by virtue of (65). On the other hand if E a were block-diagonal matrix, namely an element of $\mathrm{GL}(n_v, \mathbb{R}) \subset \mathrm{Sp}(2n_v, \mathbb{R})$, it could be reabsorbed in a local redefinition of the field strengths. Inequivalent symplectic frames are then con-

[4] The symplectic indices $M,\ N, \ldots$ are raised (and lowered) with the symplectic matrix $\mathbb{C}^{MN}$ ($\mathbb{C}_{MN}$) using north-west south-east conventions: $X^M = \mathbb{C}^{MN} X_N$ (and $X_M = \mathbb{C}_{NM} X^N$).

nected by symplectic matrices E defined modulo redefinitions of the scalar and vector fields, namely by matrices in the coset [5]:

$$E \in \mathrm{GL}(n_v, \mathbb{R})\backslash \mathrm{Sp}(2n_v, \mathbb{R})/\mathscr{R}_{v*}[G], \tag{69}$$

where the quotient is defined with respect to the left-action of $\mathrm{GL}(n_v, \mathbb{R})$ (local vector redefinitions) and to the right-action of $\mathscr{R}_{v*}[G]$ (isometry action on the scalar fields).

A change in the symplectic frame amounts to choosing a different embedding $\mathscr{R}_v$ of G inside $\mathrm{Sp}(2n_v, \mathbb{R})$, which is not unique. This affects the form of the action, in particular the coupling of the scalar fields to the vectors. However, at the ungauged level, it only amounts to a redefinition of the vector field strengths and their duals which has no physical implication. In the presence of a gauging, namely if vectors are minimally coupled to the other fields, the symplectic frame becomes physically relevant and may lead to different vacuum-structures of the scalar potential.

We emphasize here that the existence of this symplectic structure on the scalar manifold is a general feature of all extended supergravites, including those $\mathrm{N} = 2$ models in which the scalar manifold is not even homogeneous (i.e. the isometry group, if it exists, does not act transitively on the manifold itself). In the $\mathrm{N} = 2$ case, only the scalar fields belonging to the vector multiplets are non-minimally coupled to the vector fields, namely enter the matrices $\mathcal{I}(\phi)$, $\mathcal{R}(\phi)$, and they span a *special Kähler* manifold. On this manifold a flat symplectic bundle is defined,[5] which fixes the scalar dependence of the matrices $\mathcal{I}(\phi)$, $\mathcal{R}(\phi)$, aside from an initial choice of the symplectic frame, and the matrix $\mathcal{M}(\phi)$ defined in (55) satisfies the property (65).

If the scalar manifold is homogeneous, we can consider at any point the coset representative $L(\phi) \in G$ in the symplectic, $2n_v$-dimensional representation $\mathscr{R}_v$:

$$L(\phi) \stackrel{\mathscr{R}_v}{\longrightarrow} \mathscr{R}_v[L(\phi)] \in \mathrm{Sp}(2n_v, \mathbb{R}). \tag{70}$$

In general the representation $\mathscr{R}_v[H]$ of the isotropy group H may not be orthogonal, that is $\mathscr{R}_v[H] \nsubseteq \mathrm{SO}(2n_v)$. In this case we can always change the basis of the representation[6] by means of a matrix $\mathcal{S}$

$$\mathcal{S} = (\mathcal{S}^N{}_{\underline{M}}) \in \mathrm{Sp}(2n_v, \mathbb{R})/\mathrm{U}(n) \tag{71}$$

such that, in the rotated representation $\underline{\mathscr{R}}_v \equiv \mathcal{S}^{-1}\mathscr{R}_v\,\mathcal{S}$:

$$\underline{\mathscr{R}}_v[H] \equiv \mathcal{S}^{-1}\mathscr{R}_v[H]\,\mathcal{S} \subset \mathrm{SO}(2n_v) \quad \Leftrightarrow \quad \underline{\mathscr{R}}_v[h]^T\underline{\mathscr{R}}_v[h] = \mathbf{I}\,, \quad \forall h \in H. \tag{72}$$

For any point ϕ on the scalar manifold define now the *hybrid coset-representative matrix* $\mathbb{L}(\phi) = (\mathbb{L}(\phi)^M{}_{\underline{N}})$ as follows:

---

[5]A special Kähler manifold is in general characterized by the product of a U(1)-bundle, associated with its Kähler structure (with respect to which the manifold is Hodge Kähler), and a flat symplectic bundle. See for instance [19] for an in depth account of this issue.

[6]We label the new basis by underlined indices.

$$\mathbb{L}(\phi) \equiv \mathscr{R}_v[L(\phi)]\mathcal{S} \quad \Leftrightarrow \quad \mathbb{L}(\phi)^M{}_{\underline{N}} \equiv \mathscr{R}_v[L(\phi)]^M{}_N \mathcal{S}^N{}_{\underline{N}}. \tag{73}$$

We also define the matrix

$$\mathbb{L}(\phi)_M{}^{\underline{N}} \;\equiv\; \mathbb{C}_{MP}\,\mathbb{C}^{\underline{NQ}}\,\mathbb{L}(\phi)^P{}_{\underline{Q}}\,. \tag{74}$$

Notice that, as a consequence of the fact that the two indices of $\mathbb{L}$ refer to two different symplectic bases, $\mathbb{L}$ itself is not a matrix representation of the coset representative L. From (11), the property of $\mathscr{R}_v$ of being a representation and the definition (73) we have:

$$\forall \mathbf{g} \in G \;:\quad \mathscr{R}_v[\mathbf{g}]\,\mathbb{L}(\phi) = \mathbb{L}(\mathbf{g}\star\phi)\,\underline{\mathscr{R}}_v[h], \tag{75}$$

where $h \equiv h(\phi, \mathbf{g})$ is the compensating transformation. The hybrid index structure of $\mathbb{L}$ poses no consistency problem since, by (75), the coset representative is acted on to the left and to the right by two different groups: G and H, respectively. Therefore, in our notations, underlined symplectic indices $\underline{M}, \underline{N}, \dots$ are acted on by H while non-underlined ones by G.

The $\mathcal{M}(\phi)$ is then expressed in terms of the coset representative as follows:

$$\mathcal{M}(\phi)_{MN} = \mathbb{C}_{MP}\mathbb{L}(\phi)^P{}_{\underline{L}}\mathbb{L}(\phi)^R{}_{\underline{L}}\,\mathbb{C}_{RN} \quad \Leftrightarrow \quad \mathcal{M}(\phi) = \mathbb{C}\mathbb{L}(\phi)\,\mathbb{L}(\phi)^T\,\mathbb{C}, \tag{76}$$

where summation over the index$\underline{L}$ is understood. The reader can easily verify that the definition of the matrix $\mathcal{M}(\phi)$ given above is indeed consistent, in that it is H-invariant, and thus only depends on the point ϕ, and transforms according to (65):

$$\begin{aligned}\forall g \in G \;:\quad \mathcal{M}(g\star\phi) &= \mathbb{C}\mathbb{L}(g\star\phi)\,\mathbb{L}(g\star\phi)^T\mathbb{C}\\ &= \mathbb{C}\mathscr{R}_v[g]\,\mathbb{L}(\phi)(\underline{\mathscr{R}}_v[h]^{-1}\,\underline{\mathscr{R}}_v[h]^{-T})\mathbb{L}(\phi)^T\mathscr{R}_v[g]^T\mathbb{C}\\ &= \mathscr{R}_v[g]^{-T}\mathbb{C}\mathbb{L}(\phi)\,\mathbb{L}(\phi)^T\mathbb{C}\mathscr{R}_v[g]^{-1}\\ &= \mathscr{R}_v[g]^{-T}\mathcal{M}(\phi)\mathscr{R}_v[g]^{-1},\end{aligned} \tag{77}$$

where we have used (75), the orthogonality property (72) of $\underline{\mathscr{R}}_v[h]$ and the symplectic property of $\mathscr{R}_v[g]$. From the definition (76) of $\mathcal{M}$ in terms of the coset representative, it follows that for symmetric scalar manifolds the scalar Lagrangian (33) can also be written in the equivalent form:

$$\mathscr{L}_{\text{scal}} = \frac{e}{2}\,\mathcal{G}_{st}(\phi)\partial_\mu\phi^s\,\partial^\mu\phi^t = \frac{e}{8}\,k\,\mathrm{Tr}\big(\mathcal{M}^{-1}\partial_\mu\mathcal{M}\,\mathcal{M}^{-1}\partial^\mu\mathcal{M}\big), \tag{78}$$

where k depends on the representation $\mathscr{R}_v$ of G.

The transformation properties of the matrices $\mathcal{I}_{\Lambda\Sigma}$ and $\mathcal{R}_{\Lambda\Sigma}$ under G can be inferred from (65) and can be conveniently described by defining the complex symmetric matrix

$$\mathfrak{N}_{\Lambda\Sigma} \equiv \mathcal{R}_{\Lambda\Sigma} + i\,\mathcal{I}_{\Lambda\Sigma}. \tag{79}$$

Under the action of a generic element $g \in G$, $\mathfrak{N}$ transforms as follows:

$$\mathfrak{N}(g\star\phi) = (C[g] + D[g]\,\mathfrak{N}(\phi))(A[g] + B[g]\,\mathfrak{N}(\phi))^{-1}, \tag{80}$$

where $A[g],\ B[g],\ C[g],\ D[g]$ are the $n_v \times n_v$ blocks of the matrix $\mathscr{R}_v[g]$ defined in (63).

Parity. We have specified above that only the elements of G which belong to the identity sector, namely which are continuously connected to the identity, are associated with symplectic transformations. There may exist isometries $g \in G$ which do not belong to the identity sector and are associated with *anti-symplectic* matrices $\mathbf{A}[g]$:

$$\mathcal{M}(g\star\phi) = \mathbf{A}[g]^{-T}\,\mathcal{M}(\phi)\,\mathbf{A}[g]\ ;\quad \mathbf{A}[g]^T\mathbb{C}\mathbf{A}[g] = -\mathbb{C}. \tag{81}$$

Anti-symplectic matrices do not close a group but can be expressed as the product of a symplectic matrix $\mathbf{S}$ times a fixed anti-symplectic one $\mathbf{P}$, that is $\mathbf{A} = S\,\mathbf{P}$. In a suitable symplectic frame, the matrix $\mathbf{P}$ can be written in the following form:

$$\mathbf{P} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{pmatrix}. \tag{82}$$

Due to their being implemented by anti-symplectic duality transformations (63), these isometries leave (53) invariant up to a sign which can be offset by a *parity transformation*, since under parity one has $* \to -*$. Indeed one can show that these transformations are a symmetry of the theory provided they are combined with parity. Notice that this poses no problem with the generalized theta-term since, as parity reverses the sign of $\varepsilon^{\mu\nu\rho\sigma}F^{\Lambda}_{\mu\nu}F^{\Sigma}_{\rho\sigma}$, under $\mathbf{P}$ we have:

$$\mathcal{I}_{\Lambda\Sigma} \to \mathcal{I}_{\Lambda\Sigma}\ ;\qquad \mathcal{R}_{\Lambda\Sigma} \to -\mathcal{R}_{\Lambda\Sigma}, \tag{83}$$

see (80), so that the corresponding term $\varepsilon^{\mu\nu\rho\sigma}F^{\Lambda}_{\mu\nu}F^{\Sigma}_{\rho\sigma}\mathcal{R}_{\Lambda\Sigma}$ in the Lagrangian is invariant. The global symmetry group of the theory is therefore described by a group

$$G = G_0 \times \mathbb{Z}_2 = \{G_0,\ G_0 \cdot p\}, \tag{84}$$

where G_0 is the *proper duality* group defined by the identity sector of G and p is the element of G which corresponds, in a suitable symplectic frame, to the anti-symplectic matrix $\mathbf{P}$: $\mathbf{P} = \mathbf{A}[p]$.

Example. Let us discuss the simple example of the lower-half complex plane

$$G/H = \mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2). \tag{85}$$

This manifold is parametrized by a complex coordinate z, with $\mathrm{Im}(z) < 0$. As symplectic representation of $G = \mathrm{SL}(2, \mathbb{R})$ we can choose the fundamental representation and the following basis of generators of $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{R})$:

$$\mathfrak{sl}(2, \mathbb{R}) = \{\sigma^1,\ i\,\sigma^2, \sigma^3\} = \left\{\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\ \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},\ \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right\}. \tag{86}$$

The subalgebra $\mathscr{S}$ of upper-triangular generators

$$\mathscr{S} = \{\sigma^3,\ \sigma^+\}, \quad \sigma^+ \equiv \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}. \tag{87}$$

defines the solvable parametrization $\phi^s = (\varphi,\ \chi)$, in which the coset representative $\mathbb{L}$ has the following form:

$$\mathbb{L}(\varphi,\ \chi) \equiv e^{\chi\sigma^+}\, e^{\frac{\varphi}{2}\sigma^3} = \begin{pmatrix}1 & \chi\\ 0 & 1\end{pmatrix}\begin{pmatrix}e^{\varphi/2} & 0\\ 0 & e^{-\varphi/2}\end{pmatrix} \in\ e^{\mathscr{S}}. \tag{88}$$

The relation between the solvable coordinates and z is

$$z\ =\ z_1 + i\,z_2\ =\ \chi - i\,e^{\varphi}. \tag{89}$$

The metric reads:

$$ds^2 = \frac{d\varphi^2}{2} + \frac{1}{2}d\chi^2 e^{-2\varphi} = \frac{1}{2z_2^2}\,dzd\bar{z}\ ; \tag{90}$$

and the matrix $\mathcal{M}(\phi)_{MN}$ reads:

$$\mathcal{M}(z,\ \bar{z})_{MN} = \mathbb{C}_{MP}\,\mathbb{L}(\phi)^P{}_{\underline{L}}\,\mathbb{L}(\phi)^R{}_{\underline{L}}\,\mathbb{C}_{RN} = \frac{1}{z_2}\begin{pmatrix}1 & -z_1\\ -z_1 & |z|^2\end{pmatrix}. \tag{91}$$

The generic isometry which is continuously connected to the identity is a holomorphic transformation of the form

$$z \to z' = \frac{az+b}{cz+d}, \qquad ad - bc = 1, \tag{92}$$

corresponding to the $\mathrm{SL}(2, \mathbb{R})$ transformation $S = \begin{pmatrix}a & b\\ c & d\end{pmatrix}$ with $\det(S) = 1$. The reader can easily verify that:

$$\mathcal{M}(z',\ \bar{z}') = S^{-T}\mathcal{M}(z,\ \bar{z})S^{-1}. \tag{93}$$

We also have the following isometry:

$$z \to -\bar{z}, \tag{94}$$

which is not in the identity sector of the isometry group, and corresponds to the anti-symplectic transformation $\mathbf{P} = \mathrm{diag}(1, -1)$ in that:

$$\mathcal{M}(-\bar{z}, -z) = \mathbf{P}^{-T}\mathcal{M}(z, \bar{z})\mathbf{P}^{-1}. \tag{95}$$

This corresponds to a parity transformation whose effect is to change the sign of the pseudo-scalar χ while leaving the scalar φ inert:

$$\text{parity}: \quad \chi \to -\chi, \quad \varphi \to \varphi. \tag{96}$$

Notice that the correspondence between the linear transformation $\mathbf{P}$ and the isometry (94) exists since $\mathbf{P}$ is an *outer-automorphism* of the isometry algebra $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{R})$, namely:

$$\mathbf{P}^{-1}\mathfrak{sl}(2, \mathbb{R})\mathbf{P} \;=\; \mathfrak{sl}(2, \mathbb{R}), \tag{97}$$

while $\mathbf{P}$ is *not* in $\mathrm{SL}(2, \mathbb{R})$ and the above transformation cannot be offset by any conjugation by $\mathrm{SL}(2, \mathbb{R})$ elements. Analogous outer-automorphisms implementing parity can be found in other extended supergravities, including the maximal one in which $G = \mathrm{E}_{7(7)} \times \mathbb{Z}_2$ [21].

Solitonic solutions, electric-magnetic charges and duality. Ungauged supergravities only contain fields which are neutral with respect to the $\mathrm{U}(1)^{n_v}$ gauge-symmetry of the vector fields. These theories however feature *solitonic solutions*, namely configurations of neutral fields which carry $\mathrm{U}(1)^{n_v}$ electric-magnetic charges. These solutions are typically black holes in four dimensions or black branes in higher and have been extensively studied in the literature. On a charged dyonic solution of this kind, we define the electric and magnetic charges as the integrals[7]:

$$\begin{aligned} e_\Lambda &\equiv \int_{S^2} \mathcal{G}_\Lambda = \frac{1}{2}\int_{S^2} \mathcal{G}_{\Lambda\,\mu\nu}\, dx^\mu \wedge dx^\nu, \\ m^\Lambda &\equiv \int_{S^2} F^\Lambda = \frac{1}{2}\int_{S^2} F^\Lambda{}_{\mu\nu}\, dx^\mu \wedge dx^\nu, \end{aligned} \tag{98}$$

where S^2 is a spatial two-sphere. They define a symplectic vector Γ^M:

$$\Gamma = (\Gamma^M) = \begin{pmatrix} m^\Lambda \\ e_\Lambda \end{pmatrix} = \int_{S^2} \mathbb{F}^M. \tag{99}$$

[7]The electric and magnetic charges (e, m) are expressed in the rationalized-Heaviside-Lorentz (RHL) system of units.

These are the *quantized charges*, namely they satisfy the Dirac-Schwinger-Zwanziger quantization condition for dyonic particles [22–24]:

$$\Gamma_2^T \mathbb{C} \Gamma_1 = m_2^\Lambda \, e_{1\Lambda} - m_1^\Lambda \, e_{2\Lambda} = 2\pi \, \hbar \, c \, n; \quad n \in \mathbb{Z}. \tag{100}$$

At the quantum level, the dyonic charges therefore belong to a symplectic lattice and this breaks the duality group G to a suitable discrete subgroup $G(\mathbb{Z})$ which leaves this symplectic lattice invariant:

$$G(\mathbb{Z}) \equiv G \cap \mathrm{Sp}(2n_v, \mathbb{Z}). \tag{101}$$

This discrete symmetry group of surviving quantum corrections (or a suitable extension thereof) was conjectured in [3] to encode all known string/M-theory dualities.

2.5 Symplectic Frames and Lagrangians

As pointed out earlier, the duality action $\mathscr{R}_v[G]$ of G depends on which elements, in a basis of the representation space, are chosen to be the n_v electric vector fields (appearing in the Lagrangian) and which their magnetic duals namely on the choice of the *symplectic frame* which determines the embedding of the group G inside $\mathrm{Sp}(2n_v, \mathbb{R})$. Different choices of the symplectic frame may yield inequivalent Lagrangians (that is Lagrangians that are not related by local field redefinitions) with different global symmetries. Indeed, the global symmetry group of the Lagrangian[8] is defined as the subgroup $G_{el} \subset G$, whose duality action is linear on the electric field strengths

$$g \in G_{el} \;:\quad \mathscr{R}_v[g] = \begin{pmatrix} A^\Lambda{}_\Sigma & \mathbf{0} \\ C_{\Lambda\Sigma} & D_\Lambda{}^\Sigma \end{pmatrix}, \tag{102}$$

where $D = A^{-T}$ by the symplectic condition, so that

$$\begin{aligned} g \in G_{el} \;:\; F^\Lambda &\to F'^\Lambda = A^\Lambda{}_\Sigma \, F^\Sigma \,, \\ \mathcal{G}_\Lambda &\to \mathcal{G}'_\Lambda = C_{\Lambda\Sigma} \, F^\Sigma + D_\Lambda{}^\Sigma \, \mathcal{G}_\Sigma. \end{aligned} \tag{103}$$

Indeed, as the reader can verify using (80), under the above transformation the matrices $\mathcal{I}, \, \mathcal{R}$ transform as follows:

$$\mathcal{I}_{\Lambda\Sigma} \to D_\Lambda{}^\Pi D_\Sigma{}^\Delta \, \mathcal{I}_{\Pi\Delta}; \quad \mathcal{R}_{\Lambda\Sigma} \to D_\Lambda{}^\Pi D_\Sigma{}^\Delta \, \mathcal{R}_{\Pi\Delta} + C_{\Lambda\Pi} \, D_\Sigma{}^\Pi, \tag{104}$$

[8]Here we only consider *local* transformations on the fields.

and the consequent variation of the Lagrangian reads

$$\mathscr{L}_{\rm B} = \frac{1}{8}\, C_{\Lambda\Pi}\, D_{\Sigma}{}^{\Pi}\, \varepsilon^{\mu\nu\rho\sigma}\, F^{\Lambda}_{\mu\nu} F^{\Sigma}_{\rho\sigma}\,, \tag{105}$$

which is a *total derivative* since $C_{\Lambda\Pi}\, D_{\Sigma}{}^{\Pi}$ is constant. These transformations are called *Peccei-Quinn transformations* and follow from shifts in certain axionic scalar fields. They are a symmetry of the classical action, while invariance of the perturbative path-integral requires the variation (105), integrated over space-time, to be proportional through an integer to $2\pi\hbar$. This constrains the symmetries to belong to a discrete subgroup $G(\mathbb{Z})$ of G whose duality action is implemented by integer-valued matrices $\mathscr{R}_v[g]$. Such restriction of G to $G(\mathbb{Z})$ in the quantum theory was discussed earlier as a consequence of the Dirac-Schwinger-Zwanziger quantization condition for dyonic particles (100).

From (103) we see that, while the vector field strengths $F^{\Lambda}_{\mu\nu}$ and their duals $\mathcal{G}_{\Lambda\,\mu\nu}$ transform together under G in the ($2n_v$–dimensional) symplectic representation $\mathscr{R}_v$, the vector field strengths alone transform linearly under the action of G_{el} in a smaller representation $\mathbf{n_v}$, defined by the A-block in (102).

Different symplectic frames of a same ungauged theory may originate from different compactifications. A distinction here is in order. In $\mathrm{N} \geq 3$ theories, scalar fields always enter the same multiplets as the vector fields. Supersymmetry then implies their non-minimal coupling to the latter and that the scalar manifold be endowed with a symplectic structure associating with each isometry a constant symplectic matrix. In $\mathrm{N} = 2$ theories, scalar fields may sit in vector multiplets or hypermultiplets. The former span a *special Kähler manifold*, the latter a *quaternionic Kähler* one, so that the scalar manifold is always factorized in the product of the two:

$$\mathscr{M}^{(\mathrm{N}=2)}_{\rm scal} \;=\; \mathscr{M}_{\rm SK} \times \mathscr{M}_{\rm QK}\,. \tag{106}$$

The scalar fields in the hypermultiplets are not connected to vector fields through supersymmetry and consequently they do not enter the matrices $\mathcal{I}(\phi)$ and $\mathcal{R}(\phi)$. As a consequence of this the isometries of the Quaternionic-Kähler manifolds spanned by these scalars are associated with trivial duality transformations

$$g \,\in\, \text{isom. of } \mathscr{M}_{\rm QK} \quad \Rightarrow \quad \mathscr{R}_v[g] = \mathbf{I}\,, \tag{107}$$

while only $\mathscr{M}_{\rm SK}$ features a flat symplectic structure which defines the embedding of its isometry group inside $\mathrm{Sp}(2n_v, \mathbb{R})$ and the couplings of the vector multiplet-scalars to the vector fields through the matrix $\mathcal{M}(\phi)$. It is important to remark that such structure on a special Kähler manifold exists even if the manifold itself is not homogeneous. This means that one can still define the symplectic matrix $\mathbb{L}(\phi)$ and, in terms of the components $\mathcal{I}_{\Lambda\Sigma}$ and $\mathcal{R}_{\Lambda\Sigma}$, also the matrix $\mathcal{M}(\phi)$ as in (76), although $\mathbb{L}(\phi)$ has no longer the interpretation of a coset representative for non-homogeneous manifolds.

It is convenient for later purposes to rewrite the transformation properties of the bosonic fields the group G, discussed in this section, in the following infinitesimal form:

$$G\ :\quad \begin{cases} \delta\,\mathbb{L} = \Lambda^\alpha\, t_\alpha\, \mathbb{L}\ , \\ \delta\mathbb{F}^M_{\mu\nu} = -\Lambda^\alpha\, (t_\alpha)_N{}^M\, \mathbb{F}^N_{\mu\nu}, \end{cases}$$

in terms of the infinitesimal generators t_α of G introduced earlier and, satisfying the relation (36). The matrices $(t_\alpha)_M{}^N$ define the infinitesimal duality action of G and are symplectic generators

$$(t_\alpha)_M{}^N\,\mathbb{C}_{NP} = (t_\alpha)_P{}^N\,\mathbb{C}_{NM} \qquad M,\ N, \ldots = 1, \ldots,\ 2n_v\ . \tag{108}$$

This is equivalently stated as the property of the tensor $t_{\alpha\,MN} \equiv (t_\alpha)_M{}^P\,\mathbb{C}_{PN}$ of being symmetric in $M\,N$:

$$(t_\alpha)_{MN} = (t_\alpha)_{NM}. \tag{109}$$

2.6 The Fermionic Sector

Fermions in supergravity transform covariantly with respect to the isotropy group H of the scalar manifold, which has the general form (9), while they do not transform under G, as opposed to the bosonic fields. Bosons and fermions have therefore definite transformation properties with respect to different groups of internal symmetry. The matrix $\mathbb{L}$, defining the coset representative for homogeneous scalar manifolds, transforms under the action of G to the left and of H to the right, according to (11)

$$G\ \rightarrow\ \mathbb{L}\ \leftarrow\ H, \tag{110}$$

and thus has the right index structure to "mediates" in the Lagrangian between bosons and fermions. This means that we can construct G-invariant terms by contracting $\mathbb{L}$ to the left by bosons (scalars, vectors and their derivatives), and to the right by fermions

$$(\text{Bosons}) \star \mathbb{L}(\phi) \star (\text{Fermions}), \tag{111}$$

the two $\star$ symbols denote some contraction of indices: G-invariant to the left and H-invariant to the right. The "Boson" part of (111) may also contain $\mathbb{L}$ and its derivatives. These are the kind of terms occurring in the field equations. If under a transformation $g \in G$, symbolically:

$$\text{Bosons}\ \rightarrow\ \text{Bosons}' = \text{Bosons} \star g^{-1}, \tag{112}$$

and the *fermions are made to transform under the compensating transformation* $h(\phi, g)$ in (11):

$$\text{Fermions} \;\rightarrow\; \text{Fermions}' = h(\phi, g) \star \text{Fermions}. \qquad (113)$$

Using (11) we see that (111) remains invariant:

$$(\text{Bosons})' \star \mathbb{L}(\mathrm{g} \star \phi) \star (\text{Fermions}') = (\text{Bosons}) \star \mathbb{L}(\phi) \star (\text{Fermions}). \qquad (114)$$

The Lagrangian is manifestly invariant under local H-transformations since the covariant derivatives on the fermion fields contain the H-connection[9] w_μ:

$$\mathscr{D}_\mu \xi = \nabla_\mu \xi + w_\mu \star \xi, \qquad (115)$$

where, as usual, the $\star$ symbol denotes the action of the $\mathfrak{H}$-valued connection w_μ on ξ in the corresponding H-representation. The reader can verify that (115) is indeed covariant under local H-transformations (113), provided w is transformed according to (24). As opposed to the gauge groups we are going to introduce by the gauging procedure, which involve minimal couplings to the vector fields of the theory, the local H-symmetry group of the ungauged theory is not gauged by the vector fields, but by a *composite connection* w_μ, which is a function of the scalar fields and their derivatives. The minimal coupling $w_\mu \star \xi$ is an example of the boson-fermion interaction term (111).

It is useful to write the coupling (111) in the following form:

$$\mathbf{f}(\phi, \text{Bosons}) \star (\text{Fermions}), \qquad (116)$$

where we have introduced the H-covariant *composite field*:

$$\mathbf{f}(\phi, \text{Bosons}) \equiv (\text{Bosons}) \star \mathbb{L}(\phi), \qquad (117)$$

obtained by *dressing* the bosonic fields and their derivatives with the coset-representative so as to obtain an H-covariant quantity with the correct H-index structure to contract with fermionic currents. Indeed under a G-transformation

$$\mathbf{f}(g \star \phi, \text{Bosons}') \equiv \mathbf{f}(\phi, \text{Bosons}) \star h(\phi, g)^{-1}, \qquad (118)$$

The manifest H-invariance of the supergravity theory requires the supersymmetry transformation properties of the femionic fields to be H-covariant. Indeed such transformation rules, which in rigid supersymmetric theories (i.e. theories which are invariant only under global supersymmetry) can be schematically described as follows[10]:

[9] We define $w_\mu \equiv w_s \, \partial_\mu \phi^s$.

[10] This is a schematic representation in which we have suppressed the Lorentz indices and gamma-matrices.

$$\delta\text{Fermion} = \sum_{\text{Bosons}} \partial\text{Boson} \cdot \varepsilon, \tag{119}$$

and in supergravity theories have the following general H-covariant form[11]

$$\delta\text{Fermion} = \sum_{\text{Bosons}} \mathbf{f}(\phi, \text{Bosons}) \cdot \varepsilon, \tag{120}$$

where the space-time derivatives of the bosonic fields are dressed with the scalars in the definition of $\mathbf{f}(\phi, \text{Bosons})$. Examples of composite fields $\mathbf{f}(\phi, \text{Bosons})$ are the vielbein of the scalar manifold (pulled back on space-time) $\mathcal{P}_\mu \equiv \mathcal{P}_s\, \partial_\mu \phi^s$, the H-connection w_μ in (115), the dressed vector field-strengths

$$\mathbf{F}(\phi, \partial A)^{\underline{M}}_{\mu\nu} \equiv -(\mathbb{L}(\phi)^{-1})^{\underline{M}}{}_N\, \mathbb{F}^M_{\mu\nu}, \tag{121}$$

or the $\mathbb{T}$-*tensor*, to be introduced later, in which the bosonic field to be dressed by the coset representative is the *embedding tensor* Θ defining the choice of the gauge algebra.

3 Gauging Supergravities

We have reviewed the field content and the Lagrangian of ungauged supergravity, as well as the action of the global symmetry group G. Now we want to discuss how to construct a gauged theory from an ungauged one.

In the following, we will employ a covariant formalism in which the possible gaugings will be encoded into an object called embedding tensor, that can be characterized group-theoretically [5, 16, 17].

3.1 The Gauging Procedure Step-by-Step

As anticipated in the Introduction, the gauging procedure consists in promoting a suitable global symmetry group $G_g \subset G_{el}$ of the Lagrangian to a local symmetry gauged by the vector fields of the theory. This requirement gives us a preliminary condition

$$\dim(G_g) \;\leq\; n_v\ . \tag{122}$$

[11] The gravitino field has an additional term $\mathscr{D}\varepsilon$ which is its variation as the gauge field of local supersymmetry.

As explained in Sect. 2.5, *different symplectic frames correspond to ungauged Lagrangians with different global symmetry groups G_{el} and thus to different choices for the possible gauge groups.*

The first condition for the global symmetry subgroup G_g to become a viable gauge group, is that there should exist a subset $\{A^{\hat{\Lambda}}\}$ of the vector fields[12] which transform under the co-adjoint representation of the duality action of G_g. These fields will become the *gauge vectors* associated with the *generators* $X_{\hat{\Lambda}}$ of the subgroup G_g.

We shall name *electric frame* the symplectic frame defined by our ungauged Lagrangian and labeled by hatted indices.

Note that, once the gauge group is chosen within G_{el}, its action on the various fields is fixed, being it defined by the action of G_g as a global symmetry group of the ungauged theory (duality action on the vector field strengths, non-linear action on the scalar fields and indirect action through H-compensators on the fermionic fields): fields are thus automatically associated with representations of G_g.

After the initial choice of G_g in G_{el}, the first part of the procedure is quite standard in the construction of non-abelian gauge theories: we introduce a gauge-connection, gauge-curvature (i.e. non-abelian field strengths) and covariant derivatives. We will also need to introduce an extra topological term needed for the gauging of the Peccei-Quinn transformations (105). This will lead us to construct a gauged Lagrangian $\mathscr{L}^{(0)}_{\rm gauged}$ with manifest local G_g-invariance. Consistency of the construction will imply constraints on the possible choices of G_g inside G. The minimal couplings will however break supersymmetry.

The second part of the gauging procedure consists in further deforming the Lagrangian $\mathscr{L}^{(0)}_{\rm gauged}$ in order to restore the original supersymmetry of the ungauged theory and, at the same time, preserving local G_g-invariance.

Step 1. *Choice of the gauge algebra*. We start by introducing the *gauge connection*:

$$\Omega_g = \Omega_{g\,\mu} dx^\mu \ ; \quad \Omega_{g\,\mu} \equiv g\, A^{\hat{\Lambda}}_\mu\, X_{\hat{\Lambda}}, \tag{123}$$

g being the coupling constant. The gauge-algebra relations can be written in the form

$$\left[X_{\hat{\Lambda}},\, X_{\hat{\Sigma}}\right] \ = f_{\hat{\Lambda}\hat{\Sigma}}{}^{\hat{\Gamma}}\, X_{\hat{\Gamma}}, \tag{124}$$

and are characterized by the structure constants $f_{\hat{\Lambda}\hat{\Sigma}}{}^{\hat{\Gamma}}$. This closure condition should be regarded as a constraint on $X_{\hat{\Lambda}}$, since the structure constants are not generic but fixed in terms of the action of the gauge generators on the vector fields as global symmetry generators of the original ungauged theory. To understand this, let us recall that G_g is a subgroup of G_{el} and thus its electric-magnetic duality action, as a global symmetry group, will have the form (102). Therefore the duality action on the vector field strengths and their duals of the infinitesimal generators $X_{\hat{\Lambda}}$ will then by represented by a symplectic matrix of the form (see Eq. (102))

[12] We describe by hatted-indices those pertaining to the symplectic frame in which the Lagrangian is defined.

$$\left(X_{\hat{\Lambda}}\right)^{\hat{M}}{}_{\hat{N}} = \begin{pmatrix} X_{\hat{\Lambda}}{}^{\hat{\Gamma}}{}_{\hat{\Sigma}} & \mathbf{0} \\ X_{\hat{\Lambda}\,\hat{\Gamma}\hat{\Sigma}} & X_{\hat{\Lambda}\,\hat{\Gamma}}{}^{\hat{\Delta}} \end{pmatrix}, \tag{125}$$

where $X_{\hat{\Lambda}}{}^{\hat{\Gamma}}{}_{\hat{\Sigma}}$ and $X_{\hat{\Lambda}\,\hat{\Gamma}}{}^{\hat{\Delta}}$ are the infinitesimal generators of the A and D-blocks in (102) respectively, while $X_{\hat{\Lambda}\,\hat{\Gamma}\hat{\Sigma}}$ describes the infinitesimal C-block. It is worth emphasizing here that we do not identify the generator $X_{\hat{\Lambda}}$ with the symplectic matrix defining its electric-magnetic duality action. As pointed our in Sect. 2.5, there are isometries in N $= 2$ models which do not have duality action, see (107), namely for which the matrix in (125) is null.

The variation of the field strengths under an infinitesimal transformation $\xi^{\hat{\Lambda}} X_{\hat{\Lambda}}$, whose duality action is described by (125), is:

$$\delta\mathbb{F}^{\hat{M}} = \xi^{\hat{\Lambda}}\,(X_{\hat{\Lambda}})^{\hat{M}}{}_{\hat{N}}\,\mathbb{F}^{\hat{N}} \quad \Rightarrow \quad \begin{cases} \delta F^{\hat{\Lambda}} = \xi^{\hat{\Gamma}} X_{\hat{\Gamma}}{}^{\hat{\Lambda}}{}_{\hat{\Sigma}}\,F^{\hat{\Sigma}}, \\ \delta\mathcal{G}_{\hat{\Lambda}} = \xi^{\hat{\Gamma}} X_{\hat{\Gamma}\,\hat{\Lambda}\hat{\Sigma}} F^{\hat{\Sigma}} + \xi^{\hat{\Gamma}} X_{\hat{\Gamma}\hat{\Lambda}}{}^{\hat{\Sigma}}\,\mathcal{G}_{\hat{\Sigma}}. \end{cases} \tag{126}$$

The symplectic condition on the matrix $X_{\hat{\Lambda}}$ implies the properties:

$$X_{\hat{\Lambda}\hat{M}}{}^{\hat{P}}\,\mathbb{C}_{\hat{N}\hat{P}} = X_{\hat{\Lambda}\hat{N}}{}^{\hat{P}}\,\mathbb{C}_{\hat{M}\hat{P}} \quad \Leftrightarrow \quad \begin{cases} X_{\hat{\Lambda}}{}^{\hat{\Sigma}}{}_{\hat{\Gamma}} = -X_{\hat{\Lambda}\hat{\Gamma}}{}^{\hat{\Sigma}}, \\ X_{\hat{\Lambda}\,\hat{\Gamma}\hat{\Sigma}} = X_{\hat{\Lambda}\,\hat{\Sigma}\hat{\Gamma}}. \end{cases} \tag{127}$$

The condition that $A_\mu^{\hat{\Lambda}}$ transform in the co-adjoint representation of the gauge group:

$$\delta F^{\hat{\Lambda}} = \xi^{\hat{\Gamma}} f_{\hat{\Gamma}\hat{\Sigma}}{}^{\hat{\Lambda}} F^{\hat{\Sigma}}, \tag{128}$$

together with the transformation properties (126), lead us to identify the structure constants of the gauge group in (124) with the diagonal blocks of the symplectic matrices $X_{\hat{\Lambda}}$:

$$f_{\hat{\Gamma}\hat{\Sigma}}{}^{\hat{\Lambda}} = -X_{\hat{\Gamma}\hat{\Sigma}}{}^{\hat{\Lambda}}, \tag{129}$$

so that the closure condition reads

$$\left[X_{\hat{\Lambda}},\,X_{\hat{\Sigma}}\right] = -X_{\hat{\Lambda}\hat{\Sigma}}{}^{\hat{\Gamma}}\,X_{\hat{\Gamma}}, \tag{130}$$

and is a quadratic constraint on the tensor $X_{\hat{\Lambda}}{}^{\hat{M}}{}_{\hat{N}}$. The identification (129) also implies

$$X_{(\hat{\Gamma}\hat{\Sigma})}{}^{\hat{\Lambda}} = 0. \tag{131}$$

The closure condition (130) can thus be interpreted in two equivalent ways:

- the vector fields $A_\mu^{\hat{\Lambda}}$ transform in the co-adjoint representation of G_g under its action as global symmetry, namely

$$\mathbf{n_v} = \text{co-adj}(G_g); \tag{132}$$

◦ the gauge generators $X_{\hat{\Lambda}}$ are invariant under the action of G_g itself:

$$\delta_{\hat{\Lambda}} X_{\hat{\Sigma}} \equiv [X_{\hat{\Lambda}}, X_{\hat{\Sigma}}] + X_{\hat{\Lambda}\hat{\Sigma}}{}^{\hat{\Gamma}} X_{\hat{\Gamma}} = 0. \tag{133}$$

Step 2. *Introducing gauge curvatures and covariant derivatives*. Having defined the gauge connection (123) we also define its transformation property under a local G_g-transformation $\mathbf{g}(x) \in G_g$:

$$\Omega_g \;\rightarrow\; \Omega'_g = \mathbf{g}\,\Omega_g\,\mathbf{g}^{-1} + d\mathbf{g}\,\mathbf{g}^{-1} = g\,A'^{\hat{\Lambda}}\,X_{\hat{\Lambda}}. \tag{134}$$

Under an infinitesimal transformation $\mathbf{g}(x) \equiv \mathbf{I} + g\,\zeta^{\hat{\Lambda}}(x)\,X_{\hat{\Lambda}}$, (134) implies the following transformation property of the gauge vectors:

$$\delta A_\mu^{\hat{\Lambda}} \;=\; \mathcal{D}_\mu \zeta^{\hat{\Lambda}} \;\equiv\; \partial_\mu \zeta^{\hat{\Lambda}} + g\,A_\mu^{\hat{\Sigma}} X_{\hat{\Sigma}\hat{\Gamma}}{}^{\hat{\Lambda}}\,\zeta^{\hat{\Gamma}}, \tag{135}$$

where we have introduced the G_g-covariant derivative of the gauge parameter $\mathcal{D}_\mu \zeta^{\hat{\Lambda}}$.

As usual in the construction of non-abelian gauge-theories, we define the gauge curvature[13]

$$g\,\mathcal{F} = g\,F^{\hat{\Lambda}}\,X_{\hat{\Lambda}} = \frac{g}{2}\,F_{\mu\nu}^{\hat{\Lambda}}\,dx^\mu \wedge dx^\nu\,X_{\hat{\Lambda}} \equiv d\Omega_g - \Omega_g \wedge \Omega_g, \tag{136}$$

which, in components, reads:

$$F_{\mu\nu}^{\hat{\Lambda}} \;=\; \partial_\mu A_\nu^{\hat{\Lambda}} - \partial_\nu A_\mu^{\hat{\Lambda}} - g f_{\hat{\Gamma}\hat{\Sigma}}{}^{\hat{\Lambda}}\,A_\mu^{\hat{\Gamma}}\,A_\nu^{\hat{\Sigma}} \;=\; \partial_\mu A_\nu^{\hat{\Lambda}} - \partial_\nu A_\mu^{\hat{\Lambda}} + g\,X_{\hat{\Gamma}\hat{\Sigma}}{}^{\hat{\Lambda}}\,A_\mu^{\hat{\Gamma}}\,A_\nu^{\hat{\Sigma}}\,. \tag{137}$$

The gauge curvature transforms covariantly under a transformation $\mathbf{g}(x) \in G_g$:

$$\mathcal{F} \;\rightarrow\; \mathcal{F}' = \mathbf{g}\,\mathcal{F}\,\mathbf{g}^{-1}, \tag{138}$$

and satisfies the Bianchi identity:

$$\mathcal{D}\mathcal{F} \equiv d\mathcal{F} - \Omega_g \wedge \mathcal{F} + \mathcal{F} \wedge \Omega_g = 0 \;\;\Leftrightarrow\;\; \mathcal{D}F^{\hat{\Lambda}} \equiv dF^{\hat{\Lambda}} + g\,X_{\hat{\Sigma}\hat{\Gamma}}{}^{\hat{\Lambda}} A^{\hat{\Sigma}} \wedge F^{\hat{\Lambda}} = 0, \tag{139}$$

where we have denoted by $\mathcal{D}F^{\hat{\Lambda}}$ the G_g-covariant derivative acting on $F^{\hat{\Lambda}}$. In the original ungauged Lagrangian we then replace the abelian field strengths by the new G_g-covariant ones:

---

[13] Here we use the following convention for the definition of the components of a form: $\omega_{(p)} = \frac{1}{p!}\,\omega_{\mu_1\ldots\mu_p}\,dx^{\mu_1} \wedge \ldots dx^{\mu_p}$.

$$\partial_\mu A^{\hat{\Lambda}}_\nu - \partial_\nu A^{\hat{\Lambda}}_\mu \quad\rightarrow\quad \partial_\mu A^{\hat{\Lambda}}_\nu - \partial_\nu A^{\hat{\Lambda}}_\mu + g\,X_{\hat{\Gamma}\hat{\Sigma}}{}^{\hat{\Lambda}}\,A^{\hat{\Gamma}}_\mu\,A^{\hat{\Sigma}}_\nu\,. \tag{140}$$

After having given the gauge fields a G_g-covariant description in the Lagrangian through the non-abelian field strengths, we now move to the other fields. The next step in order to achieve local invariance of the Lagrangian under G_g consists in replacing ordinary derivatives by covariant ones

$$\partial_\mu \quad\longrightarrow\quad \mathcal{D}_\mu \;=\; \partial_\mu - g\,A^{\hat{\Lambda}}_\mu\,X_{\hat{\Lambda}}. \tag{141}$$

As it can be easily ascertained, the covariant derivatives satisfy the identity which is well known from gauge theories:

$$\mathcal{D}^2 = -g\,\mathcal{F} = -g\,F^{\hat{\Lambda}}\,X_{\hat{\Lambda}} \quad\Leftrightarrow\quad [\mathcal{D}_\mu,\,\mathcal{D}_\nu] = -g\,F^{\hat{\Lambda}}_{\mu\nu}\,X_{\hat{\Lambda}}. \tag{142}$$

Aside from the vectors and the metric, the remaining bosonic fields are the scalars ϕ^s, whose derivatives are covariantized using the Killing vectors $k_{\hat{\Lambda}}$ associated with the action of the gauge generator $X_{\hat{\Lambda}}$ as an isometry:

$$\partial_\mu \quad\longrightarrow\quad \mathcal{D}_\mu\phi^s \;=\; \partial_\mu\phi^s - g\,A^{\hat{\Lambda}}\,k^s_{\hat{\Lambda}}(\phi), \tag{143}$$

The replacement (141), and in particular (143), amounts to the *introduction of minimal couplings* for the vector fields.

Care is needed for the fermion fields which, as we have discussed above, do not transform directly under G, but under the corresponding compensating transformations in H. This was taken into account by writing the H-connection w in the fermion H-covariant derivatives. Now we need to promote such derivatives to G_g-covariant ones, by minimally coupling the fermions to the gauge fields. This is effected by modifying the H-connection.

For homogeneous scalar manifolds redefine the left-invariant 1-form Ω (pulled-back on space-time), defined on them in (18), by a *gauged* one obtained by covariantizing the derivative on the coset representative:

$$\Omega_\mu = L^{-1}\partial_\mu L \quad\longrightarrow\quad \hat{\Omega}_\mu \equiv L^{-1}\mathcal{D}L = L^{-1}\left(\partial_\mu - g\,A^{\hat{\Lambda}}_\mu\,X_{\hat{\Lambda}}\right)L = \hat{\mathcal{P}}_\mu + \hat{w}_\mu \tag{144}$$

where, as usual, the space-time dependence of the coset representative is defined by the scalar fields $\phi^s(x)$: $\partial_\mu L \equiv \partial_s L\,\partial_\mu\phi^s$.

The *gauged* vielbein and connection are related to the ungauged ones as follows:

$$\hat{\mathcal{P}}_\mu = \mathcal{P}_\mu - g\,A^{\hat{\Lambda}}_\mu\,\mathcal{P}_{\hat{\Lambda}}; \qquad \hat{w}_\mu = w_\mu - g\,A^{\hat{\Lambda}}_\mu\,w_{\hat{\Lambda}}. \tag{145}$$

The matrices $\mathcal{P}_{\hat{\Lambda}},\ w_{\hat{\Lambda}}$ begin the projections onto $\mathfrak{K}$ and $\mathfrak{H}$, respectively, of $L^{-1}X_{\hat{\Lambda}}L$:

$$\mathcal{P}_{\hat{\Lambda}} \equiv L^{-1}X_{\hat{\Lambda}}L\big|_{\mathfrak{K}}; \quad w_{\hat{\Lambda}} \equiv L^{-1}X_{\hat{\Lambda}}L\big|_{\mathfrak{H}}\,. \tag{146}$$

Using (43) we can express the above quantities as follows:

$$\mathcal{P}_{\hat{\Lambda}} = k^s_{\hat{\Lambda}}\, V_s{}^{\underline{s}} K_{\underline{s}}\,; \quad w_{\hat{\Lambda}} = -\frac{1}{2}\,\mathscr{P}^{\mathbf{a}}_{\hat{\Lambda}}\, J_{\mathbf{a}} - \frac{1}{2}\,\mathscr{P}^{\mathbf{m}}_{\hat{\Lambda}}\, J_{\mathbf{m}}, \tag{147}$$

where $\mathscr{P}^{\mathbf{a}}_{\hat{\Lambda}}$ were defined in Sect. 2.1.

For non-homogeneous scalar manifolds we cannot use the construction (144) based on the coset representative. Nevertheless we can still define $\mathscr{P}^{\mathbf{m}}_{\hat{\Lambda}}$, $\mathscr{P}^{\mathbf{a}}_{\hat{\Lambda}}$ in terms of the Killing vectors, see discussion below (45). From these quantities one then defines gauged vielbein $\hat{\mathcal{P}}_\mu$ and H-connection $\hat{w}_\mu$ using (145) and (147), where now $K_{\underline{s}}$ should be intended as a basis of the tangent space to the manifold at the origin (and not as isometry generators) and $\{J_{\mathbf{a}},\, J_{\mathbf{m}}\}$ a basis of the holonomy group.

Notice that, as a consequence of (147) and (145), the gauged vielbein 1-forms (pulled-back on space-time) can be written as the ungauged ones in which the derivatives on the scalar fields are replaced by the covariant ones (143). This is readily seen by applying the general formula (40) for homogeneous manifolds to the isometry $X_{\hat{\Lambda}}$ in (144), and projecting both sides of this equation on the coset space $\mathfrak{K}$:

$$\hat{\mathcal{P}}_\mu = \mathcal{P}_s\, \mathcal{D}_\mu z^s. \tag{148}$$

Consequently the replacement (143) is effected by replacing everywhere in the Lagrangian $\mathcal{P}_\mu$ by $\hat{\mathcal{P}}_\mu$.

Consider now a local G_g-transformation $\mathbf{g}(x)$ whose effect on the scalars is described by (11): $\mathbf{g}L(\phi) = L(\mathbf{g} \star \phi)\, h(\phi, \mathbf{g})$. From (144) and from the fact that $\mathcal{D}$ is the G-covariant derivative, the reader can easily verify that:

$$\hat{\Omega}_\mu(g \star \phi) = h\, \hat{\Omega}_\mu(\phi)\, h^{-1} + hdh^{-1} \quad \Rightarrow \quad \begin{cases} \hat{\mathcal{P}}(g \star \phi) = h\, \hat{\mathcal{P}}(\phi)\, h^{-1}, \\ \hat{w}(g \star \phi) = h\, \hat{w}(\phi)\, h^{-1} + hdh^{-1}, \end{cases} \tag{149}$$

where $h = h(\phi, \mathbf{g})$. By deriving (144) we find the *gauged* Maurer-Cartan equations:

$$d\hat{\Omega} + \hat{\Omega} \wedge \hat{\Omega} = -g\, L^{-1} \mathcal{F} L, \tag{150}$$

where we have used (142). Projecting the above equation onto $\mathfrak{K}$ and $\mathfrak{H}$ we find the gauged version of (27), (28):

$$\mathscr{D}\hat{\mathcal{P}} \equiv d\hat{\mathcal{P}} + \hat{w} \wedge \mathcal{P} + \mathcal{P} \wedge \hat{w} = -g\, F^{\hat{\Lambda}}\, \mathcal{P}_{\hat{\Lambda}}, \tag{151}$$

$$\hat{R}(\hat{w}) \equiv d\hat{w} + \hat{w} \wedge \hat{w} = -\mathcal{P} \wedge \mathcal{P} - g\, F^{\hat{\Lambda}}\, w_{\hat{\Lambda}}. \tag{152}$$

The above equations are manifestly G_g-invariant. Using (148) one can easily verify that the gauged curvature 2-form (with value in $\mathfrak{H}$) can be written in terms of the curvature components R_{rs} of the manifold, given in (29), as follows:

$$\hat{R}(\hat{w}) = \frac{1}{2} R_{rs}\, \mathcal{D}\phi^r \wedge \mathcal{D}\phi^s - g\, F^{\hat{\Lambda}}\, w_{\hat{\Lambda}}. \tag{153}$$

The gauge-covariant derivatives, when acting on a generic fermion field ξ, is defined using $\hat{w}_\mu$, so that (115) is replaced by

$$\mathcal{D}_\mu \xi = \nabla_\mu \xi + \hat{w}_\mu \star \xi. \tag{154}$$

Summarizing, local invariance of the action under G_g requires replacing everywhere in the Lagrangian the abelian field strengths by the non abelian ones, (140) and the ungauged vielbein $\mathcal{P}_\mu$ and H-connection w_μ by the gauged ones:

$$\mathcal{P}_\mu \;\rightarrow\; \hat{\mathcal{P}}_\mu; \qquad w_\mu \;\rightarrow\; \hat{w}_\mu. \tag{155}$$

Clearly supersymmetry of the gauged action would require as a necessary, though not sufficient, condition to perform the above replacements also in the supersymmetry transformation laws of the fields.

Step 3. *Introducing topological terms*. If the symplectic duality action (125) of $X_{\hat{\Lambda}}$ has a non-vanishing off-diagonal block $X_{\hat{\Lambda}\hat{\Gamma}\hat{\Sigma}}$, that is if the gauge transformations include Peccei-Quinn shifts, then an infinitesimal (local) gauge transformation $\xi^{\hat{\Lambda}}(x)\, X_{\hat{\Lambda}}$ would produce a variation of the Lagrangian of the form (105):

$$\delta\mathscr{L}_{\mathrm{B}} = -\frac{g}{8}\, \xi^{\hat{\Lambda}}(x) X_{\hat{\Lambda}\hat{\Gamma}\hat{\Sigma}} \varepsilon^{\mu\nu\rho\sigma}\, F^{\hat{\Gamma}}_{\mu\nu} F^{\hat{\Sigma}}_{\rho\sigma}. \tag{156}$$

Being $\xi^{\hat{\Lambda}}(x)$ a local parameter, the above term is no longer a total derivative and thus the transformation is not a symmetry of the action. In [25] it was proven that the variation (156) can be canceled by adding to the Lagrangian a topological term of the form

$$\mathscr{L}_{\mathrm{top.}} = -\frac{1}{3}\, g\, \varepsilon^{\mu\nu\rho\sigma}\, X_{\hat{\Lambda}\hat{\Gamma}\hat{\Sigma}}\, A^{\hat{\Lambda}}_\mu A^{\hat{\Sigma}}_\nu \left(\partial_\rho A^{\hat{\Gamma}}_\sigma + \frac{3}{8}\, g\, X_{\hat{\Delta}\hat{\Pi}}{}^{\hat{\Gamma}}\, A^{\hat{\Delta}}_\rho A^{\hat{\Pi}}_\sigma \right), \tag{157}$$

provided the following condition holds

$$X_{(\hat{\Lambda}\hat{\Gamma}\hat{\Sigma})} \;=\; 0. \tag{158}$$

We will see in the following that condition (158), together with the closure constraint (130), is part of a set of constraints on the gauge algebra which are also implied by supersymmetry. Indeed, even if the Lagrangian $\mathscr{L}^{(0)}_g$ constructed so far is locally G_g-invariant, the presence of minimal couplings explicitly breaks both supersymmetry and the duality global symmetry G.

Choice of the gauge algebra and the embedding tensor. We have seen that the gauging procedure corresponds to promoting some suitable subgroup $G_g \subset G_{el}$ to a local symmetry. This subgroup is defined selecting a subset of generators within the global symmetry algebra $\mathfrak{g}$ of G. Now, all the information about the gauge algebra can be encoded in a G_{el}-covariant object θ, which expresses the gauge generators as linear combinations of the global symmetry generators t_α of the subgroup $G_{el} \subset G$

$$X_{\hat{\Lambda}} = \theta_{\hat{\Lambda}}{}^{\sigma}\, t_\sigma; \qquad \theta_{\hat{\Lambda}}{}^{\sigma} \in \mathbf{n_v} \times \mathrm{adj}(G_{el}), \tag{159}$$

with $\hat{\Lambda} = 1, \ldots, n_v$ and with $\sigma = 1, \ldots, \dim(G_{el})$. The advantage of this description is that the G_{el}-invariance of the original ungauged Lagrangian $\mathscr{L}$ is restored at the level of the gauged Lagrangian $\mathscr{L}_{\mathrm{gauged}}$, to be constructed below, provided $\theta_{\hat{\Lambda}}{}^{\sigma}$ is transformed under G_{el} as well. However, the full global symmetry group G of the field equations and Bianchi identities is still broken, since the parameters $\theta_{\hat{\Lambda}}{}^{\sigma}$ can be viewed as of electric charges, whose presence manifestly break electric-magnetic duality invariance. In other words we are working in a specific symplectic frame defined by the ungauged Lagrangian we started from.

We shall give later on a definition of the gauging procedure which is completely freed from the choice of the symplectic frame. For the time being, it is useful to give a description of the gauge algebra (and of the consistency constraints on it) which does not depend on the original symplectic frame, namely which is manifestly G-covariant. This is done by encoding all information on the initial symplectic frame in a symplectic matrix $E \equiv (E_M{}^N)$ and writing the gauge generators, through this matrix, in terms of new generators

$$X_M = (X_\Lambda,\, X^\Lambda) \tag{160}$$

which are at least twice as many as the $X_{\hat{\Lambda}}$:

$$\begin{pmatrix} X_{\hat{\Lambda}} \\ 0 \end{pmatrix} = E \begin{pmatrix} X_\Lambda \\ X^\Lambda \end{pmatrix} . \tag{161}$$

This description is clearly redundant and this is the price we have to pay in order to have a manifestly symplectic covariant formalism. We can then rewrite the gauge connection in a symplectic invariant fashion

$$A^{\hat{\Lambda}}\, X_{\hat{\Lambda}} = A^{\hat{\Lambda}}\, E_{\hat{\Lambda}}{}^{\Lambda}\, X_\Lambda + A^{\hat{\Lambda}}\, E_{\hat{\Lambda}\,\Lambda}\, X^\Lambda = A^\Lambda_\mu\, X_\Lambda + A_{\Lambda\,\mu}\, X^\Lambda = A^M_\mu\, X_M, \tag{162}$$

where we have introduced the vector fields A^Λ_μ and the corresponding dual ones $A_{\Lambda\,\mu}$, that can be regarded as components of a symplectic vector

$$A^M_\mu \equiv (A^\Lambda_\mu,\, A_{\Lambda\,\mu}). \tag{163}$$

These are clearly not independent, since they are all expressed in terms of the only electric vector fields $A^{\hat{\Lambda}}$ of our theory (those entering the vector kinetic terms):

$$A^\Lambda_\mu = E_{\hat{\Lambda}}{}^{\Lambda}\, A^{\hat{\Lambda}}_\mu\,, \qquad A_{\Lambda\,\mu} = E_{\hat{\Lambda}\,\Lambda}\, A^{\hat{\Lambda}}_\mu\,. \tag{164}$$

In what follows, it is useful to adopt this symplectic covariant description in terms of $2n_v$ vector fields A^M_μ and $2n_v$ generators X_M, bearing in mind the above definitions through the matrix E, which connects our initial symplectic frame to a generic one.

The components of the symplectic vector X_M are generators in the isometry algebra $\mathfrak{g}$ and thus can be expanded in a basis t_α of generators of G:

$$X_M = \Theta_M{}^\alpha\, t_\alpha\,, \qquad \alpha = 1, \ldots, \dim(G). \tag{165}$$

The coefficients of this expansion $\Theta_M{}^\alpha$ represent an extension of the definition of θ to a G-covariant tensor:

$$\theta_\Lambda{}^\sigma \quad \dashrightarrow \quad \Theta_M{}^\alpha \equiv (\theta^{\Lambda\,\alpha},\ \theta_\Lambda{}^\alpha); \qquad \Theta_M{}^\alpha\ \in \mathscr{R}_{v*} \times \mathrm{adj}(G), \tag{166}$$

which describes the explicit embedding of the gauge group G_g into the global symmetry group G, and combines the full set of deformation parameters of the original ungauged Lagrangian. The advantage of this description is that it allows to recast all the consistency conditions on the choice of the gauge group into G-covariant (and thus independent of the symplectic frame) constraints on Θ.

We should however bear in mind that, just as the redundant set of vectors A^M_μ, also the components of $\Theta_M{}^\alpha$ are not independent since, by (161),

$$\theta_{\hat\Lambda}{}^\alpha = E_{\hat\Lambda}{}^M\, \Theta_M{}^\alpha\ , \qquad 0 = E^{\hat\Lambda\, M}\, \Theta_M{}^\alpha, \tag{167}$$

so that

$$\dim(G_g)\ =\ \mathrm{rank}(\theta) = \mathrm{rank}(\Theta). \tag{168}$$

The above relations (167) imply for $\Theta_M{}^\alpha$ the following symplectic-covariant condition:

$$\Theta_\Lambda{}^\alpha\, \Theta^{\Lambda\,\beta} - \Theta_\Lambda{}^\beta\, \Theta^{\Lambda\,\alpha} = 0 \quad \Leftrightarrow \quad \mathbb{C}^{MN}\Theta_M{}^\alpha\Theta_N{}^\beta = 0. \tag{169}$$

Vice versa, one can show that if $\Theta_M{}^\alpha$ satisfies the above conditions, there exists a symplectic matrix E which can rotate it to an electric frame, namely such that (167) are satisfied for some $\theta_{\hat\Lambda}{}^\alpha$. Equations (169) define the so-called *locality constraint* on the embedding tensor $\Theta_M{}^\alpha$ and they clearly imply:

$$\dim(G_g) = \mathrm{rank}(\Theta) \le n_v\ , \tag{170}$$

which is the preliminary consistency condition (122).

The electric-magnetic duality action of X_M, in the generic symplectic frame defined by the matrix E, is described by the tensor:

$$X_{MN}{}^P\ \equiv\ \Theta_M{}^\alpha\, t_{\alpha\,N}{}^P\ =\ E^{-1}{}_M{}^{\hat M} E^{-1}{}_N{}^{\hat N}\, X_{\hat M\hat N}{}^{\hat P}\, E_{\hat P}{}^P. \tag{171}$$

For each value of the index M, the tensor $X_{MN}{}^P$ should generate symplectic transformations. This implies that:

$$X_{MNP} \equiv X_{MN}{}^Q \mathbb{C}_{QP} = X_{MPN}, \tag{172}$$

which is equivalent to (127). The remaining linear constraints (131), (158) on the gauge algebra can be recast in terms of $X_{MN}{}^P$ in the following symplectic-covariant form:

$$X_{(MNP)} = 0 \quad \Leftrightarrow \quad \begin{cases} 2\,X_{(\Lambda\Sigma)}{}^\Gamma = X^\Gamma{}_{\Lambda\Sigma}, \\ 2\,X^{(\Lambda\Sigma)}{}_\Gamma = X_\Gamma{}^{\Lambda\Sigma}, \\ X_{(\Lambda\Sigma\Gamma)} = 0. \end{cases} \tag{173}$$

Notice that the second of equations (173) implies that, in the electric frame in which $X^{\hat{\Lambda}} = 0$, also the B-block (i.e. the upper-right one) of the infinitesimal gauge generators $\mathscr{R}_v[X_{\hat{\Lambda}}]$ vanishes, being $X_{\hat{\Gamma}}{}^{\hat{\Lambda}\hat{\Sigma}} = 0$, so that the gauge transformations are indeed in G_{el}. Moreover from the first of equation (173), equation (131) follows in the electric frame.

Finally, the closure constraints (130) can be written, in the generic frame, in the following form:

$$[X_M, X_N] = -X_{MN}{}^P X_P \quad \Leftrightarrow \quad \Theta_M{}^\alpha \Theta_N{}^\beta \mathrm{f}_{\alpha\beta}{}^\gamma + \Theta_M{}^\alpha\, t_{\alpha\,N}{}^P \Theta_P{}^\gamma = 0. \tag{174}$$

The above condition can be rephrased, in a G-covariant fashion, as the condition that the *embedding tensor $\Theta_M{}^\alpha$ be invariant under the action of the gauge group it defines*:

$$\delta_M \Theta_N{}^\alpha = 0. \tag{175}$$

Summarizing we have found that consistency of the gauging requires the following set of linear and quadratic algebraic, G-covariant constraints to be satisfied by the embedding tensor:

- *Linear constraint:*

$$X_{(MNP)} = 0, \tag{176}$$

- *Quadratic constraints:*

$$\mathbb{C}^{MN} \Theta_M{}^\alpha \Theta_N{}^\beta = 0, \tag{177}$$

$$[X_M, X_N] = -X_{MN}{}^P X_P. \tag{178}$$

The linear constraint (176) amounts to a projection of the embedding tensor on a specific G-representation $\mathscr{R}_\Theta$ in the decomposition of the product $\mathscr{R}_{v*} \times \mathrm{Adj}(G)$ with respect to G

$$\mathscr{R}_{v*} \times \mathrm{Adj}(G) \quad \xrightarrow{G} \quad \mathscr{R}_\Theta + \ldots \tag{179}$$

and thus can be formally written as follows:

$$\mathbb{P}_\Theta \cdot \Theta = \Theta, \tag{180}$$

where $\mathbb{P}_\Theta$ denotes the projection on the representation $\mathscr{R}_\Theta$. For this reason (176) is also named *representation constraint*.

The first quadratic constraint (177) guarantees that a symplectic matrix E exists which rotates the embedding tensor $\Theta_M{}^\alpha$ to an electric frame in which the *magnetic components* $\Theta^{\hat{\Lambda}\,\alpha}$ vanish. The second one (178) is the condition that the gauge algebra close within the global symmetry one $\mathfrak{g}$ and implies that Θ is a singlet with respect to G_g.

The second part of the gauging procedure, which we are going to discuss below, has to do with restoring supersymmetry after minimal couplings have been introduced and the G_g-invariant Lagrangian $\mathscr{L}^{(0)}_{\text{gauged}}$ have been constructed. As we shall see, the supersymmetric completion of $\mathscr{L}^{(0)}_{\text{gauged}}$ requires no more constraints on G_g (i.e. on Θ) than the linear (176) and quadratic ones (177), (178) discussed above.

As a final remark let us prove that the locality constraint (177) is independent of the others only in theories featuring scalar isometries with no duality action, namely in which the symplectic duality representation $\mathscr{R}_v$ of the isometry algebra $\mathfrak{g}$ is *not faithful*. This is the case of the quaternionic isometries in $N = 2$ theories, see (107) of Sect. 2.5. Let us split the generators t_α of G into t_ℓ, which have a non-trivial duality action, and t_m, which do not:

$$(t_\ell)_M{}^N \neq 0\;; \qquad (t_m)_M{}^N = 0. \tag{181}$$

From (178) we derive, upon symmetrization of the M, N indices, the following condition:

$$X_{(MN)}{}^P\, X_P = X_{(MN)}{}^P\, \Theta_P{}^\alpha\, t_\alpha = 0, \tag{182}$$

where t_α on the right hand side are *not* evaluated in the $\mathscr{R}_v$ representation and thus are all non-vanishing. Using the linear constraint (176) we can then rewrite $X_{(MN)}{}^P$ as follows:

$$X_{(MN)}{}^P = -\frac{1}{2}\,\mathbb{C}^{PQ}\, X_{QMN} = -\frac{1}{2}\,\mathbb{C}^{PQ}\, \Theta_Q{}^\ell t_{\ell\, MN}, \tag{183}$$

so that (182) reads

$$\mathbb{C}^{QP}\, \Theta_Q{}^\ell \Theta_P{}^\alpha\, t_\alpha\, t_{\ell\, MN} = 0. \tag{184}$$

Being t_α and $t_{\ell\, MN}$ independent for any α and ℓ, conditions (176) and (178) only imply *part of* the locality constraint (177):

$$\mathbb{C}^{QP}\, \Theta_Q{}^\ell \Theta_P{}^\alpha = 0, \tag{185}$$

while the remaining constraints (177)

$$\mathbb{C}^{QP}\,\Theta_Q{}^m\Theta_P{}^n = 0, \tag{186}$$

need to be imposed independently. Therefore in theories in which all scalar fields sit in the same supermultiplets as the vector ones, as it is the case of $N > 2$ or $N = 2$ with no hypermultiplets, the locality condition (178) is not independent but follows from the other constraints.

3.2 *The Gauged Lagrangian*

The three steps described above allow us to construct a Lagrangian $\mathscr{L}^{(0)}_{\text{gauged}}$ which is locally G_g-invariant starting from the ungauged one. Now we have to check if this deformation is compatible with local supersymmetry. As it stands, as emphasized above the Lagrangian $\mathscr{L}^{(0)}_{\text{gauged}}$ is no longer invariant under supersymmetry, due to the extra contributions that arise from variation of the vector fields in the covariant derivatives.

Consider, for instance, the supersymmetry variation of the (gauged) Rarita-Schwinger term in the Lagrangian

$$\mathscr{L}_{\text{RS}} = i\,e\,\bar{\psi}^A_\mu\gamma^{\mu\nu\rho}\mathcal{D}_\nu\psi_{A\,\rho}\;+\;\text{h.c.}\,, \tag{187}$$

where $\mathcal{D}_\nu$ is the gauged covariant derivative defined in (154). Under supersymmetry variation of ψ_μ:

$$\delta\psi_\mu = \mathcal{D}_\mu\varepsilon + \cdots\,, \tag{188}$$

ε being the local supersymmetry parameter.[14] The variation of $\mathscr{L}_{\text{RS}}$ produces a term

$$\begin{aligned}\delta\mathscr{L}_{\text{RS}} &= \;\cdots + 2i\,e\,\bar{\psi}^A_\mu\gamma^{\mu\nu\rho}\mathcal{D}_\nu\mathcal{D}_\rho\varepsilon_A\;+\;\text{h.c.}\\ &= \;-\,i\,g\,e\,\bar{\psi}^A_\mu\gamma^{\mu\nu\rho}F^{\hat{\Lambda}}_{\nu\rho}\,(w_{\hat{\Lambda}}\varepsilon)_A\;+\;\text{h.c.}\,,\end{aligned} \tag{189}$$

where we have used the property (142) of the gauge covariant derivative. Similarly we can consider the supersymmetry variation of the spin-1/2 fields:

$$\delta\lambda^{\mathcal{I}} = i\,\hat{\mathcal{P}}^{\mathcal{I}A}_\mu\,\gamma^\mu\varepsilon_A + \ldots, \tag{190}$$

where the dots denote terms containing the vector fields and $\hat{\mathcal{P}}^{\mathcal{I}A}_\mu$ is a specific component of the $\mathfrak{K}$-valued matrix $\hat{\mathcal{P}}_\mu$. The resulting variation of the corresponding kinetic

[14] The ellipses refer to terms containing the vector field strengths.

Lagrangian contains terms of the following form:

$$\delta\left(-i/2e\,\bar{\lambda}_{\mathcal{I}}\gamma^{\mu}\mathcal{D}_{\mu}\lambda^{\mathcal{I}}\,+\,\text{h.c.}\right)=\cdots-i\,e\,\bar{\lambda}_{\mathcal{I}}\gamma^{\mu\nu}\mathcal{D}_{\mu}\hat{\mathcal{P}}_{\nu}^{\mathcal{I}A}\,\varepsilon_{A}\,+\,\text{h.c.}$$
$$=\cdots+i/2g\,e\,\bar{\lambda}_{\mathcal{I}}\gamma^{\mu\nu}F^{\hat{\Lambda}}_{\mu\nu}\,\mathcal{P}^{\mathcal{I}A}_{\hat{\Lambda}}\,\varepsilon_{A}\,+\,\text{h.c.}\qquad(191)$$

We see that the supersymmetry variation of the minimal couplings in the fermion kinetic terms have produced $O(g)$-terms which contain the tensor

$$F^{\hat{\Lambda}}_{\mu\nu}\,L^{-1}X_{\hat{\Lambda}}L=\mathbb{F}^{M}_{\mu\nu}\,L^{-1}X_{M}L\qquad(192)$$

projected on $\mathfrak{H}$ and contracted with the $\bar{\psi}\varepsilon$ current in (189), or restricted to $\mathfrak{K}$ and contracted with the $\bar{\lambda}\varepsilon$ current in the second case (191). On the right hand side of (192) the summation over the gauge generators has been written in the symplectic invariant form defined in (162): $\mathbb{F}^{M}\,X_{M}\equiv F^{\hat{\Lambda}}\,E_{\hat{\Lambda}}{}^{M}\,X_{M}$. These are instances of the various terms occurring in the supersymmetry variation $\delta\mathscr{L}^{(0)}_{\text{gauged}}$. Just as (189) and (191), these terms are proportional to an H-tensor defined as follows[15]:

$$\mathbb{T}(\Theta,\phi)_{\underline{M}}\;\equiv\;\frac{1}{2}\,\mathbb{L}(\phi)^{-1}{}_{\underline{M}}{}^{N}\,L(\phi)^{-1}X_{N}\,L(\phi)\;=\;\frac{1}{2}\,\mathbb{L}(\phi)^{-1}{}_{\underline{M}}{}^{N}\,\Theta_{N}{}^{\beta}\,L(\phi)_{\beta}{}^{\alpha}\,t_{\alpha}$$
$$=\;\mathbb{T}(\Theta,\phi)_{\underline{M}}{}^{\alpha}\,t_{\alpha},\qquad(194)$$

where

$$\mathbb{T}(\Theta,\phi)_{\underline{M}}{}^{\alpha}\equiv\frac{1}{2}\,\mathbb{L}(\phi)^{-1}{}_{\underline{M}}{}^{N}\,\Theta_{N}{}^{\beta}L(\phi)_{\beta}{}^{\alpha}=\frac{1}{2}\,(L^{-1}(\phi)\star\Theta)_{\underline{M}}{}^{\alpha},\qquad(195)$$

where $\star$ denotes the action of L^{-1} as an element of G on $\Theta_{M}{}^{\alpha}$ in the corresponding $\mathscr{R}_{\Theta}$-representation. The tensor $\mathbb{T}(\phi,\,\Theta)=\frac{1}{2}L^{-1}(\phi)\star\Theta$ is called the $\mathbb{T}$-*tensor* and was first introduced in [4].

If Θ and ϕ are simultaneously transformed with G, the $\mathbb{T}$-tensor transforms under the corresponding H-compensator:

$$\forall\mathbf{g}\in G\;:\;\;\mathbb{T}(\mathbf{g}\star\phi,\;\mathbf{g}\star\Theta)=\frac{1}{2}\,L^{-1}(\mathbf{g}\star\phi)\star(\mathbf{g}\star\Theta)$$
$$=\frac{1}{2}\,(h(\mathbf{g},\phi)L^{-1}(\phi)\mathbf{g}^{-1})\star(\mathbf{g}\star\Theta)=h(\mathbf{g},\phi)\star\mathbb{T}(\phi,\;\Theta).\qquad(196)$$

[15] In the formulas below we use the coset representative in which the first index (acted on by G) is in the generic symplectic frame defined by the matrix E and which is then related to the same matrix in the electric frame (labeled by hatted indices) as follows:

$$L(\phi)_{\hat{M}}{}^{\underline{N}}\;=\;E_{\hat{M}}{}^{P}\,L(\phi)_{P}{}^{\underline{N}}\quad\Rightarrow\quad\mathcal{M}(\phi)_{\hat{M}\hat{N}}=E_{\hat{M}}{}^{P}E_{\hat{N}}{}^{Q}\mathcal{M}(\phi)_{PQ},\qquad(193)$$

last equation being (68).

This quantity $\mathbb{T}$ naturally belongs to a representation of the group H and is an example of *composite field* discussed at the end of Sect. 2.6.

If, on the other hand, we fix ϕ and only transform Θ, $\mathbb{T}$ transforms in the same G-representation $\mathscr{R}_\Theta$ as Θ, being $\mathbb{T}$ defined (aside for the factor 1/2) by acting on the embedding tensor with the G-element $\mathbb{L}^{-1}$. As a consequence of this, $\mathbb{T}$ satisfies the same constraints (176), (177) and (178) as Θ:

$$\begin{aligned} \mathbb{T}_{\underline{NM}}{}^{\underline{N}} = \mathbb{T}_{(\underline{MNP})} &= 0, \\ \mathbb{C}^{\underline{MN}}\,\mathbb{T}_{\underline{M}}{}^{\alpha}\,\mathbb{T}_{\underline{N}}{}^{\beta} &= 0, \\ [\mathbb{T}_{\underline{M}},\,\mathbb{T}_{\underline{N}}] + \mathbb{T}_{\underline{MN}}{}^{\underline{P}}\,\mathbb{T}_{\underline{P}} &= 0, \end{aligned} \tag{197}$$

where we have defined $\mathbb{T}_{\underline{MN}}{}^{\underline{P}} \equiv \mathbb{T}_{\underline{M}}{}^{\alpha}\, t_{\alpha\underline{N}}{}^{\underline{P}}$. Equations (197) have been originally derived within maximal supergravity in [4], and dubbed $\mathbb{T}$-*identities*.[16]

Notice that, using (146) and (147) we can rewrite the $\mathbb{T}$-tensor in the following form:

$$\mathbb{T}_{\underline{M}} = \frac{1}{2}\,\mathbb{L}^{-1}{}_{\underline{M}}{}^{N}\,\Theta_N{}^{\alpha}\left(k_\alpha^s\,V_s{}^{\underline{s}}\,K_{\underline{s}} - \frac{1}{2}\,\mathscr{P}_\alpha^{\mathbf{a}}\,J_{\mathbf{a}} - \frac{1}{2}\,\mathscr{P}_\alpha^{\mathbf{m}}\,J_{\mathbf{m}}\right), \tag{198}$$

which can be extended to $N = 2$ theories with non-homogeneous scalar manifolds, see discussion at the end of this section.

To cancel the supersymmetry variations of $\mathscr{L}^{(0)}_{\text{gauged}}$ and to construct a gauged Lagrangian $\mathscr{L}_{\text{gauged}}$ preserving the original supersymmetries, one can apply the general Noether method (see [26] for a general review) which consists in adding new terms to $\mathscr{L}^{(0)}_{\text{gauged}}$ and to the supersymmetry transformation laws, iteratively in the gauge coupling constant. In our case the procedure converges by adding terms of order one ($\Delta\mathscr{L}^{(1)}_{\text{gauged}}$) and two ($\Delta\mathscr{L}^{(2)}_{\text{gauged}}$) in g, so that

$$\mathscr{L}_{\text{gauged}} = \mathscr{L}^{(0)}_{\text{gauged}} + \Delta\mathscr{L}^{(1)}_{\text{gauged}} + \Delta\mathscr{L}^{(2)}_{\text{gauged}}. \tag{199}$$

The additional $O(g)$-terms are of *Yukawa type* and have the general form:

$$\begin{aligned} &e^{-1}\,\Delta\mathscr{L}^{(1)}_{\text{gauged}} \\ &\quad = g\left(2\bar{\psi}^A_\mu\,\gamma^{\mu\nu}\,\psi^B_\nu\,\mathbb{S}_{AB} + i\,\bar{\lambda}^{\mathcal{I}}\,\gamma^\mu\,\psi_{\mu A}\,\mathbb{N}_{\mathcal{I}}{}^{A} + \bar{\lambda}^{\mathcal{I}}\,\lambda^{\mathcal{J}}\,\mathbb{M}_{\mathcal{I}\mathcal{J}}\right) + \text{h.c.}, \end{aligned} \tag{200}$$

characterized by the scalar-dependent matrices $\mathbb{S}_{AB}$ and $\mathbb{N}^{\mathcal{I}A}$ called *fermion shift matrices*, and a matrix $\mathbb{M}^{\mathcal{I}\mathcal{J}}$ that can be rewritten in terms of the previous mixed mass tensor $\mathbb{N}^{\mathcal{I}A}$ (see the subsequent sections).

The $O(g^2)$-terms consist of a scalar potential:

$$e^{-1}\,\Delta\mathscr{L}^{(2)}_{\text{gauged}} = -g^2\,V(\phi). \tag{201}$$

---

[16] Recall that in maximal supergravity the locality constraint follows from the linear and the closure ones.

At the same time the fermionic supersymmetry transformations need to be suitably modified. To this end, we shall *add order–g terms to the fermion supersymmetry transformation rules* of the gravitino ($\psi_{\mu A}$) and of the other fermions ($\chi^{\mathcal{I}}$)

$$\begin{aligned} \delta_\varepsilon \psi_{\mu A} &= \mathcal{D}_\mu \varepsilon_A + i\, g\, \mathbb{S}_{AB}\, \gamma_\mu\, \varepsilon^B + \cdots\,, \\ \delta_\varepsilon \lambda_{\mathcal{I}} &= g\, \mathbb{N}_{\mathcal{I}}{}^A\, \varepsilon_A + \cdots \end{aligned} \tag{202}$$

depending on the same matrices $\mathbb{S}_{AB}$, $\mathbb{N}_{\mathcal{I}}{}^A$ entering the mass terms. The fermion shift-matrices are composite fields belonging to some appropriate representations $\mathscr{R}_S$, $\mathscr{R}_N$ of the H group, such that (200) is H-invariant.

These additional terms in the Lagrangian and supersymmetry transformation laws are enough to cancel the original $O(g)$ variations in $\delta\mathscr{L}^{(0)}_{\text{gauged}}$—like (189) and (191), together with new $O(g)$ terms depending on $\mathbb{S}$ and $\mathbb{N}$ in the supersymmetry variation of $\mathscr{L}^{(0)}_{\text{gauged}}$—provided the shift-tensors $\mathbb{S}_{AB}$, $\mathbb{N}^{\mathcal{I}A}$ are identified with suitable H-covariant components of the $\mathbb{T}$-tensor:

$$\mathscr{R}_\Theta \;\xrightarrow{H}\; \mathscr{R}_N + \mathscr{R}_S + \mathscr{R}_{\text{other}}, \tag{203}$$

and that additional H-representations $\mathscr{R}_{\text{other}}$ in the $\mathbb{T}$-tensor do not enter the supersymmetry variations of the Lagrangian. This can be formulated as a G-covariant restriction on the representation $\mathscr{R}_\Theta$ of the $\mathbb{T}$-tensor or, equivalently, of embedding tensor, which can be shown to be no more than the representation constraint (176) discussed earlier.

The identification with components of the $\mathbb{T}$-tensor defines the expression fermion shift-tensors as H-covariant composite fields in terms of the embedding tensor and the scalar fields:

$$\mathbb{S}_{AB} = \mathbb{S}_{AB}(\phi, \Theta) = \mathbb{T}(\phi, \Theta)|_{\mathscr{R}_S}\,; \quad \mathbb{N}_{\mathcal{I}}{}^A = \mathbb{N}_{\mathcal{I}}{}^A(\phi, \Theta) = \mathbb{T}(\phi, \Theta)|_{\mathscr{R}_N}\,. \tag{204}$$

Finally, in order to cancel the $O(g^2)$-contributions resulting from the variations (202) in (200), we need to add an *order-g^2 scalar potential $V(\phi)$* whose expression is totally determined by supersymmetry as a bilinear in the shift matrices by the condition

$$\delta_B{}^A\, V(\phi) \;=\; g^2\, \left(\mathbb{N}_{\mathcal{I}}{}^A\, \mathbb{N}^{\mathcal{I}}{}_B - 12\, \mathbb{S}^{AC}\, \mathbb{S}_{BC}\right), \tag{205}$$

where we have defined $\mathbb{N}^{\mathcal{I}}{}_A \equiv (\mathbb{N}_{\mathcal{I}}{}^A)^*$ and $\mathbb{S}^{AB} \equiv (\mathbb{S}_{AB})^*$. The above condition is called *potential Ward identity* [27, 28] (for a comprehensive discussion of the supersymmetry constraints on the fermion shifts see [29]). This identity defines the scalar potential as a quadratic function of the embedding tensor and non-linear function of the scalar fields. As a constraint on the fermion shifts, once these have been identified with components of the $\mathbb{T}$-tensor, it follows from the $\mathbb{T}$-identities (197) or, equivalently, from the quadratic constraints (177), (178) on Θ. The derivation of quadratic supersymmetry constraints on the fermion shifts in maximal supergravity from algebraic constraints (i.e. scalar field independent) on the embedding tensor, was originally accomplished in [16], though in a specific symplectic frame, and in maximal $D = 3$ theory in [17]. In [5] the four-dimensional result was extended to a generic symplectic frame of the $N = 8$ model, i.e. using the G-covariant constraint (176), (177), (178) on the embedding tensor.[17]

Let us comment on the case of $N = 2$ theories with a non-homogeneous scalar manifold (106). In this case we cannot define a coset representative. However, as mentioned earlier, one can still define a symplectic matrix $\mathbb{L}^M{}_{\underline{N}}$ depending on the complex scalar fields in the vector multiplets (which has no longer the interpretation of a coset representative). We can then define the $\mathbb{T}$-tensor in these theories as in (198) where $\{K_{\underline{s}}\}$ should be intended as a basis of the tangent space to the origin (and not as isometry generators), while $\{J_I\} = \{J_{\mathbf{a}},\ J_{\mathbf{m}}\}$ are holonomy group generators.[18] Recall that $\{\mathscr{P}_\alpha^{\mathbf{a}},\ \mathscr{P}_\alpha^{\mathbf{m}}\}$ enter the definition of the gauged composite connection (147) on the scalar manifold and, as mentioned earlier, are related to the Killing vectors by general properties of the spacial Kähler and quaternionic Kähler geometries [19].

It is a characteristic of supergravity theories that—in contrast to globally supersymmetric ones—by virtue of the negative contribution due to the gravitino shift-matrix, the scalar potential is in general not positive definite, but may, in particular, feature AdS vacua. These are maximally symmetric solutions whose negative cosmological constant is given by the value of the potential at the corresponding extremum: $\Lambda = V_0 < 0$. Such vacua are interesting in the light of the AdS/CFT holography conjecture [30], according to which stable AdS solutions describe conformal critical points of a suitable gauge theory defined on the boundary of the space. In this perspective, domain wall solutions to the gauged supergravity interpolating between AdS critical points of the potential describe renormalization group (RG) flow (from an ultra-violet to an infra-red fixed point) of the dual gauge theory and give important insights into its non-perturbative properties. The spatial evolution of such holographic flows is determined by the scalar potential $V(\phi)$ of the gauged theory.

In some cases the effective scalar potential $V(\phi)$, at the classical level, is non–negative and defines vacua with vanishing cosmological constant in which supersymmetry is spontaneously broken and part of the moduli are fixed. Models of this type are generalizations of the so called “no–scale” models [31–33] which were subject to intense study during the eighties.

---

[17]In a generic gauged model, supersymmetry further require the fermion shifts to be related by differential “gradient flow” relations [29] which can e shown to follow from the identification of the shifts with components of the $\mathbb{T}$-tensor and the geometry of the scalar manifold.

[18]The $H_{\mathrm{R}} = \mathrm{U}(2)$-generators $\{J_{\mathbf{a}}\}$ naturally split into a U(1)-generator J_0 of the Kähler transformations on $\mathscr{M}_{\mathrm{SK}}$ and SU(2)-generators J_x ($x = 1, 2, 3$) in the holonomy group of the quaternionic Kähler manifold $\mathscr{M}_{\mathrm{QK}}$.

3.3 Dualities and Flux Compactifications

Let us summarize what we have learned so far.

- The most general local internal symmetry group G_g which can be introduced in an extended supergravity is defined by an embedding tensor Θ, covariant with respect to the on-shell global symmetry group G of the ungauged model and defining the embedding of G_g inside G. Since a scalar potential $V(\phi)$ can only be introduced through the gauging procedure, Θ also defines the most general choice for $V = V(\phi, \Theta)$.
- Consistency of the gauging at the level of the bosonic action requires Θ to satisfy a number of (linear and quadratic) G-covariant constraints. The latter, besides completely determining the gauged bosonic action, also allow for its consistent (unique) supersymmetric extension.
- Once we find a solution $\Theta_M{}^\alpha$ to these algebraic constraints, a suitable symplectic matrix E, which exists by virtue of (177), will define the corresponding electric frame, in which its magnetic components vanish.

Although we have freed our choice of the gauge group from the original symplectic frame, the resulting gauged theory is still defined in an electric frame and thus depends on the matrix E: whatever solution Θ to the constraints is chosen for the gauging, the kinetic terms of the gauged Lagrangian are always written in terms of the only *electric* vector fields $A^{\hat{\Lambda}}_\mu$, namely of the vectors effectively involved in the minimal couplings, see (162). We shall discuss in the next section a more general formulation of the gauging which no longer depends on the matrix E.

Dual gauged supergravities. All the deformations of the ungauged model required by the gauging procedure depend on Θ in a manifestly G-covariant way. This means that, if we transform all the fields Φ (bosons and fermions) of the model under G (the fermions transforming under corresponding compensating transformations in H) and at the same time transform Θ and the matrix E, the field equations and Bianchi identities—collectively denoted by $\mathscr{E}(E,\, \Phi,\, \Theta) = 0$—are left invariant:

$$\forall g \in G \;:\; \mathscr{E}(E,\, \Phi,\, \Theta) = 0 \quad \Leftrightarrow \quad \mathscr{E}(E',\, g \star \Phi,\, g \star \Theta) = 0$$
$$(\text{with } E' = E\, \mathscr{R}_v[g]^T). \tag{206}$$

Since the embedding tensor Θ is a *spurionic*, namely non-dynamical, object, the above on-shell invariance should not be regarded as a symmetry of a single theory, but rather as an equivalence (or proper duality) between two different theories, one defined by Θ and the other by $g \star \Theta$. Gauged supergravities are therefore classified in *orbits* with respect to the action of G (or better $G(\mathbb{Z})$) on Θ. This property has an important bearing on the study of flux compactifications mentioned in the Introduction. Indeed, in all instances of flux compactifications, the internal fluxes manifest themselves in the lower-dimensional effective gauged supergravity as components of the embedding tensor defining the gauging [6, 34, 35]:

$$\Theta = \text{Internal Fluxes.} \tag{207}$$

This allows us to formulate a precise correspondence between general fluxes (form, geometric and non-geometric) and the gauging of the resulting supergravity. Moreover, using this identification, the quadratic constraints (177), (178) precisely reproduce the consistency conditions on the internal fluxes deriving from the Bianchi identities and field equations in the higher dimensional theory such as, in the presence of RR fluxes, the tadpole cancelation condition [6, 13, 34].

Consider the limit in which the lower-dimensional gauged theory provides a reliable description of the low-energy string or M-theory dynamics on a flux background. This limit is defined by the condition that the flux-induced masses in the effective action be much smaller than the scale of the Kaluza–Klein masses (of order $1/R$, where R is the size of the internal manifold)[19]:

$$\text{Flux-induced masses} \ll \frac{1}{R}. \tag{208}$$

In this case, fields and fluxes in the lower-dimensional supergravity arrange in representations with respect to the characteristic symmetry group G_{int} the internal manifold would have in the absence of fluxes. In the case of compactifications on T^n, such characteristic group is $\mathrm{GL}(n, \mathbb{R})$, acting transitively on the internal metric moduli.

In general, in the absence of fluxes, G_{int} is a global symmetry group of the action: $G_{int} \subset G_{el}$. By branching $\mathscr{R}_\Theta$ with respect to G_{int}, we can identify within Θ the components corresponding to the various internal fluxes. The effect of any such background quantities in the compactification is reproduced by simply switching on the corresponding components of Θ. The gauging procedure does the rest and the resulting gauged model is thus uniquely determined. Since, as mentioned earlier at the end of Sect. 2.4, a suitable subgroup $G(\mathbb{Z})$ of G was conjectured to encode all known string/M-theory dualities, the embedding tensor formulation of the gauging procedure provides an ideal theoretical laboratory where to systematically study the effects of these dualities on fluxes. Some elements of $G(\mathbb{Z})$ will map gauged supergravity descriptions of known compactifications into one another, see Fig. 1.

Other elements of $G(\mathbb{Z})$ will map gauged supergravities, originating from known compactifications, into theories whose string or M-theory origin is unknown, see Fig. 2.

In this case we can use the duality between the corresponding low-energy descriptions to make sense of new compactifications as "dual" to known ones.

The so-called *non-geometric* fluxes naturally fit in the above description as dual to certain compactifications with NS-NS H-flux. If we consider superstring theory compactified to four-simensions on a six-torus T^6 without fluxes, the resulting (classical) ungauged supergravity features a characteristic $\mathrm{O}(6, 6)$ global symmetry group, which contains the T-duality group $\mathrm{O}(6, 6; \mathbb{Z})$ and which acts transitively on the

[19]For string theory compactifications we should also require this latter scale to be negligible compared to the mass-scale of the string excitations (order $1/\sqrt{\alpha'}$).

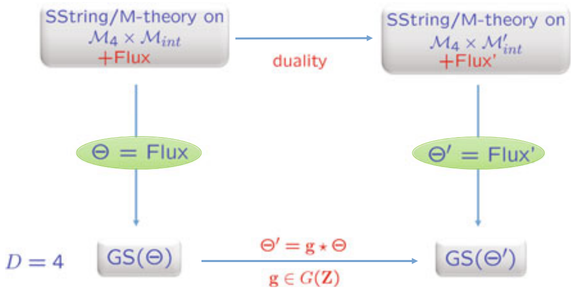


Fig. 1 Dualities between known flux compactifications ("GS" stands for "gauged supergravity")

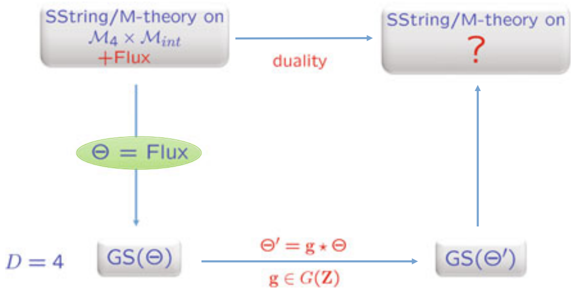


Fig. 2 Dualities connecting known flux compactifications to unknown ones

moduli originating from the metric and Kalb-Ramond B-field in ten dimensions. The G-representation $\mathscr{R}_\Theta$ of the embedding tensor, defining the most general gauging, contains the representation **220** of O(6, 6)

$$\mathscr{R}_\Theta \xrightarrow{\mathrm{O}(6,6)} \mathbf{220} + \ldots \tag{209}$$

which in turn branches with respect to the characteristic group $G_{int} = \mathrm{GL}(6, \mathbb{R})$ of the torus as follows:

$$\mathbf{220} \xrightarrow{\mathrm{GL}(6,\mathbb{R})} \mathbf{20}_{-3} + (\mathbf{84} + \mathbf{6})_{-1} + (\mathbf{84}' + \mathbf{6}')_{+1} + \mathbf{20}_{+3}. \tag{210}$$

The component $\mathbf{20}_{-3}$ can be identified with the H-flux $H_{\alpha\beta\gamma}$ (that is the flux of the field strength of the Kalb-Ramond field B) along a 3-cycle of the torus. Switching on only the $\mathbf{20}_{-3}$ representation in Θ, the gauging procedure correctly reproduces the couplings originating from a toroidal dimensional reduction with H-flux. What (210) tells us is that the action of the T-duality group O(6, 6; $\mathbb{Z}$) will generate, from an H-flux in the $\mathbf{20}_{-3}$, all the other representations:

$$\begin{aligned} (\mathbf{84} + \mathbf{6})_{-1} &: \tau_{\alpha\beta}{}^{\gamma}, \\ (\mathbf{84}' + \mathbf{6}')_{+1} &: Q_\alpha{}^{\beta\gamma}, \\ \mathbf{20}_{+3} &: R^{\alpha\beta\gamma}. \end{aligned} \tag{211}$$

The first tensor $\tau_{\alpha\beta}{}^{\gamma}$ is an instance of *geometric flux*, being a background quantity which characterizes the geometry of the internal manifold. It describes a compactification on a space which is no longer a torus, but is locally described by a group manifold [36] with structure constants $\tau_{\alpha\beta}{}^{\gamma}$. The constraint (178) indeed implies for $\tau_{\alpha\beta}{}^{\gamma}$ the Jacobi identity: $\tau_{[\alpha\beta}{}^{\gamma}\tau_{\sigma]\gamma}{}^{\delta} = 0$. This new internal manifold is called *twisted torus* [37] (see also [13] and references therein).

The T-duality picture is completed by the remaining two representations, described by the tensors $Q_{\alpha}{}^{\beta\gamma}$, $R^{\alpha\beta\gamma}$. Their interpretation as originating from a string theory compactification is more problematic, since in their presence the internal space cannot be given a global or even local description as a differentiable manifold. For this reason they are called *non-geometric* fluxes [38–40] (see also [13] and references therein). The H, τ, Q, R-fluxes can all be given a unified description as quantities defining the geometry of more general internal manifolds, having the T-duality group as structure group. Such manifolds are defined in the context of *generalized geometry* [41, 42] (see also [13] and references therein), by doubling the tangent space to the internal manifold in order to accommodate a representation of O(6, 6) and introducing on it additional geometric structures, or of *double geometry/double field theory* [2, 43–45], in which the internal manifold itself is enlarged, and parametrized by twice as many coordinates as the original one.

Finally there are gauged supergravities which are not $G(\mathbb{Z})$-dual to models with a known string or M-theory origin, Fig. 3. Finding an ultra-violet completion of these theories, which are sometimes called *intrinsically non-geometric*, in the context of string/M-theory is an open challenge of theoretical high-energy physics. Progress in this direction has been achieved in the context of extended generalized geometry [46, 47] or exceptional field theory [1, 48, 49].

If the hierarchy condition (208) is not met, the gauged supergravity cannot be intended as a description of the low-energy string/M-theory dynamics, but just as a *consistent truncation* of it, as in the case of the spontaneous compactification of $D = 11$ supergravity on $AdS_4 \times S^7$. In this case, the back-reaction of the fluxes on the internal geometry will manifest in extra geometric fluxes, to be identified with additional components of Θ.

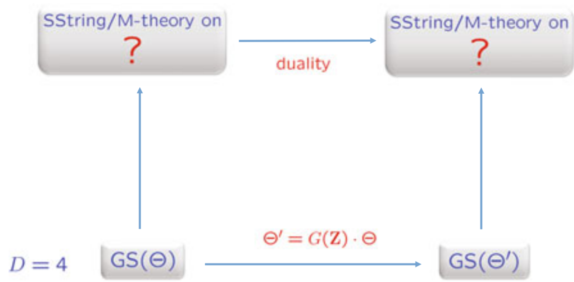


Fig. 3 Intrinsically non-geometric theories

Vacua and dualities. The scalar potential

$$V(\phi, \Theta) = \frac{g^2}{\mathcal{N}} \left(\mathbb{N}_{\mathcal{I}}{}^{A} \, \mathbb{N}^{\mathcal{I}}{}_{B} - 12 \, \mathbb{S}^{AC} \, \mathbb{S}_{BC} \right), \tag{212}$$

being expressed as an H-invariant combination of composite fields (the fermion shifts), is invariant under the simultaneous action of G on Θ and ϕ^s:

$$\forall g \in G \; : \quad V(g \star \phi, \; g \star \Theta) = V(\phi, \; \Theta). \tag{213}$$

This means that, if $V(\phi, \; \Theta)$ has an extremum in ϕ_0

$$\frac{\partial}{\partial \phi^s} V(\phi, \; \Theta) \bigg|_{\phi_0} = 0, \tag{214}$$

$V(\phi, g \star \Theta)$ has an extremum at $\phi_0' = g \star \phi_0$ with the same properties (value of the potential at the extremum and its derivatives):

$$\frac{\partial}{\partial \phi^s} V(\phi, g \star \Theta) \bigg|_{g \star \phi_0} = 0, \qquad g \in G \, . \tag{215}$$

If the scalar manifold is homogeneous, we can map any point ϕ_0 to the origin $\mathcal{O}$, where all scalars vanish, by the inverse of the coset representative $L(\phi_0)^{-1} \in G$. We can then map a generic vacuum ϕ_0 of a given theory (defined by an embedding tensor Θ) to the origin of the theory defined by $\Theta' = L(\phi_0)^{-1} \star \Theta$. As a consequence of this, when looking for vacua with given properties (residual (super)symmetry, cosmological constant, mass spectrum etc.), with no loss of generality we can compute all quantities defining the gauged theory—fermion shifts and mass matrices—at the origin:

$$\mathbb{N}(\mathcal{O}, \; \Theta), \;\; \mathbb{S}(\mathcal{O}, \; \Theta), \;\; \mathbb{M}(\mathcal{O}, \; \Theta), \tag{216}$$

and translate the properties of the vacuum in conditions on Θ. In this way, we can search for the vacua by scanning though all possible gaugings [50–52].

3.4 Gauging $N = 8$, $D = 4$

Ungauged action. The four dimensional maximal supergravity is characterized by having $N = 8$ supersymmetry (that is 32 supercharges), which is the maximal amount of supersymmetry allowed by a consistent theory of gravity.

We shall restrict ourselves to the (ungauged) $N = 8$ theory with no antisymmetric tensor field—which would eventually be dualized to scalars. The theory, firstly constructed in [53, 54], describes a single massless graviton supermultiplet consisting of the graviton $g_{\mu\nu}$, 8 spin-3/2 gravitini ψ_μ^A ($A = 1, \dots, 8$) transforming in the fundamental representation of the R–symmetry group SU(8), 28 vector fields A_μ^Λ (with $\Lambda = 0, \dots, 27$), 56 spin-1/2 dilatini χ_{ABC} in the **56** of SU(8) and 70 real scalar fields ϕ^r:

$$\Big[\ 1 \times \underbrace{g_{\mu\nu}}_{j=2}\,,\quad 8 \times \underbrace{\psi_\mu^A}_{j=\frac{3}{2}}\,,\quad 28 \times \underbrace{A_\mu^\Lambda}_{j=1}\,,\quad 56 \times \underbrace{\chi_{ABC}}_{j=\frac{1}{2}}\,,\quad 70 \times \underbrace{\phi^r}_{j=0}\ \Big]. \tag{217}$$

The scalar fields are described by a non-linear σ-model on the Riemannian manifold $\mathscr{M}_{\rm scal}$, that in the $N = 8$ model has the form

$$\mathscr{M}_{\rm scal} = \frac{G}{H} = \frac{{\rm E}_{7(7)}}{{\rm SU}(8)}, \tag{218}$$

the isometry group being $G = {\rm E}_{7(7)}$, and $H = {\rm SU}(8)$ being the R–symmetry group. The bosonic Lagrangian has the usual form (5). The global symmetry group of the maximal four-dimensional theory $G = {\rm E}_{7(7)}$ has 133 generators t_α. The (abelian) vector field strengths $F^\Lambda = dA^\Lambda$ and their magnetic duals $\mathcal{G}_\Lambda$ together transform in the $\mathscr{R}_v = \mathbf{56}$ fundamental representation of the ${\rm E}_{7(7)}$ duality group with generators $(t_\alpha)_M{}^N$, so that

$$\delta\mathbb{F}^M_{\mu\nu} = \begin{pmatrix}\delta F^\Lambda_{\mu\nu}\\ \delta\mathcal{G}_{\Lambda\,\mu\nu}\end{pmatrix} = -\,\Lambda^\alpha\,(t_\alpha)_N{}^M\,\mathbb{F}^N_{\mu\nu}\,. \tag{219}$$

Gauging. According to our general discussion of Sect. 3.1, the most general gauge group G_g which can be introduced in this theory is defined by an embedding tensor $\Theta_M{}^\alpha$ ($M = 1, \dots, 56$ and $\alpha = 1, \dots, 133$), which expresses the gauge generators X_M as linear combinations of the global symmetry group ones t_α (165). The embedding tensor encodes all parameters (couplings and mass deformations) of the gauged theory. This object is solution to the G-covariant constraints (176), (177), (178).

The embedding tensor formally belongs to the product

$$\Theta_M{}^\alpha \in \mathscr{R}_v \otimes {\rm adj}(G) = \mathbf{56} \otimes \mathbf{133} = \mathbf{56} \oplus \mathbf{912} \oplus \mathbf{6480}. \tag{220}$$

The linear constraint (176) sets to zero all the representation in the above decomposition which are contained in the 3-fold symmetric product of the **56** representation:

$$X_{(MNP)} \in (\mathbf{56} \otimes \mathbf{56} \otimes \mathbf{56})_{\rm sym.} \to \mathbf{56} \oplus \mathbf{6480} \oplus \mathbf{24320}. \tag{221}$$

The representation constraint therefore selects the **912** as the representation $\mathscr{R}_\Theta$ of the embedding tensor.[20]

[20] We can relax this constraint by extending this representation to include the **56** in (220). Consistency however would require the gauging of the scaling symmetry of the theory (which is never an off-shell symmetry), also called *trombone symmetry* [55, 56]. This however leads to gauged theories which do not have an action. We shall not discuss these gaugings here.

The quadratic constraints pose further restrictions on the $E_{7(7)}$-orbits of the **912** representation which $\Theta_M{}^\alpha$ should belong to. In particular the locality constraint implies that the embedding tensor can be rotated to an electric frame through a suitable symplectic matrix E, see (167).

Steps 1, 2 and 3 allow to construct the bosonic gauged Lagrangian in this electric frame. We shall discuss in Sect. 4 a frame-independent formulation of the gauging procedure in which, for a given solution Θ to the constraints, we no longer need to switch to the corresponding electric frame.

The complete supersymmetric gauged Lagrangian is then obtained by adding fermion mass terms, a scalar potential and additional terms in the fermion supersymmetry transformation rules, according to the prescription given in Step 4. All these deformations depend on the fermion shift matrices $\mathbb{S}_{AB}$, $\mathbb{N}_{\mathcal{I}}{}^A$. In the maximal theory $\mathcal{I} = [ABC]$ labels the spin-1/2 fields χ_{ABC} and the two fermion shift-matrices are conventionally denoted by the symbols $A_1 = (A_{AB})$, $A_2 = (A^D{}_{ABC})$. The precise correspondence is[21]:

$$\mathbb{S}_{AB} = -\frac{1}{\sqrt{2}} A_{AB}; \quad \mathbb{N}_{ABC}{}^D = -\sqrt{2} A^D{}_{ABC}, \tag{223}$$

where

$$A_{AB} = A_{BA}; \quad A_{ABC}{}^D = A_{[ABC]}{}^D; \quad A_{DBC}{}^D = 0. \tag{224}$$

The above properties identify the SU(8) representations of the two tensors:

$$A_{AB} \in \mathbf{36}; \quad A_{ABC}{}^D \in \mathbf{420}. \tag{225}$$

[21] In the previous sections we have used, for the supergravity fields, notations which are different from those used in the literature of maximal supergravity (e.g. in [18]) in order to make contact with the literature of gauged $N < 8$ theories, in particular $N = 2$ ones [19]. Denoting by a hat the quantities in [18], the correspondence between the two notations is:

$$\begin{aligned}
&\hat{\gamma}^\mu = i\gamma^\mu; \quad \hat{\gamma}_5 = \gamma_5,\\
&\hat{\varepsilon}_i = \frac{1}{\sqrt{2}}\,\varepsilon^A; \quad \hat{\varepsilon}^i = \frac{1}{\sqrt{2}}\,\varepsilon_A; \quad (i = A),\\
&\hat{\psi}_{i\mu} = \sqrt{2}\,\psi^A_\mu; \quad \hat{\psi}^i_\mu = \sqrt{2}\,\psi_{A\,\mu}; \quad (i = A),\\
&\hat{\chi}_{ijk} = \chi^{ABC}; \quad \hat{\chi}^{ijk} = \chi_{ABC}; \quad ([ijk] = [ABC]),\\
&\hat{A}_{ij} = (\hat{A}_{ij})^* = A^{AB}; \quad \hat{A}_i{}^{jkl} = (\hat{A}^i{}_{jkl})^* = A^A{}_{BCD}; \quad (i = A,\, j = B,\, k = C,\, l = D),\\
&\mathcal{V}^{\Lambda\, ij} = -\frac{i}{\sqrt{2}}\,\mathbb{L}^\Lambda{}_{AB}; \quad \mathcal{V}_\Lambda{}^{ij} = \frac{i}{\sqrt{2}}\,\mathbb{L}_{\Lambda AB}; \quad (i = A,\, j = B),
\end{aligned} \tag{222}$$

where in the last line the 28 × 28 blocks of $\mathcal{V}_M{}^{\underline{N}}$ have been put in correspondence with those of $\mathbb{L}^M{}_{\underline{N}}$. The factor $\sqrt{2}$ originates from a different convention with the contraction of antisymmetric couples of SU(8)-indices: $\hat{V}_{ij}\hat{V}^{ij} = \frac{1}{2}\, V^{AB}\, V_{AB}$.

The $\mathbb{T}$-tensor, defined in (194) as an $E_{7(7)}$-object, transforms in $\mathscr{R}_\Theta = \mathbf{912}$, while as an SU(8)-tensor it belongs to the following sum of representations:

$$\mathbb{T} \in \mathbf{912} \xrightarrow{\mathrm{SU}(8)} \mathbf{36} \oplus \overline{\mathbf{36}} \oplus \mathbf{420} \oplus \overline{\mathbf{420}}\,, \tag{226}$$

which are precisely the representations of the fermion shift-matrices and their conjugates $A_{AB}\,A^{AB},\ A^A{}_{BCD},\ A_A{}^{BCD}$. This guarantees that the $O(g)$-terms in the supersymmetry variation of $\mathscr{L}^{(0)}_{\mathrm{gauged}}$, which depend on the $\mathbb{T}$-tensor, only contain SU(8)-structures which can be canceled by the new terms containing the fermion shift-matrices. This shows that the linear condition $\Theta \in \mathscr{R}_\Theta$ is also required by supersymmetry.

The same holds for the quadratic constraints, in particular for (178), which implies the $\mathbb{T}$-identities and also the Ward identity (205) for the potential [4, 18]:

$$V(\phi)\,\delta_A^B = \frac{g^2}{6}\,\mathbb{N}^{CDE}{}_A\mathbb{N}_{CDE}{}^B - 12\,g^2\,\mathbb{S}_{AC}\mathbb{S}^{BC} = \frac{g^2}{3}A^B{}_{CDE}A_A{}^{CDE} - 6\,g^2\,A_{AC}\,A^{BC}, \tag{227}$$

from which we derive:

$$V(\phi) = g^2\left(\frac{1}{24}\,|A^B{}_{CDE}|^2 - \frac{3}{4}\,|A_{AB}|^2\right). \tag{228}$$

The scalar potential can also be given in a manifestly G-invariant form [18]:

$$V(\phi) = -\frac{g^2}{672}\left(X_{MN}{}^R\,X_{PQ}{}^S\,\mathcal{M}^{MP}\,\mathcal{M}^{NQ}\,\mathcal{M}_{RS} + 7\,X_{MN}{}^Q\,X_{PQ}{}^N\,\mathcal{M}^{MP}\right), \tag{229}$$

where $\mathcal{M}^{MN}$ is the inverse of the (negative definite) matrix $\mathcal{M}_{MN}$ defined in (55) and, as usual, $X_{MN}{}^R$ describe the symplectic duality action of the generators X_M in the $\mathscr{R}_{v*}$-representation: $X_{MN}{}^R \equiv \mathscr{R}_{v*}[X_M]_N{}^R$.

3.5 Brief Account of Old and New Gaugings

As mentioned in Sect. 3.1, different symplectic frames (i.e. different ungauged Lagrangians) correspond to different choices for the viable gauge groups and may originate from different compactifications (see [5] for a study of the different symplectic frames for the ungauged maximal theory).

The toroidal compactification of eleven dimensional theory performed in [53], upon dualization of all form-fields to lower order ones, yields an ungauged Lagrangian with global symmetry $G_{el} = \mathrm{SL}(8,\mathbb{R})$. We shall refer to this symplectic frame as the $\mathrm{SL}(8,\mathbb{R})$-frame. The first gauging of the maximal theory was performed in this symplectic frame by choosing $G_g = \mathrm{SO}(8) \subset \mathrm{SL}(8,\mathbb{R})$ [4]. The scalar potential features a maximally supersymmetric anti-de Sitter vacuum which corresponds [57] to the spontaneous compactification of eleven dimensional supergravity on $AdS_4 \times S^7$. The range of possible gaugings in the $\mathrm{SL}(8,\mathbb{R})$-frame was extended to include non-compact and non semisimple groups $G_g = \mathrm{CSO}(p,q,r)$ (with $p+q+r=8$) [10]. These were shown in [16] to exhaust all possible gaugings in this frame.

The discovery of inequivalent Lagrangian formulations of the ungauged maximal theory broadened the choice of possible gauge groups. Flat-gaugings in $D=4$ describing Scherk-Schwarz reductions of maximal $D=5$ supergravity [58] and yielding no-scale models, were first constructed in [59]. The corresponding symplectic frame is the one originating from direct dimensional reduction of the maximal five-dimensional theory on a circle and has a manifest off-shell symmetry which contains the global symmetry group of the parent model[22] $\mathrm{E}_{6(6)}$: one has in fact $G_{el} = \mathrm{O}(1,1) \times \mathrm{E}_{6(6)}$.

Exploiting the freedom in the initial choice of the sympectic frame, it was recently possible to discover a new class of gauging generalizing the original $\mathrm{CSO}(p,q,r)$ ones [60–62]. These models are obtained by gauging, in a different frame, the same $\mathrm{CSO}(p,q,r)$.

Consider two inequivalent frames admitting $G_g = \mathrm{CSO}(p,q,r)$ as gauge group, namely for each of which $\mathrm{CSO}(p,q,r) \subset G_{el}$. Let $\hat{\mathscr{R}}_v$ and $\mathscr{R}_v$ be the corresponding symplectic duality representations of G. We can safely consider one of them ($\hat{\mathscr{R}}_v$) as electric. The duality action of the gauge generators $\hat{\mathscr{R}}_{v*}$ and $\mathscr{R}_{v*}$ are described by two tensors $X_{\hat{M}\hat{N}}{}^{\hat{P}}$ and $X_{MN}{}^P$, respectively, related by a suitable matrix E (171):

$$X_{\hat{M}\hat{N}}{}^{\hat{P}} = E_{\hat{M}}{}^M \, E_{\hat{N}}{}^N \, (E^{-1})_P{}^{\hat{P}} \, X_{MN}{}^P . \tag{230}$$

The matrices $\mathcal{M}(\phi)$ in the two frames are then related by (68). The two embedding tensors describe the same gauge group provided that $\{X_M\}$ and $\{E\,X_M\,E^{-1}\}$ define different bases of the same gauge algebra $\mathfrak{g}_g = \mathfrak{cso}(p,q,r)$ in the Lie algebra $\mathfrak{e}_{7(7)}$ of $\mathrm{E}_{7(7)}$. In other words, E should belong to the *normalizer* of $\mathfrak{cso}(p,q,r)$ in $\mathrm{Sp}(2n_v,\mathbb{R})$. At the same time the effect of E should not be offset by local (vector and scalar field) redefinitions, see (69). The duality action of G_g in both $\hat{\mathscr{R}}_{v*}$ and $\mathscr{R}_{v*}$ is block-diagonal:

$$\hat{\mathscr{R}}_{v*}[G_g] = \mathscr{R}_{v*}[G_g] = \begin{pmatrix} G_g & \mathbf{0} \\ \mathbf{0} & G_g^{-T} \end{pmatrix} . \tag{231}$$

For semisimple gauge groups $G_g = \mathrm{SO}(p,q)$ (with $p+q=8$), it was shown in [62] that the most general E belongs to an $\mathrm{SL}(2,\mathbb{R})$-subgroup of $\mathrm{Sp}(56,\mathbb{R})$ and has the general form:

$$E = \begin{pmatrix} a\,\mathbf{I} & b\,\eta \\ c\,\eta & d\,\mathbf{I} \end{pmatrix} \in \mathrm{Sp}(56,\mathbb{R}) \, ; \quad ad - bc = 1 , \tag{232}$$

[22] See Table 2 at the end of Sect. 4.

where $\eta_{\Lambda\Sigma}$ is the $\mathfrak{so}(p,q)$-Cartan Killing metric, normalized so that $\eta^2 = \mathbf{I}$. The most general SL(2, $\mathbb{R}$)-matrix can be written, using the Iwasawa decomposition, as follows:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix} \begin{pmatrix} 1 & \vartheta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & \cos(\omega) \end{pmatrix}. \quad (233)$$

The leftmost block corresponds in E to an unphysical rescaling of the vectors (in GL(28, $\mathbb{R}$)). The middle block realizes, in going from the unhatted frame to the hatted one, a constant shift in the generalized θ-angle matrix $\mathcal{R}$: $\mathcal{R} \to \mathcal{R} + \vartheta\, \eta$. This can have effects at the quantum level, but does not affect field equations [62].

The rightmost block has, on the other hand, important bearing on the physics of the classical theory. Let $E(\omega)$ be the symplectic image (232) of this block only, and let $\mathscr{R}_v$ be the SL(8, $\mathbb{R}$)-frame, where the CSO(p, q, r) gaugings were originally constructed and in which the matrices $\mathbb{L}$ and $\mathcal{M}$ are given by well know general formulas [4, 53]. For $\omega \neq 0$, this frame is no longer electric, but is related to the electric one by $E(\omega)$. Using (167) we can write:

$$X_{\hat{\Lambda}} = \cos(\omega) X_\Lambda + \sin(\omega)\eta_{\Lambda\Sigma} X^\Sigma; \quad 0 = -\sin(\omega)\eta^{\Lambda\Sigma} X_\Sigma + \cos(\omega) X^\Lambda, \quad (234)$$

where $(\eta^{\Lambda\Sigma}) \equiv \eta^{-1} = \eta$. The above relation is easily inverted:

$$X_\Lambda = \cos(\omega)\, X_{\hat{\Lambda}}, \quad X^\Lambda = \sin(\omega)\, \eta^{\Lambda\Sigma} X_{\hat{\Sigma}}. \quad (235)$$

We can then write the symplectic invariant connection (162) in the following way:

$$\Omega_{g\,\mu} = A^M_\mu X_M = A^\Lambda_\mu X_\Lambda + A_{\Lambda\,\mu} X^\Lambda = (\cos\omega A^\Lambda_\mu + \sin(\omega) A_{\Lambda\,\mu}) X_{\hat{\Lambda}} = A^{\hat{\Lambda}}_\mu X_{\hat{\Lambda}}. \quad (236)$$

In other words, the gauging defined by X_M amounts to gauge, in the SL(8, $\mathbb{R}$)-frame, the same SO(p, q)-generators by a linear combination of the electric A^Λ_μ and magnetic $A_{\Lambda\,\mu}$ vector fields. The true electric vectors are all and only those entering the gauge connection, that is $A^{\hat{\Lambda}}_\mu$, and define the electric frame. We shall denote by $\Theta[\omega]$ the corresponding embedding tensor.

The gauged model can be constructed either directly in the SL(8, $\mathbb{R}$)-frame, using the covariant formulation to be discussed in Sect. 4, or in the electric frame, along the lines described in Sect. 3. The range of values of ω is restricted by the discrete symmetries of the theory. One of these is parity (see Sect. 2.4), whose duality representation $\mathbf{P}$ in the SL(8, $\mathbb{R}$)-frame has the form (82) [21]. The reader can verify that its effect on the $\mathbb{T}$-tensor (194) is:

$$\mathbb{T}(\Theta[\omega], \phi)_{\underline{M}} = \mathbf{P} \star \mathbb{T}(\Theta[-\omega], \phi_p) \quad (237)$$

by using the properties

$$\mathbf{P}_{\hat{M}}{}^{\hat{N}}\, \mathbf{P}^{-1} X_{\hat{N}} \mathbf{P} = X_{\hat{M}}; \quad \mathbf{P}^{-1} E(\omega) \mathbf{P} = E(-\omega); \quad \mathbf{P}^{-1} \mathbb{L}(\phi) \mathbf{P} = \mathbb{L}(\phi_p), \quad (238)$$

where ϕ_p denote the parity-transformed scalar fields. Equation (237) shows that parity maps ϕ into ϕ_p and ω in $-\omega$. In other words ω is *parity-odd parameter*. The overall $\mathbf{P}$ transformation on $\mathbb{T}$ in (237) is ineffective, since it will cancel everywhere in the Lagrangian, being $\mathbf{P}$ an $\mathrm{O}(2n_v)$-transformation. Similarly, we can use other discrete global symmetries of the ungauged theory, which include the SO(8)-triality transformations $S_3 \subset \mathrm{E}_{7(7)}$ for the SO(8)-gauging, to further restrict the range of values of ω. One finds that [61, 62]:

$$\begin{aligned} \omega &\in \left(0, \frac{\pi}{8}\right) , \quad \text{SO(8)-gauging,} \\ \omega &\in \left(0, \frac{\pi}{4}\right) , \quad \text{non-compact SO}(p,q)\text{-gaugings.} \end{aligned} \tag{239}$$

These are called "ω-rotated" $\mathrm{SO}(p,q)$-models, or simply $\mathrm{SO}(p,q)_\omega$-models. The SO(8) ones, in particular, came as a surprise since they contradicted the common belief that the original de Wit-Nicolai SO(8)-gauged model was unique.

For the non-semisimple $\mathrm{CSO}(p,q,r)$-gaugings, the non-trivial matrix E does not depend on continuous parameters but is fixed, thus yielding for each gauge group only one rotated-model [60, 62].

Even more surprisingly, these new class of gauged theories feature a broader range of vacua than the original models. In this sense the $\omega \to 0$ limit can be considered a singular one, in which some of the vacua move to the boundary of the moduli space at infinity and thus disappear.

Consider for instance the $\mathrm{SO}(8)_\omega$-models. They all feature an AdS_4, $N = 8$ vacuum at the origin with the same cosmological constant and mass spectrum as the original SO(8) theory. The parameter ω manifests itself in the higher order interactions of the effective theory. They also feature new vacua, which do not have counterparts in the $\omega = 0$ model. Figure 4 illustrates some of the vacua of the de Wit-Nicolai model ($\omega = 0$), namely those which feature a residual symmetry group $G_2 \subset \mathrm{SO}(8)$.

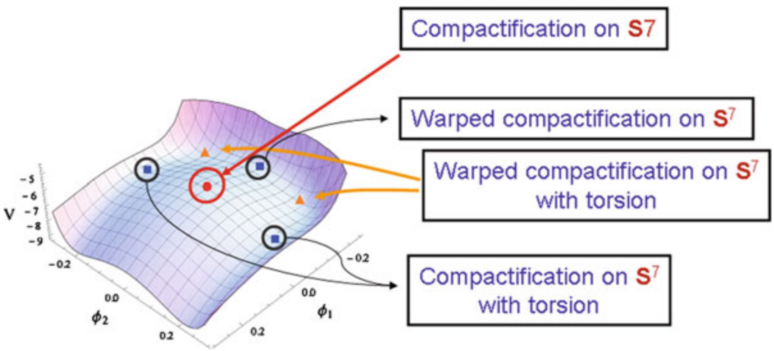


Fig. 4 The G_2-invariant vacua of the de Wit-Nicolai model, with their interpretation in terms of compactifications of the eleven-dimensional theory

Figure 5 shows the G_2-invariant vacua of a particular $\mathrm{SO}(8)_\omega$ model and the disappearance of one of the vacua in the $\omega \to 0$ limit [61]. The vacua of these models have been extensively studied [63–66] also in the context of renormalization group flows interpolating between (or simply originating from) *AdS* vacua [67, 68] and *AdS* black holes [69–71].

Determining a string or M-theory origin of the ω-rotated models is, to date, an open problem [72]. They seem to provide examples of what we named *intrinsically non-geometric* models in Sect. 3.3. The only exception so far is the dyonic ISO(7) which was related to compactifications of massive Type IIA theory [73].

4 Duality Covariant Gauging

Let us discuss in this section a formulation of the gauging procedure in four-dimensions which was developed in [8, 18] and which no longer depends on the matrix E, so that the kinetic terms are not written in terms of the vector fields in the electric frame.

Step 1, 2 and 3 revisited. We start from a symplectic-invariant gauge connection of the form[23]:

$$\Omega_{g\mu} \equiv A^M_\mu\, X_M = A^\Lambda_\mu\, X_\Lambda + A^\mu_\Lambda\, X_\Lambda = A^M_\mu\, \Theta_M{}^\alpha\, t_\alpha, \tag{240}$$

where $\Theta_M{}^\alpha$ satisfies the constraints (176), (177), (178). The fields A^Λ_μ and $A_{\Lambda\,\mu}$ are now taken to be independent. This is clearly a redundant choice and, as we shall see, half of them play to role of auxiliary fields. Equation (177) still implies that at most n_v linear combinations $A^{\hat{\Lambda}}_\mu$ of the $2n_v$ vectors A^Λ_μ, $A_{\Lambda\,\mu}$ effectively enter the gauge connection (and thus the minimal couplings):

$$A^M_\mu\, X_M = A^{\hat{\Lambda}}_\mu\, X_{\hat{\Lambda}}, \tag{241}$$

where $X_{\hat{\Lambda}}$ are defined in (167) through the matrix E, whose existence is guaranteed by (177), and where $A^{\hat{\Lambda}}_\mu \equiv E^{-1}{}_M{}^{\hat{\Lambda}}\, A^M_\mu$.

In the new formulation we wish to discuss, however, the vectors A^Λ_μ instead of $A^{\hat{\Lambda}}_\mu$ enter the kinetic terms. The covariant derivatives are then defined in terms of (240) as in Step 2 of the Sect. 3.1, and, as prescribed there, should replace ordinary derivative everywhere in the action. The infinitesimal gauge variation of A^M reads:

$$\delta A^M_\mu = \mathcal{D}_\mu \zeta^M \equiv \partial_\mu \zeta^M + A^N_\mu X_{NP}{}^M\, \zeta^P, \tag{242}$$

where, as usual, $X_{MP}{}^R \equiv \mathscr{R}_{v*}[X_M]_P{}^R$. We define for this set of electric-magnetic vector fields a symplectic covariant generalization $\mathbb{F}^M$ of the non-abelian field strengths $F^{\hat{\Lambda}}$ (137):

[23] Here, for the sake of simplicity, we reabsorb the gauge coupling constant g into Θ: $g\,\Theta \to \Theta$.

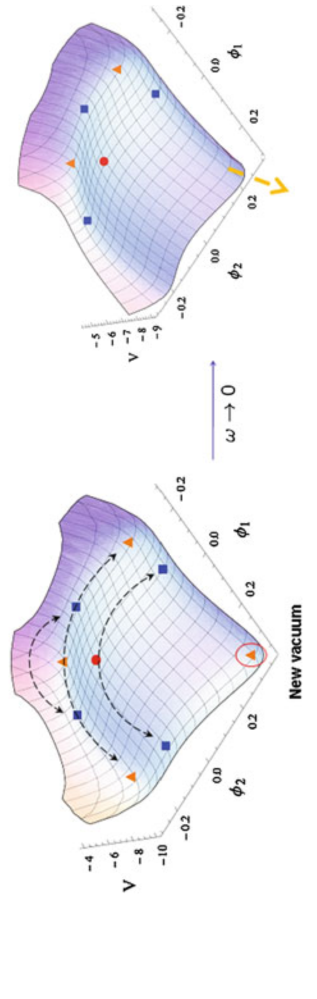


Fig. 5 On the *left* the G_2-invariant vacua of the $SO(8)_\omega$ model, with $\omega = \frac{\pi}{8}$. The *dashed lines* represent identifications of vacua due to a discrete symmetry of the theory which is a combination of triality and parity. All of them have an $\omega = 0$ counterpart, except the lowest one, marked by a *circle*, which disappears in the $\omega \to 0$ limit

$$F^M_{\mu\nu} \equiv \partial_\mu A^M_\nu - \partial_\nu A^M_\mu + X_{[NP]}{}^M A^N_\mu A^P_\nu \quad \Leftrightarrow \quad F^M \equiv dA^M + \frac{g}{2} X_{NP}{}^M A^N \wedge A^P, \tag{243}$$

where in the last equation we have used the form-notation for the fields strengths. The gauge algebra-valued curvature $\mathcal{F}$ is defined as in (136):

$$\mathcal{F} \equiv F^M X_M. \tag{244}$$

The first problem one encounters in describing the vectors A^Λ_μ in the kinetic terms is that, in a symplectic frame which is not the electric one, such fields are not well defined, since their curvatures fail to satisfy the Bianchi identity. This comes with no surprise, since the components $\Theta^{\Lambda\alpha}$ of the embedding tensor are nothing but *magnetic charges*. One can indeed verify that:

$$\mathcal{D}F^M \equiv dF^M + X_{NP}{}^M A^N \wedge F^P = X_{(PQ)}{}^M A^P \wedge \left(dA^Q + \frac{g}{3} X_{RS}{}^Q A^R \wedge A^S\right) \neq 0. \tag{245}$$

In particular $\mathcal{D}F^\Lambda \neq 0$ since $X_{(PQ)}{}^\Lambda = -\frac{1}{2}\Theta^{\Lambda\alpha} t_{\alpha M}{}^P \mathbb{C}_{PN} \neq 0$, being in the non-electric frame $\Theta^{\Lambda\alpha} \neq 0$. To deduce (245) we have used the quadratic constraint (178) on the gauge generators X_M in the $\mathscr{R}_{v*}$-representation, which reads:

$$X_{MP}{}^R X_{NR}{}^Q - X_{NP}{}^R X_{MR}{}^Q + X_{MN}{}^R X_{RP}{}^Q = 0. \tag{246}$$

From the above identity, after some algebra, one finds:

$$X_{[MP]}{}^R X_{[NR]}{}^Q + X_{[PN]}{}^R X_{[MR]}{}^Q + X_{[NM]}{}^R X_{[PR]}{}^Q = -(X_{NM}{}^R X_{(PR)}{}^Q)_{[MNP]}, \tag{247}$$

that is the *generalized structure constants* $X_{[MP]}{}^R$ entering the definition (243) do not satisfy the Jacobi identity, and this feature is at the root of (245). Related to this is the non-gauge covariance of F^M. The reader can indeed verify that (using the form-notation):

$$\delta F^M = -X_{NP}{}^M \zeta^N F^P + \left(2 X_{(NP)}{}^M \zeta^N F^P - X_{(NP)}{}^M A^N \wedge \delta A^P\right) \neq -X_{NP}{}^M \zeta^N F^P, \tag{248}$$

where δA^M is given by (242) and where we have used the general property

$$\delta F^M = \mathcal{D}\delta A^M - X_{(PQ)}{}^M A^P \wedge \delta A^Q, \tag{249}$$

valid for generic δA^M. We also observe that the obstruction to the Bianchi identity (245), as well as the non-gauge covariant terms in (248), are proportional to a same tensor $X_{(MN)}{}^P$. This quantity, as a consequence of (178) and (182), vanishes if contracted with the gauge generators X_M, namely with the first index of the embedding tensor: $X_{(MN)}{}^P \Theta_P{}^\alpha = 0$. Therefore the true electric vector fields $A^{\hat\Lambda}_\mu$ and the gauge connection which only depends on them, are perfectly well defined. Indeed, one can easily show using the matrix E that the gauge curvature (244) only contains the field strengths $F^{\hat{\Lambda}}$ associated with $A^{\hat{\Lambda}}$ and defined in (137):

$$\mathcal{F} \;\equiv\; F^M\,X_M \;=\; F^{\hat{\Lambda}}\,X_{\hat{\Lambda}}. \tag{250}$$

On the other hand, using (245) and (182) we have:

$$\mathcal{D}\mathcal{F} \;=\; \mathcal{D}F^M\,X_M \;=\; 0. \tag{251}$$

The gauge covariance (138) of $\mathcal{F}$, and thus of $F^{\hat{\Lambda}}$, is also easily verified by the same token, together with (142): $\mathcal{D}^2 = -\mathcal{F}$.

In order to construct gauge-covariant quantities describing the vector fields, we combine the vector field strengths $F^M_{\mu\nu}$ with a set of massless antisymmetric tensor fields[24] $B_{\alpha\,\mu\nu}$ in the adjoint representation of G through the matrix

$$Z^{M\,\alpha} \equiv \frac{1}{2}\,\mathbb{C}^{MN}\,\Theta_N{}^\alpha, \tag{252}$$

and define the following new field strengths:

$$\mathcal{H}^M_{\mu\nu} \;\equiv\; F^M_{\mu\nu} + Z^{M\,\alpha}\,B_{\alpha\,\mu\nu} \;:\; \begin{cases}\mathcal{H}^\Lambda = dA^\Lambda + \frac{1}{2}\,\Theta^{\Lambda\alpha}\,B_\alpha,\\ \mathcal{H}_\Lambda = dA_\Lambda - \frac{1}{2}\,\Theta_\Lambda{}^\alpha\,B_\alpha.\end{cases} \tag{253}$$

From the definition (252) and (177) we have:

$$Z^{M\,\alpha}\,\Theta_M{}^\beta = 0 \quad\Leftrightarrow\quad Z^{M\,\alpha}\,X_M = 0. \tag{254}$$

The reader can verify, using the linear constraint (176), that:

$$X_{(NP)}{}^M = -\frac{1}{2}\,\mathbb{C}^{MQ}\,X_{QN}{}^R\mathbb{C}_{RP} = -\frac{1}{2}\,\mathbb{C}^{MQ}\,\Theta_Q{}^\alpha\,t_{\alpha\,N}{}^R\mathbb{C}_{RP} = -Z^{M\,\alpha}\,t_{\alpha\,NP}, \tag{255}$$

where, as usual, we have defined $t_{\alpha\,NP} \equiv t_{\alpha\,N}{}^R\mathbb{C}_{RP}$.

The reason for considering the combination (253) is that the non-covariant terms in the gauge variation of $F^M_{\mu\nu}$, being proportional to $X_{(NP)}{}^M$, that is to $Z^{M\,\alpha}$, can be canceled by a corresponding variation of the tensor fields $\delta B_{\alpha\mu\nu}$:

$$\begin{aligned}\delta\mathcal{H}^M &= X_{PN}{}^M\,\zeta^N\,F^P + Z^{M\alpha}\,\left(\delta B_\alpha + t_{\alpha NP}\,A^N\wedge\delta A^P\right)\\ &= X_{PN}{}^M\,\zeta^N\,\mathcal{H}^P + Z^{M\alpha}\,\left(\delta B_\alpha + t_{\alpha NP}\,A^N\wedge\delta A^P\right)\\ &= -X_{NP}{}^M\,\zeta^N\,\mathcal{H}^P + 2\,X_{(NP)}{}^M\,\zeta^N\,\mathcal{H}^P + Z^{M\alpha}\,\left(\delta B_\alpha + t_{\alpha NP}\,A^N\wedge\delta A^P\right)\end{aligned}$$

[24] These fields will also be described as 2-forms $B_\alpha \equiv \frac{1}{2}\,B_{\mu\nu}\,dx^\mu\wedge dx^\nu$.

$$= -X_{NP}{}^{M}\,\zeta^{N}\,\mathcal{H}^{P} + Z^{M\alpha}\,\left[\delta B_{\alpha} + t_{\alpha NP}\,(A^{N}\wedge\delta A^{P} - 2\,\zeta^{N}\,\mathcal{H}^{P})\right], \qquad (256)$$

where, in going from the first to the second line, we have used (254), so that $X_{PN}{}^{M}\,F^{P} = X_{PN}{}^{M}\,\mathcal{H}^{P}$. If we define:

$$\delta B_{\alpha} \;\equiv\; t_{\alpha NP}\,(2\,\zeta^{N}\,\mathcal{H}^{P} - A^{N}\wedge\delta A^{P}), \qquad (257)$$

the term proportional to $Z^{M\alpha}$ vanishes and $\mathcal{H}^{M}$ transforms covariantly. The kinetic terms in the Lagrangian are then written in terms of $\mathcal{H}^{\Lambda}_{\mu\nu}$:

$$\frac{1}{e}\mathscr{L}_{v,\,\mathrm{kin}} = \frac{1}{4}\,\mathcal{I}_{\Lambda\Sigma}(\phi)\,\mathcal{H}^{\Lambda}_{\mu\nu}\,\mathcal{H}^{\Sigma\,\mu\nu} + \frac{1}{8\,e}\,\mathcal{R}_{\Lambda\Sigma}(\phi)\,\varepsilon^{\mu\nu\rho\sigma}\,\mathcal{H}^{\Lambda}_{\mu\nu}\,\mathcal{H}^{\Sigma}_{\rho\sigma}. \qquad (258)$$

The above transformation property (257) should however be modified since the quantity we want to transform covariantly is not quite $\mathcal{H}^{M}$, but rather the symplectic vector:

$$\mathcal{G}^{M} \equiv \begin{pmatrix}\mathcal{H}^{\Lambda}\\ \mathcal{G}_{\Lambda}\end{pmatrix}\;;\quad \mathcal{G}_{\Lambda\,\mu\nu} \equiv -\varepsilon_{\mu\nu\rho\sigma}\,\frac{\partial\mathscr{L}}{\partial\mathcal{H}^{\Lambda}_{\rho\sigma}}, \qquad (259)$$

corresponding, in the ungauged theory, to the field-strength-vector $\mathbb{F}^{M}$ of (51), and which contains inside $\mathcal{G}_{\Lambda}$ fermion bilinears coming from Pauli terms in the Lagrangian. Consistency of the construction will then imply that the two quantities $\mathcal{H}^{M}$ and $\mathcal{G}^{M}$, which are off-shell different since the former depends on the magnetic vector fields A_{Λ} as opposed to the latter, *will be identified on-shell* by the equation

$$(\mathcal{H}^{M} - \mathcal{G}^{M})\,\Theta_{M}{}^{\alpha} = (\mathcal{H}_{\Lambda} - \mathcal{G}_{\Lambda})\,\Theta^{\Lambda\alpha} = 0. \qquad (260)$$

These equations will in particular identify the field strengths of the auxiliary fields A_{Λ} in $\mathcal{H}_{\Lambda}$ with the duals to $\mathcal{H}^{\Lambda}$. The best that we can do is to make $\mathcal{G}^{M}$ on-shell covariant under G_{g}, namely upon use of (260). To this end, we modify (257) as follows:

$$\delta B_{\alpha} \equiv t_{\alpha NP}\,(2\,\zeta^{N}\,\mathcal{G}^{P} - A^{N}\wedge\delta A^{P}), \qquad (261)$$

so that the variations of the symplectic vectors $\mathcal{H}^{M}$ and $\mathcal{G}^{M}$ read:

$$\begin{aligned}\delta\mathcal{H}^{M} &= -X_{NP}{}^{M}\,\zeta^{N}\,\mathcal{H}^{P} \;+\; \text{non-covariant terms},\\ \delta\mathcal{G}^{M} &= -X_{NP}{}^{M}\,\zeta^{N}\,\mathcal{G}^{P} \;+\; \text{non-covariant but on-shell vanishing terms}.\end{aligned} \qquad (262)$$

Consistent definition of B_{α} requires the theory to be gauge-invariant with respect to transformations parametrized by 1-forms: $\Xi_{\alpha} = \Xi_{\alpha\mu}\,dx^{\mu}$. Such transformations should in turn be G_{g}-invariant and leave $\mathcal{H}^{M}$ unaltered:

$$A^{M} \to A^{M} + \delta_{\Xi}A^{M}\;;\quad B_{\alpha} \to B_{\alpha} + \delta_{\Xi}B_{\alpha} \quad\Rightarrow\quad \delta_{\Xi}\mathcal{H}^{M} = 0. \qquad (263)$$

Let us use (249) then to write

$$\delta_\Xi \mathcal{H}^M = \mathcal{D}\delta_\Xi A^M + Z^{M\,\alpha}\,\left(\delta_\Xi B_\alpha + t_{\alpha NP}\,A^N \wedge \delta_\Xi A^P\right). \tag{264}$$

If we set

$$\delta_\Xi A^M = -Z^{M\alpha}\;\Xi_\alpha, \tag{265}$$

the invariance of $\mathcal{H}^M$ implies:

$$\delta_\Xi B_\alpha = \mathcal{D}\Xi_\alpha - t_{\alpha NP}\,A^N \wedge \delta_\Xi A^P, \tag{266}$$

where

$$\mathcal{D}\Xi_\alpha \equiv d\,\Xi_\alpha + \Theta_M{}^\beta\,\mathrm{f}_{\beta\alpha}{}^\gamma A^M \wedge \Xi_\gamma. \tag{267}$$

Let us now introduce field strengths for the 2-forms:

$$\mathcal{H}^{(3)}_\alpha \equiv \mathcal{D}B_\alpha - t_{\alpha PQ}A^P \wedge \left(dA^Q + \frac{1}{3}\,X_{RS}{}^Q\,A^R \wedge A^S\right). \tag{268}$$

Writing the forms in components,

$$\mathcal{H}^{(3)}_\alpha = \frac{1}{3!}\,\mathcal{H}_{\alpha\,\mu\nu\rho}\,dx^\mu \wedge dx^\nu \wedge dx^\rho\ ;\quad \mathcal{D}B_\alpha = \frac{1}{2}\,\mathcal{D}_\mu B_{\alpha\,\nu\rho}\,dx^\mu \wedge dx^\nu \wedge dx^\rho, \tag{269}$$

we have:

$$\mathcal{H}_{\alpha\,\mu\nu\rho} = 3\,\mathcal{D}_{[\mu}B_{\alpha\,\nu\rho]} - 6\,t_{\alpha PQ}\left(A^P_{[\mu}\partial_\nu A^Q_{\rho]} + \frac{1}{3}\,X_{RS}{}^Q\,A^P_{[\mu}A^R_\nu A^S_{\rho]}\right). \tag{270}$$

The reader can verify that the following Bianchi identities hold:

$$\mathcal{D}\mathcal{H}^M = Z^{M\alpha}\,\mathcal{H}^{(3)}_\alpha, \tag{271}$$

$$\mathcal{D}\mathcal{H}^{(3)}_\alpha = X_{NP}{}^M\,\mathcal{H}^N \wedge \mathcal{H}^P. \tag{272}$$

Just as in Step 3 of Sect. 3.1, gauge invariance of the bosonic action requires the introduction of topological terms, so that the final gauged bosonic Lagrangian reads:

$$\begin{aligned}
\mathscr{L}_\mathrm{B} = &-\frac{e}{2}\,R + \frac{e}{2}\,\mathcal{G}_{st}(\phi)\,\mathcal{D}_\mu\phi^s\,\mathcal{D}^\mu\phi^t \\
&+ \frac{e}{4}\,\mathcal{I}_{\Lambda\Sigma}\,\mathcal{H}_{\mu\nu}{}^\Lambda\mathcal{H}^{\mu\nu\,\Sigma} + \frac{1}{8}\mathcal{R}_{\Lambda\Sigma}\,\varepsilon^{\mu\nu\rho\sigma}\mathcal{H}_{\mu\nu}{}^\Lambda\mathcal{H}_{\rho\sigma}{}^\Sigma \\
&- \frac{1}{8}\,\varepsilon^{\mu\nu\rho\sigma}\,\theta^{\Lambda\alpha}\,B_{\mu\nu\,\alpha}\left(2\,\partial_\rho A_{\sigma\,\Lambda} + X_{MN\,\Lambda}\,A_\rho{}^M A_\sigma{}^N - \frac{1}{4}\,\theta_\Lambda{}^\beta B_{\rho\sigma\,\beta}\right) \\
&- \frac{1}{3}\,\varepsilon^{\mu\nu\rho\sigma}X_{MN\,\Lambda}\,A_\mu{}^M A_\nu{}^N\left(\partial_\rho A_\sigma{}^\Lambda + \frac{1}{4}X_{PQ}{}^\Lambda A_\rho{}^P A_\sigma{}^Q\right)
\end{aligned}$$

$$-\frac{1}{6}\,\varepsilon^{\mu\nu\rho\sigma}X_{MN}{}^{\Lambda}\,A_{\mu}{}^{M}A_{\nu}{}^{N}\Big(\partial_{\rho}A_{\sigma\,\Lambda}+\frac{1}{4}\,X_{PQ\Lambda}A_{\rho}{}^{P}A_{\sigma}{}^{Q}\Big). \tag{273}$$

The Chern-Simons terms in the last two lines generalize those in (157). On top of them, gauge invariance of the action requires the introduction of new topological terms, depending on the B-fields, which appear in the third line of (273). Notice that if the magnetic charges $\Theta^{\Lambda\,\alpha}$ vanish (i.e. we are in the electric frame), B_{α} disappear from the action, since the second line of (273) vanish as well as the B-dependent Stueckelberg term in $\mathcal{H}^{\Lambda}$.

The constraints (176), (177) and (178) are needed for the consistent construction of the gauged bosonic action, which is uniquely determined. Just as discussed in Sect. 3.1, they are also enough to guarantee its consistent supersymmetric completion through Step 4, which equally applies to this more general construction.

Some comments are in order.

(i) The construction we are discussing in this Section requires the introduction of additional fields: n_v magnetic potentials $A_{\Lambda\mu}$ and a set of antisymmetric tensors $B_{\alpha\,\mu\nu}$. These new fields come together with extra gauge-invariances (242), (265), (266), which guarantee the correct counting of physical degrees of freedom. As we shall discuss below these fields can be disposed of using their equations of motion.

(ii) It is known that in D-dimensions there is a duality that relates p-forms to $(D-p-2)$-forms, the corresponding field strengths having complementary order and being related by a Hodge-like duality. In four dimensions vectors are dual to vectors, while scalars are dual to antisymmetric tensor fields. From this point of view, we can understand the 2-forms B_{α} as "dual" to the scalars in the same way as A_{Λ} are "dual" to A^{Λ}. This relation can be schematically illustrated as follows:

$$\partial_{[\mu}B_{\nu\rho]}\;\propto\;e\,\varepsilon_{\mu\nu\rho\sigma}\partial^{\sigma}\phi\;+\;\ldots. \tag{274}$$

More precisely, we can write the non-local relation between B_{α} and ϕ^{s} in a G-covariant fashion as a Hodge-like duality between $\mathcal{H}^{(3)}_{\alpha}$ and the Noether current $\mathbf{j}_{\alpha}$ of the sigma model describing the scalar fields, associated with the generator t_{α}:

$$\mathcal{H}_{\alpha\,\mu\nu\rho}\propto e\,\varepsilon_{\mu\nu\rho\sigma}\,\mathbf{j}^{\sigma}_{\alpha}\;;\quad\mathbf{j}^{\mu}_{\alpha}\equiv\frac{\delta\mathscr{L}_{\rm B}}{\delta\partial_{\mu}\phi^{s}}\,k^{s}_{\alpha}, \tag{275}$$

k^{s}_{α} being the Killing vector corresponding to t_{α}. This motivated the choice of the 2-forms in the adjoint representation of G. In the gauged theory we will find a G_g-invariant version of (275), see discussion below.

(iii) It can be shown that the presence of the extra fields B_{α} and A_{Λ} in the action is related to non-vanishing magnetic components $\Theta^{\Lambda\,\alpha}$ of the embedding tensor. In the electric frame in which $\Theta^{\Lambda\,\alpha}=0$, these fields disappear altogether from the Lagrangian and we are back to the gauged action described in Sect. 3.1.

(iv) The kinetic terms in the Lagrangian only describe fields in the ungauged theory, while the extra fields enter topological terms or Stueckelberg-like couplings and satisfy first order equations, see discussion below. This feature is common to the G-covariant construction of gauged supergravities in any dimensions [7, 74–76].

(v) The dyonic embedding tensor $\Theta_M{}^\alpha$ determines a splitting of the $2n_v$ vector fields A^M_μ into the truly electric ones $A^{\hat{\Lambda}}_\mu$, which are singled out by the combination $A^M_\mu \Theta_M{}^\alpha$ and thus define the gauge connection. The remaining ones $\tilde{A}^M_\mu$ correspond to non-vanishing components of $Z^{M\alpha}$, that is to the components along which the Jacobi identity is not satisfied, see (247). These latter vectors, of which there are at most n_v independent, can be then written as $\tilde{A}^M_\mu = Z^{M\alpha} A_{\alpha\,\mu}$ and are ill-defined, since the corresponding field strengths do not satisfy the Bianchi identity. An other problem with the vectors $\tilde{A}^M_\mu$ is that they are not part of the gauge connection, but in general are charged under the gauge group, that is are minimally coupled to $A^{\hat{\Lambda}}_\mu$. These fields cannot therefore be consistently described as vector fields. However, this poses no consistency problem for the theory, since $\tilde{A}^M_\mu$ can be gauged away by a transformation (265), (266) proportional to Ξ_α. In a vacuum, they provide the two degrees of freedom needed by some of the tensor fields B_α to become massive, according to the *anti-Higgs* mechanism [77, 78]. In the electric frame, these vectors become magnetic ($A_{\hat{\Lambda}\,\mu}$) and disappear from the action. This phenomenon also occurs in higher dimensions: the vectors $\tilde{A}^M_\mu$ which do not participate in the gauge connection but are charged with respect to the gauge group, are gauged away by a transformation associated with some of the antisymmetric tensor fields which, in a vacuum, become massive.

(vi) An important role in this construction was played by the linear constraint (176), in particular by the property (255) implied by it, which allowed to cancel the non-covariant terms in the gauge variation of F^Λ by a corresponding variation of the antisymmetric tensor fields. It turns out that a condition analogous to (255) represents the relevant linear constraint on the embedding tensor needed for the construction of gauged theories in higher dimensions [7, 74–76].

Let us now briefly discuss the bosonic field equations for the antisymmetric tensor fields and the vectors. The variation of the action with respect to $B_{\alpha\,\mu\nu}$ yields equations (260). By fixing the Ξ_α-gauge freedom, we can gauge away the ill-defined vectors $\tilde{A}^M_\mu = Z^{M\alpha} A_{\alpha\,\mu}$ and then solve (260) in B_α as a function of the remaining field strengths, which are a combination of the $F^{\hat{\Lambda}}$ only. Substituting this solution in the action, the latter will only describe the $A^{\hat{\Lambda}}_\mu$ vector fields and no longer contain magnetic ones or antisymmetric tensors. In other words by eliminating B_α through equations (260) we effectively perform the rotation to the electric frame and find the action discussed in Sect. 3.1.

By varying the action with respect to A^M_μ we find the following equations:

$$\mathcal{D}_{[\mu}\mathcal{G}^M_{\rho\sigma]} = -2\,e\,\mathbb{C}^{MN}\,\varepsilon_{\mu\nu\rho\sigma}\,\mathcal{D}^\sigma\phi^s\mathcal{G}_{sr}\,k^r_N = -2\,e\,\mathbb{C}^{MN}\,\varepsilon_{\mu\nu\rho\sigma}\,\mathbf{j}^\sigma_N, \tag{276}$$

which are the manifestly G-covariant form of the Maxwell equations. The right-hand-side is proportional to the electric current

$$\mathbf{j}^\sigma_N \;\equiv\; \mathcal{D}^\sigma\phi^s\mathcal{G}_{sr}\,k^r_N \;=\; \Theta_N{}^\alpha\,\mathcal{D}^\sigma\phi^s\mathcal{G}_{sr}\,k^r_\alpha \;=\; \Theta_N{}^\alpha\,\mathbf{j}^\sigma_\alpha. \tag{277}$$

If we contract both sides of (276) with $\Theta_M{}^\alpha$, we are singling out the Bianchi identity for the fields strengths $F^{\hat{\Lambda}}$ of the vectors which actually participate in the minimal couplings. By using the locality condition on Θ, we find:

$$\mathcal{D}_{[\mu}\mathcal{G}^M_{\rho\sigma]}\,\Theta_M{}^\alpha = -2\,e\,\mathbb{C}^{MN}\,\Theta_M{}^\alpha\,\Theta_N{}^\beta\varepsilon_{\mu\nu\rho\sigma}\,\mathcal{D}^\sigma\phi^s\mathcal{G}_{sr}\,k^r_\beta = 0, \tag{278}$$

which are nothing but the Bianchi identities for $F^{\hat{\Lambda}}$. This is consistent with our earlier discussion, see (251), in which we showed that the locality condition implies that the Bianchi identity for the gauge curvature have no magnetic source term, so that the gauge connection is well defined.[25]

Now we can use the Bianchi identity (271) to rewrite (278) as a dualization equation generalizing (275). To this end, we consider only the upper components of (278), corresponding to the field equations for $A_{\Lambda\,\mu}$:

$$Z^{\Lambda\alpha}\,\mathcal{H}_{\alpha\,\mu\nu\rho} = -12\,e\,Z^{\Lambda\alpha}\,\varepsilon_{\mu\nu\rho\sigma}\,\mathcal{D}^\sigma\phi^s\mathcal{G}_{sr}\,k^r_\alpha. \tag{279}$$

When the gauging involves translational isometries [8], $\phi^I \to \phi^I + c^I$, the above equations can be solved in the fields A_Λ contained in the covariant derivative. This is done by first using the ζ-gauge freedom associated with A_Λ to gauge away the scalar fields ϕ^I acted on by the translational isometries. Equations (279) are then solved in the fields A_Λ, which are expressed in terms of the remaining scalars, the vectors A^Λ and the field strengths of the antisymmetric tensors. Substituting this solution in the action, we obtain a theory in which no vectors A_Λ appear and the scalar fields ϕ^I have been effectively dualized to corresponding tensor fields $B_{I\,\mu\nu}$. The latter become dynamical and are described by kinetic terms. These theories were first constructed in the framework of $N = 2$ supergravity in [79, 80], generalizing previous results [81].

The gauged theory we have discussed in this section features a number of non-dynamical extra fields. This is the price we have to pay for a manifest G-covariance of the field equations and Bianchi identities. The embedding tensor then defines how the physical degrees of freedom are distributed within this larger set of fields, by fixing

[25] In our earlier discussion we showed that $\mathcal{D}\mathcal{H}^M\,\Theta_M{}^\alpha = \mathcal{D}F^M\,\Theta_M{}^\alpha = 0$. This is consistent with (278) since on-shell $\mathcal{H}^M\,\Theta_M{}^\alpha = \mathcal{G}^M\,\Theta_M{}^\alpha$.

the gauge symmetry associated with the extra fields and solving the corresponding non-dynamical field equations (260), (279).

A view on higher dimensions. As mentioned in point (ii) above, there are equivalent formulations of ungauged supergravities in D-dimensions obtained from one another by dualizing certain p-forms $C_{(p)}$ (i.e. rank-p antisymmetric tensor fields) into $(D-p-2)$-forms $C_{(D-p-2)}$ through an equation of the type:

$$dC_{(p)} = {}^*dC_{(D-p-2)} + \cdots . \tag{280}$$

Such formulations feature in general different global symmetry groups. This phenomenon is called *Dualization of Dualities* and was studied in [82]. The scalar fields in these theories are still described by a non-linear sigma model and in $D \geq 6$ the scalar manifold is homogeneous symmetric. Just as in four dimensions, the scalars are non-minimally coupled to the p-form fields (see below) and the global symmetry group G is related to the isometry group of the scalar manifold and thus is maximal in the formulation of the theory in which the scalar sector is maximal, that is in which all forms are dualized to lower order ones. This prescription, however, does not completely fix the ambiguity related to duality in even dimensions $D = 2k$, when order-k field strengths, corresponding to rank-$(k-1)$ antisymmetric tensor fields $C_{(k-1)}$, are present. In fact, after having dualized all forms to lower-order ones, we can still dualize $(k-1)$-forms $C_{(k-1)}$ into $(k-1)$-forms $\tilde{C}_{(k-1)}$. This is the electric-magnetic duality of the four-dimensional theory, related to the vector fields, and also occurs for instance in six dimensions with the 2-forms and in eight dimensions with the 3-forms.

Duality transformations interchanging $C_{(k-1)}$ with $\tilde{C}_{(k-1)}$, and thus the corresponding field equations with Bianchi identities, are encoded in the group G, whose action on the scalar fields, just as in four dimensions, is combined with a linear action on the k-form field strengths $F_{(k)}$ and their duals $\tilde{F}_{(k)}$:

$$g \in G \;:\quad \begin{cases} F_{(k)} \to F'_{(k)} = A[g]\, F_{(k)} + B[g]\, \tilde{F}_{(k)}, \\ \tilde{F}_{(k)} \to \tilde{F}'_{(k)} = C[g]\, F_{(k)} + D[g]\, \tilde{F}_{(k)}. \end{cases} \tag{281}$$

As long as the block $B[g]$ is non-vanishing, this symmetry can only be on-shell since the Bianchi identity for the transformed $F_{(k)}$, which guarantees that the transformed elementary field $C'_{(k-1)}$ be well defined, only holds if the field equations $d\tilde{F}_{(k)} = 0$ for $C_{(k-1)}$ are satisfied [83]:

$$dF'_{(k)} = A[g]\, dF_{(k)} + B[g]\, d\tilde{F}_{(k)} = B[g]\, d\tilde{F}_{(k)} = 0. \tag{282}$$

The field strengths $F_{(k)}$ and $\tilde{F}_{(k)}$ transform in a linear representation $\mathscr{R}$ of G defined by the matrix:

$$g \in G \;\; \xrightarrow{\mathscr{R}} \;\; \mathscr{R}[g] = \begin{pmatrix} A[g] & B[g] \\ C[g] & D[g] \end{pmatrix}. \tag{283}$$

Just as in four dimensions, depending on which of the $C_{(k-1)}$ and $\tilde{C}_{(k-1)}$ are chosen to be described as elementary fields in the Lagrangian, the action will feature a different global symmetry G_{el}, though the global symmetry group G of the field equations and Bianchi identities remains the same. The constraints on $\mathscr{R}$ derive from the non-minimal couplings of the scalar fields to the $(k-1)$-forms which are a direct generalization of those in four dimensions between the scalars and the vector fields,[26] see (258)

$$\mathscr{L}_{\text{kin},\,C} = -\frac{e\varepsilon}{2k!}\left(\mathcal{I}_{\Lambda\Sigma}(\phi)\,F^{\Lambda}_{\mu_1\ldots\mu_k}\,F^{\Lambda\,\mu_1\ldots\mu_k} + \mathcal{R}_{\Lambda\Sigma}(\phi)\,F^{\Lambda}_{\mu_1\ldots\mu_k}\,{}^*F^{\Lambda\,\mu_1\ldots\mu_k}\right), \quad (285)$$

where $\mu = 0, \ldots, D-1$ and $\Lambda, \Sigma = 1, \ldots, n_k$, being n_k the number of $(k-1)$-forms $C_{(k-1)}$ and $\varepsilon \equiv (-)^{k-1}$.

The matrices $\mathcal{I}_{\Lambda\Sigma}(\phi)$, $\mathcal{R}_{\Lambda\Sigma}(\phi)$ satisfy the following properties:

$$\mathcal{I}_{\Lambda\Sigma} = \mathcal{I}_{\Sigma\Lambda} < 0, \quad \mathcal{R}_{\Lambda\Sigma} = -\varepsilon\,\mathcal{R}_{\Sigma\Lambda}. \quad (286)$$

Just as we did in four dimensions, see (47), we define dual field strengths (omitting the fermion terms):

$$\mathcal{G}_{\Lambda\,\mu_1\ldots\,\mu_k} \equiv \varepsilon\,\varepsilon_{\mu_1\ldots\,\mu_k\nu_1\ldots\nu_k}\,\frac{\delta\mathscr{L}}{\delta F^{\Lambda}_{\nu_1\ldots\nu_k}} \quad \Rightarrow \quad \mathcal{G}_{\Lambda} = -\mathcal{I}_{\Lambda\Sigma}\,{}^*F^{\Sigma} - \varepsilon\,\mathcal{R}_{\Lambda\Sigma}\,F^{\Sigma}, \quad (287)$$

and define the vector of field strengths:

$$\mathbb{F} = (\mathbb{F}^M) \equiv \begin{pmatrix} F^{\Lambda} \\ \mathcal{G}_{\Lambda} \end{pmatrix}. \quad (288)$$

The definition (287) can be equivalently written in terms of the *twisted self-duality condition* [82]:

$${}^*\mathbb{F} = -\mathbb{C}_{\varepsilon}\,\mathcal{M}(\phi)\,\mathbb{F}, \quad (289)$$

which generalizes (53), where

$$\mathbb{C}_{\varepsilon} \equiv (\mathbb{C}^{MN}) \equiv \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ \varepsilon\,\mathbf{I} & \mathbf{0} \end{pmatrix}, \quad (290)$$

[26] The Hodge dual ${}^*\omega$ of a generic q-form ω is defined as:

$$*\,\omega_{\mu_1\ldots\mu_{D-q}} = \frac{e}{q!}\,\varepsilon_{\mu_1\ldots\mu_{D-q}\nu_1\ldots\nu_q}\,\omega^{\nu_1\ldots\nu_q}, \quad (284)$$

where $\varepsilon_{01\ldots D-1} = 1$. One can easily verify that ${}^{**}\omega = (-)^{q(D-q)}\,(-)^{D-1}\,\omega$.

I, **0** being the $n_k \times n_k$ identity and zero-matrices, respectively, and

$$\mathcal{M}(\phi) = (\mathcal{M}(\phi)_{MN}) \equiv \begin{pmatrix} (\mathcal{I} - \varepsilon\,\mathcal{R}\mathcal{I}^{-1}\mathcal{R})_{\Lambda\Sigma} & -(\mathcal{R}\mathcal{I}^{-1})_{\Lambda}{}^{\Gamma} \\ \varepsilon(\mathcal{I}^{-1}\mathcal{R})^{\Delta}{}_{\Sigma} & \mathcal{I}^{-1\,\Delta\Gamma} \end{pmatrix}. \tag{291}$$

The reader can easily verify that:

$$\mathcal{M}^T\,\mathbb{C}_\varepsilon \mathcal{M} = \mathbb{C}_\varepsilon. \tag{292}$$

For $\varepsilon = -1$, which is the case of the vector fields in four dimensions, $\mathbb{C}_\varepsilon$ is the symplectic invariant matrix and $\mathcal{M}$ is a symmetric, symplectic matrix. For $\varepsilon = +1$, which is the case of 2-forms in six dimensions, $\mathbb{C}_\varepsilon$ is the $\mathrm{O}(n_k, n_k)$-invariant matrix and $\mathcal{M}$ a symmetric element of $\mathrm{O}(n_k, n_k)$.

The Maxwell equations read:

$$d\mathbb{F} = 0. \tag{293}$$

In order for (283) to be a symmetry of (289) and (293) we must have:

$$\mathcal{M}(g \star \phi) = \mathscr{R}[g]^{-T}\mathcal{M}(\phi)\mathscr{R}[g]^{-1}, \tag{294}$$

and

$$\mathscr{R}[g]^T\mathbb{C}_\varepsilon\mathscr{R}[g] = \mathbb{C}_\varepsilon. \tag{295}$$

This means that in $D = 2k$ dimensions:

$$\begin{aligned} k \text{ even} \ : &\quad \mathscr{R}[G] \subset \mathrm{Sp}(2n_k, \mathbb{R}), \\ k \text{ odd} \ : &\quad \mathscr{R}[G] \subset \mathrm{O}(n_k, n_k). \end{aligned} \tag{296}$$

All other forms of rank $p \neq k-1$, which include the vector fields in $D > 4$, will transform in linear representations of G. The corresponding kinetic Lagrangian only feature the first term of (285), with no generalized theta-term ($\mathcal{R} = 0$).

If we compactify Type IIA/IIB or eleven-dimensional supergravity on a torus down to D-dimensions, we end up with an effective ungauged, maximal theory in D dimensions, featuring form-fields of various order. Upon dualizing all form-fields to lower order ones, we end up with a formulation of the theory in which G is maximal, and is described by the non-compact real form $\mathrm{E}_{11-D(11-D)}$ of the group E_{11-D}. Here we use the symbol $\mathrm{E}_{11-D(11-D)}$ as a short-hand notation for the following groups:

$$\begin{aligned} D = 9 \ : &\quad G = \mathrm{E}_{2(2)} \equiv \mathrm{GL}(2, \mathbb{R}), \\ D = 8 \ : &\quad G = \mathrm{E}_{3(3)} \equiv \mathrm{SL}(2, \mathbb{R}) \times \mathrm{SL}(3, \mathbb{R}), \\ D = 7 \ : &\quad G = \mathrm{E}_{4(4)} \equiv \mathrm{SL}(5, \mathbb{R}), \\ D = 6 \ : &\quad G = \mathrm{E}_{5(5)} \equiv \mathrm{SO}(5, 5), \\ D = 5 \ : &\quad G = \mathrm{E}_{6(6)}, \end{aligned} \tag{297}$$

$$D = 4 : \quad G = E_{7(7)},$$
$$D = 3 : \quad G = E_{8(8)}.$$

Only for $D \leq 5$, $E_{11-D(11-D)}$ is a proper exceptional group. The ungauged four-dimensional maximal supergravity was originally obtained from compactification of the eleven-dimensional one and dualization of all form-fields to lower order ones in [53], where the $E_{7(7)}$ on-shell symmetry was found.

In $D = 10$ Type IIA and IIB theories feature different global symmetry groups: $G_{IIA} = SO(1, 1)$ and $G_{IIB} = SL(2, \mathbb{R})$, respectively. The latter encodes the conjectured S-duality symmetry of Type IIB string theory. In this theory G_{IIB} does not act as a duality group since the 5-form field strength is self-dual and is a G_{IIB}-singlet.

A G-covariant gauging [7, 74–76] is effected starting from the formulation of the ungauged theory in which G is maximal and promoting a suitable global symmetry group of the Lagrangian $G_g \subset G$ to local symmetry. The choice of the gauge group is still completely encoded in a G-covariant embedding tensor Θ:

$$\Theta \in \mathscr{R}_{v*} \times \mathrm{adj}(G), \tag{298}$$

subject to a linear constraint, generalizing (255), which singles out in the above product a certain representation $\mathscr{R}_\Theta$ for the embedding tensor, and a quadratic one expressing the G_g-invariance of Θ. In Table 2 we give, in the various D-dimensional maximal supergravities, the representations $\mathscr{R}_\Theta$ of Θ.

Just as in the duality covariant construction of the four-dimensional gaugings discussed above, one introduces all form-fields which are dual to the fields of the ungauged theory. All the form-fields will transform in representations of G and dual forms of different order will belong to conjugate representations. In $D = 2k$, in the presence of rank-$(k-1)$ antisymmetric tensors, this amounts to introducing the fields $\tilde{C}_{(k-1)\,\Lambda}$ dual to the elementary ones $C^\Lambda_{(k-1)}$, just as we did for the vector fields in four dimensions. Together they transform in the representation $\mathscr{R}$ discussed above. By consistency, each form-field is associated with its own gauge invariance. Only the fields of the original ungauged theory are described by kinetic terms, the extra fields enter in topological terms and in Stueckelberg-like combinations within the covariant field strengths. The latter, for a generic p-form field, can be schematically represented in the form (we suppress all indices)

$$F_{(p+1)} = \mathcal{D}C_{(p)} + Y_p[\Theta] \cdot C_{(p+1)} + \cdots . \tag{299}$$

where $Y_p[\Theta]$ is a constant *intertwiner* tensor constructed out of Θ and of G-invariant tensors. The gauge variation of the p-form has the following schematic expression:

$$\delta C_{(p)} = Y_p[\Theta] \cdot \Xi_{(p)} + \mathcal{D}\Xi_{(p-1)} + \cdots \tag{300}$$

The embedding tensor defines, through the tensors $Y_p[\Theta]$, a splitting of the p-forms into physical fields and unphysical ones. The former will in general become massive

Table 2 Decomposition of the embedding tensor Θ for maximal supergravities in various space-time dimensions in terms of irreducible G representations [5, 7]

D	G	H	Θ
7	SL(5)	USp(4)	$\mathbf{10} \times \mathbf{24} =$ $\mathbf{10} + \underline{\mathbf{15}} + \underline{\overline{\mathbf{40}}} + \mathbf{175}$
6	SO(5, 5)	USp(4) × USp(4)	$\mathbf{16} \times \mathbf{45} =$ $\mathbf{16} + \underline{\mathbf{144}} + \mathbf{560}$
5	$E_{6(6)}$	USp(8)	$\mathbf{27} \times \mathbf{78} =$ $\mathbf{27} + \underline{\mathbf{351}} + \overline{\mathbf{1728}}$
4	$E_{7(7)}$	SU(8)	$\mathbf{56} \times \mathbf{133} =$ $\mathbf{56} + \underline{\mathbf{912}} + \mathbf{6480}$
3	$E_{8(8)}$	SO(16)	$\mathbf{248} \times \mathbf{248} =$ $\underline{\mathbf{1}} + \mathbf{248} + \underline{\mathbf{3875}} +$ $\mathbf{27000} + \mathbf{30380}$

Only the underlined representations are allowed by supersymmetry. The R-symmetry group H is the maximal compact subgroup of G

by "eating" corresponding unphysical $(p-1)$-forms, while the latter, whose field strengths fail to satisfy the Bianchi identity, are in turn gauged away and become degrees of freedom of massive $(p+1)$-forms. The constraints on the embedding tensor and group theory guarantee the consistency of the whole construction.

Just as in the four-dimensional model discussed above, the embedding tensor defines the distribution of the physical degrees of freedom among the various fields by fixing the gauge freedom (300) and solving the non-dynamical field equations. These G-covariant selective couplings between forms of different order, determined by a single object Θ, define the so-called *tensor hierarchy* and was developed in the maximal theories, in [7, 75, 76] as a general G-covariant formulation of the gauging procedure in any dimensions. In this formalism the maximal gauged supergravity in $D = 5$ was constructed in [74], generalizing previous works [84, 85]. The general gauging of the six and seven -dimensional maximal theories were constructed in [86] and [75] respectively, extending previous works [87]. In $D = 8$ the most general gaugings were constructed in [88]. We refer to these works for the details of the construction in the different cases.

References

1. O. Hohm, H. Samtleben, Exceptional field theory. II. $E_{7(7)}$. Phys. Rev. D **89**, 066017 (2014). arXiv:1312.4542 [hep-th]
2. C. Hull, B. Zwiebach, Double field theory. JHEP **0909**, 099 (2009). arXiv:0904.4664 [hep-th]
3. C.M. Hull, P.K. Townsend, Unity of superstring dualities. Nucl. Phys. B **438**, 109 (1995). arXiv:hep-th/9410167
4. B. de Wit, H. Nicolai, N = 8 Supergravity. Nucl. Phys. B **208**, 323 (1982)
5. B. de Wit, H. Samtleben, M. Trigiante, On Lagrangians and gaugings of maximal supergravities. Nucl. Phys. B **655**, 93 (2003). arXiv:hep-th/0212239

6. B. de Wit, H. Samtleben, M. Trigiante, Maximal supergravity from IIB flux compactifications. Phys. Lett. B **583**, 338 (2004). arXiv:hep-th/0311224
7. B. de Wit, H. Samtleben, Gauged maximal supergravities and hierarchies of non Abelian vector-tensor systems. Fortsch. Phys. **53**, 442 (2005). arXiv:hep-th/0501243
8. B. de Wit, H. Samtleben, M. Trigiante, Magnetic charges in local field theory. JHEP **0509**, 016 (2005). arXiv:hep-th/0507289
9. C.M. Hull, Noncompact gaugings of N = 8 supergravity. Phys. Lett. B **142**, 39 (1984)
10. C.M. Hull, More gaugings of N = 8 supergravity. Phys. Lett. B **148**, 297 (1984)
11. M. Trigiante, *Dual Gauged Supergravities* (17th SIGRAV Conference Turin, September 4-7, Italy, 2006). arXiv:hep-th/0701218
12. H. Samtleben, Lectures on gauged supergravity and flux compactifications. Class. Quant. Grav. **25**, 214002 (2008). arXiv:0808.4076 [hep-th]
13. M. Grana, Flux compactifications in string theory: a comprehensive review. Phys. Rept. **423**, 91 (2006). arXiv:hep-th/0509003
14. R. Blumenhagen, B. Kors, D. Lust, S. Stieberger, Four-dimensional string compactifications with D-Branes, orientifolds and fluxes. Phys. Rept. **445**, 1 (2007). arXiv:hep-th/0610327
15. M.R. Douglas, S. Kachru, Flux compactification. Rev. Mod. Phys. **79**, 733 (2007). arXiv:hep-th/0610102
16. F. Cordaro, P. Fre, L. Gualtieri, P. Termonia, M. Trigiante, N = 8 gaugings revisited: an exhaustive classification. Nucl. Phys. B **532**, 245 (1998). arXiv:hep-th/9804056
17. H. Nicolai, H. Samtleben, Maximal gauged supergravity in three-dimensions. Phys. Rev. Lett. **86**, 1686 (2001). arXiv:hep-th/0010076
18. B. de Wit, H. Samtleben, M. Trigiante, The maximal D = 4 supergravities. JHEP **0706**, 049 (2007). arXiv:0705.2101 [hep-th]
19. L. Andrianopoli, M. Bertolini, A. Ceresole, R. D'Auria, S. Ferrara, P. Fre, T. Magri, N = 2 supergravity and N = 2 super Yang-Mills theory on general scalar manifolds: symplectic covariance, gaugings and the momentum map. J. Geom. Phys. **23**, 111 (1997). arxiv:hep-th/9605032
20. M.K. Gaillard, B. Zumino, Duality rotations for interacting fields. Nucl. Phys. B **193**, 221 (1981)
21. S. Ferrara, A. Marrani, E. Orazi, M. Trigiante, Dualities near the horizon. JHEP **1311**, 056 (2013). arXiv:1305.2057 [hep-th]
22. P.A.M. Dirac, Quantized singularities in the electromagnetic field. Proc. Roy. Soc. Lond. A **133**, 60 (1931)
23. J.S. Schwinger, Magnetic charge and quantum field theory. Phys. Rev. **144**, 1087 (1966)
24. D. Zwanziger, Quantum field theory of particles with both electric and magnetic charges. Phys. Rev. **176**, 1489 (1968)
25. B. de Wit, P.G. Lauwers, A. Van Proeyen, Lagrangians of N = 2 supergravity—matter systems. Nucl. Phys. B **255**, 569 (1985)
26. P. Van Nieuwenhuizen, Supergravity. Phys. Rep. **68**, 189 (1981)
27. S. Ferrara L. Maiani, An introduction to supersymmetry breaking in extended supergravity, in *5th SILARG Symposium on Relativity, Supersymmetry and Cosmology* (Bariloche, Argentina, January 4–11, 1985)
28. S. Cecotti, L. Girardello, M. Porrati, Constraints on partial superhiggs. Nucl. Phys. B **268**, 295 (1986)
29. R. D'Auria, S. Ferrara, On fermion masses, gradient flows and potential in supersymmetric theories. JHEP **0105**, 034 (2001). arXiv:hep-th/0103153
30. J.M. Maldacena, The Large N limit of superconformal field theories and supergravity. Int. J. Theor. Phys. **38**, 1113 (1999) (Adv. Theor. Math. Phys. **2**, 231 (1998)) arXiv:hep-th/9711200
31. E. Cremmer, S. Ferrara, C. Kounnas, D.V. Nanopoulos, Naturally vanishing cosmological constant in N = 1 supergravity. Phys. Lett. B **133**, 61 (1983)
32. J.R. Ellis, C. Kounnas, D.V. Nanopoulos, No scale supersymmetric guts. Nucl. Phys. B **247**, 373 (1984)
33. R. Barbieri, E. Cremmer, S. Ferrara, Flat and positive potentials in $N = 1$ supergravity. Phys. Lett. B **163**, 143 (1985)

34. C. Angelantonj, S. Ferrara, M. Trigiante, New D = 4 gauged supergravities from N = 4 orientifolds with fluxes. JHEP **0310**, 015 (2003). arXiv:hep-th/0306185
35. R. D'Auria, S. Ferrara, F. Gargiulo, M. Trigiante, S. Vaula, N = 4 supergravity Lagrangian for type IIB on T**6 / Z(2) in presence of fluxes and D3-branes. JHEP **0306**, 045 (2003). hep-th/0303049
36. J. Scherk, J.H. Schwarz, Nucl. Phys. B **153**, 61 (1979)
37. N. Kaloper, R.C. Myers, The odd story of massive supergravity. JHEP **9905**, 010 (1999). hep-th/9901045
38. V. Mathai, J.M. Rosenberg, T duality for torus bundles with H fluxes via noncommutative topology. Commun. Math. Phys. **253**, 705 (2004). hep-th/0401168
39. C.M. Hull, A geometry for non-geometric string backgrounds. JHEP **0510**, 065 (2005). hep-th/0406102
40. J. Shelton, W. Taylor, B. Wecht, Nongeometric flux compactifications. JHEP **0510**, 085 (2005). hep-th/0508133
41. N. Hitchin, Lectures on Generalized Geometry. arXiv:1008.0973 [math.DG]
42. M. Gualtieri, Generalized Complex Geometry. arXiv:math/0401221 [math-dg]
43. C.M. Hull, R.A. Reid-Edwards, Gauge symmetry, T-duality and doubled geometry. JHEP **0808**, 043 (2008). arXiv:0711.4818 [hep-th]
44. A. Dabholkar, C. Hull, Generalised T-duality and non-geometric backgrounds. JHEP **0605**, 009 (2006). arXiv:hep-th/0512005
45. O. Hohm, C. Hull, B. Zwiebach, Generalized metric formulation of double field theory. JHEP **1008**, 008 (2010). arXiv:1006.4823 [hep-th]
46. C.M. Hull, Generalised geometry for M-theory. JHEP **0707**, 079 (2007). arXiv:hep-th/0701203
47. P.P. Pacheco, D. Waldram, M-theory, exceptional generalised geometry and superpotentials. JHEP **0809**, 123 (2008). arXiv:0804.1362 [hep-th]
48. O. Hohm, H. Samtleben, Exceptional form of D = 11 supergravity. Phys. Rev. Lett. **111**, 231601 (2013). arXiv:1308.1673 [hep-th]
49. O. Hohm, H. Samtleben, Consistent Kaluza-Klein truncations via exceptional field theory. JHEP **1501**, 131 (2015). arXiv:1410.8145 [hep-th]
50. G. Dibitetto, A. Guarino, D. Roest, Charting the landscape of N = 4 flux compactifications. JHEP **1103**, 137 (2011). arXiv:1102.0239 [hep-th]
51. G. Dall'Agata, G. Inverso, On the vacua of N = 8 gauged supergravity in 4 dimensions. Nucl. Phys. B **859**, 70 (2012). arXiv:1112.3345 [hep-th]
52. G. Inverso, De Sitter Vacua in N = 8 Supergravity, MSc Thesis, Università di Padova, Italy, 14 July (2010)
53. E. Cremmer, B. Julia, The N = 8 supergravity theory. 1. The Lagrangian. Phys. Lett. B **80**, 48 (1978)
54. E. Cremmer, B. Julia, The SO(8) supergravity. Nucl. Phys. B **159**, 141 (1979)
55. A. Le Diffon, H. Samtleben, Supergravities without an action: gauging the trombone. Nucl. Phys. B **811**, 1 (2009). arXiv:0809.5180 [hep-th]
56. A. Le Diffon, H. Samtleben, M. Trigiante, N = 8 supergravity with local scaling symmetry. JHEP **1104**, 079 (2011). arXiv:1103.2785 [hep-th]
57. B. de Wit, H. Nicolai, The consistency of the S**7 truncation in D = 11 supergravity. Nucl. Phys. B **281**, 211 (1987)
58. E. Cremmer, J. Scherk, J.H. Schwarz, Spontaneously broken N = 8 supergravity. Phys. Lett. B **84**, 83 (1979)
59. L. Andrianopoli, R. D'Auria, S. Ferrara, M.A. Lledo, Gauging of flat groups in four-dimensional supergravity. JHEP **0207**, 010 (2002). arXiv:hep-th/0203206
60. G. Dall'Agata, G. Inverso, de Sitter vacua in N = 8 supergravity and slow-roll conditions. Phys. Lett. B **718**, 1132 (2013). arXiv:1211.3414 [hep-th]
61. G. Dall'Agata, G. Inverso, M. Trigiante, Evidence for a family of SO(8) gauged supergravity theories. Phys. Rev. Lett. **109**, 201301 (2012). arXiv:1209.0760 [hep-th]
62. G. Dall'Agata, G. Inverso, A. Marrani, Symplectic deformations of gauged maximal supergravity. JHEP **1407**, 133 (2014). arXiv:1405.2437 [hep-th]

63. A. Borghese, A. Guarino, D. Roest, All G_2 invariant critical points of maximal supergravity. JHEP **1212**, 108 (2012). arXiv:1209.3003 [hep-th]
64. A. Borghese, G. Dibitetto, A. Guarino, D. Roest, O. Varela, The SU(3)-invariant sector of new maximal supergravity. JHEP **1303**, 082 (2013)
65. A. Borghese, A. Guarino, D. Roest, Triality, periodicity and stability of SO(8) gauged supergravity. JHEP **1305**, 107 (2013). arXiv:1302.6057 [hep-th]
66. A. Gallerati, H. Samtleben, M. Trigiante, The $\mathcal{N} > 2$ supersymmetric AdS vacua in maximal supergravity. JHEP **1412**, 174 (2014). arXiv:1410.0711 [hep-th]
67. J. Tarrío, O. Varela, Electric/magnetic duality and RG flows in AdS_4/CFT_3. JHEP **1401**, 071 (2014). arXiv:1311.2933 [hep-th]
68. Y. Pang, C.N. Pope, J. Rong, Holographic RG flow in a new $SO(3) \times SO(3)$ sector of ω-deformed $SO(8)$ gauged $\mathcal{N} = 8$ supergravity. JHEP **1508**, 122 (2015). arXiv:1506.04270 [hep-th]
69. A. Anabalon, D. Astefanesei, Black holes in ω-defomed gauged N = 8 supergravity. Phys. Lett. B **732**, 137 (2014). arXiv:1311.7459 [hep-th]
70. H. Lü, Y. Pang, C.N. Pope, An ω deformation of gauged STU supergravity. JHEP **1404**, 175 (2014). arXiv:1402.1994 [hep-th]
71. S.Q. Wu, S. Li, Thermodynamics of static Dyonic AdS black holes in the ω-Deformed Kaluza-Klein gauged supergravity theory. Phys. Lett. B **746**, 276 (2015). arXiv:1505.00117 [hep-th]
72. K. Lee, C. Strickland-Constable, D. Waldram, New Gaugings and Non-geometry. arXiv:1506.03457 [hep-th]
73. A. Guarino, D. L. Jafferis, O. Varela, String theory origin of dyonic N=8 Supergravity and Its Chern-Simons duals. Phys. Rev. Lett. **115**(9), 091601 (2015) arXiv:1504.08009 [hep-th]
74. B. de Wit, H. Samtleben, M. Trigiante, The maximal D = 5 supergravities. Nucl. Phys. B **716**, 215 (2005). arXiv:hep-th/0412173
75. H. Samtleben, M. Weidner, The maximal D = 7 supergravities. Nucl. Phys. B **725**, 383 (2005). arXiv:hep-th/0506237
76. B. de Wit, H. Nicolai, H. Samtleben, Gauged supergravities, tensor hierarchies, and M-theory. JHEP **0802**, 044 (2008). arXiv:0801.1294 [hep-th]
77. P. K. Townsend, in *Gauge Field Theories: Theoretical Studies and Computer Simulations* ed. by W. Garazynski (Harwood Academic Chur, New York 1981)
78. S. Cecotti, S. Ferrara, L. Girardello, Massive vector multiplets from superstrings. Nucl. Phys. B **294**, 537 (1987)
79. G. Dall'Agata, R. D'Auria, L. Sommovigo, S. Vaula, D = 4, N = 2 gauged supergravity in the presence of tensor multiplets. Nucl. Phys. B **682**, 243 (2004). arXiv:hep-th/0312210
80. R. D'Auria, L. Sommovigo, S. Vaula, N = 2 supergravity Lagrangian coupled to tensor multiplets with electric and magnetic fluxes. JHEP **0411**, 028 (2004). arXiv:hep-th/0409097
81. J. Louis, A. Micu, Type 2 theories compactified on Calabi-Yau three folds in the presence of background fluxes. Nucl. Phys. B **635**, 395 (2002). arXiv:hep-th/0202168
82. E. Cremmer, B. Julia, H. Lu, C.N. Pope, Dualization of dualities. 1. Nucl. Phys. B **523**, 73 (1998). arXiv:hep-th/9710119
83. Y. Tanii, Introduction to Supergravities in Diverse Dimensions. arXiv:hep-th/9802138
84. M. Gunaydin, L.J. Romans, N.P. Warner, Compact and noncompact gauged supergravity theories in five-dimensions. Nucl. Phys. B **272**, 598 (1986)
85. L. Andrianopoli, F. Cordaro, P. Fre', L. Gualtieri, Nonsemisimple gaugings of D = 5 N = 8 supergravity and FDA.s. Class. Quant. Grav. **18**, 395 (2001). arXiv:hep-th/0009048
86. E. Bergshoeff, H. Samtleben, E. Sezgin, The gaugings of maximal D = 6 supergravity. JHEP **0803**, 068 (2008). arXiv:0712.4277 [hep-th]
87. M. Pernici, K. Pilch, P. van Nieuwenhuizen, Gauged maximally extended supergravity in seven-dimensions. Phys. Lett. B **143**, 103 (1984)
88. E. Bergshoeff, U. Gran, R. Linares, M. Nielsen, T. Ortin, D. Roest, The Bianchi classification of maximal D = 8 gauged supergravities. Class. Quant. Grav. **20**, 3997 (2003). arXiv:hep-th/0306179

Supersymmetric Black Holes and Attractors in Gauged Supergravity

Dietmar Klemm

Abstract These are notes of lectures given by the author at the school 'Theoretical frontiers in black holes and cosmology', iiP Natal (Brazil), June 2015. They are divided into three parts. The first contains a brief introduction to matter-coupled $N = 2$ gauged supergravity in four dimensions and its ingredients. Part two deals with the attractor mechanism in gauged supergravity, while in the last part we show how to construct both supersymmetric and nonextremal black holes in these theories.

1 Introduction

Black holes in gauged supergravity theories provide an important testground to address fundamental questions of gravity, both at the classical and quantum level. Among these are for instance the problems of black hole microstates, the final state of black hole evolution, uniqueness- or no hair theorems, to mention only a few of them. In gauged supergravity, the solutions typically have AdS asymptotics, and one can then try to study these issues guided by the AdS/CFT correspondence. On the other hand, black hole solutions to these theories are also relevant for a number of recent developments in high energy- and especially in condensed matter physics, since they provide the dual description of certain condensed matter systems at finite temperature, cf. [1] for a review. In particular, models that contain Einstein gravity coupled to U(1) gauge fields[1] and neutral scalars have been instrumental to study transitions from Fermi-liquid to non-Fermi-liquid behaviour, cf. [2, 3] and references therein. In AdS/condensed matter applications one is often interested in including a charged scalar operator in the dynamics, e.g. in the holographic modeling of strongly coupled superconductors [4]. This is dual to a charged scalar field in the bulk, that

[1] The necessity of a bulk U(1) gauge field arises, because a basic ingredient of realistic condensed matter systems is the presence of a finite density of charge carriers.

D. Klemm (✉)
Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano,
Via Celoria 16, 20133 Milano, Italy
e-mail: dietmar.klemm@mi.infn.it

R. Kallosh and E. Orazi (eds.), *Theoretical Frontiers in Black Holes and Cosmology*, Springer Proceedings in Physics 176,
DOI 10.1007/978-3-319-31352-8_3

typically appears in supergravity coupled to gauged hypermultiplets, which is the theory that will be considered in these lectures.

Another point of interest addressed here is the attractor mechanism [5–9], that has been the subject of extensive research in the asymptotically flat case, but for which not very much has been done for black holes with more general asymptotics. First steps towards a systematic analysis of the attractor flow in gauged supergravity were made in [10, 12] for the non-BPS and in [11, 13–15] for the BPS case. Some interesting results have been found, for instance the appearance of flat directions in the effective black hole potential for BPS flows [13], a property that does not occur in ungauged $N = 2, d = 4$ supergravity [9], at least as long as the metric of the scalar manifold is strictly positive definite.

In the second part of our lectures we extend the work of [12] to include also gauged hypermultiplets. We shall construct an effective potential V_{eff} that depends on both the usual black hole potential and the potential for the scalar fields. V_{eff} governs the attractors, in the sense that it is extremized on the horizon by all the scalar fields of the theory, and the entropy is given by the critical value of V_{eff}. As in [12], our analysis does not make use of supersymmetry, so our results are valid for any static extremal black hole in four-dimensional $N = 2$ matter-coupled supergravity with gauging of abelian isometries of the hypermultiplet scalar manifold.

These lectures are organized as follows: In the next section, we review $N = 2$, $d = 4$ gauged supergravity coupled to vector- and hypermultiplets. Section 3 contains an extension of the results of [12] on black hole attractors in gauged supergravity to the case that includes also hypermultiplets. Finally, in Sect. 4 we show how to construct both supersymmetric and nonextremal black holes in Fayet–Iliopoulos gauged $N = 2, d = 4$ supergravity.

2 Brief Introduction to $N = 2, d = 4$ Gauged Supergravity and Ingredients

In this section we shall give a brief introduction into matter-coupled $N = 2$ gauged supergravity in four dimensions. For a much more extended discussion we refer to the original paper [16], to the book [17], or to the lecture notes [18].

The gravity multiplet of $N = 2, d = 4$ supergravity can be coupled to a number n_V of vector multiplets and to n_H hypermultiplets. The bosonic sector then includes the vierbein $e^a{}_\mu$, $\bar{n} \equiv n_V + 1$ vector fields $A^\Lambda{}_\mu$ with $\Lambda = 0, \dots, n_V$ (the graviphoton plus n_V other fields from the vector multiplets), n_V complex scalar fields Z^i, $i = 1, \dots, n_V$, and $4n_H$ real hyperscalars q^u, $u = 1, \dots, 4n_H$.

The complex scalars Z^i of the vector multiplets parametrize an n_V-dimensional special Kähler manifold, i.e. a Kähler–Hodge manifold (which means that the Kähler form is of even integer cohomology), with Kähler metric $\mathcal{G}_{i\bar{j}}(Z, \bar{Z})$, which is the base of a symplectic bundle with the covariantly holomorphic sections

$$\mathcal{V} = \begin{pmatrix} \mathcal{L}^\Lambda \\ \mathcal{M}_\Lambda \end{pmatrix}, \qquad \mathcal{D}_{\bar{i}}\mathcal{V} \equiv \partial_{\bar{i}}\mathcal{V} - \frac{1}{2}\left(\partial_{\bar{j}}\mathcal{K}\right)\mathcal{V} = 0, \tag{1}$$

obeying the constraint

$$\langle \mathcal{V}|\bar{\mathcal{V}}\rangle \equiv \bar{\mathcal{L}}^\Lambda \mathcal{M}_\Lambda - \mathcal{L}^\Lambda \bar{\mathcal{M}}_\Lambda = -i, \tag{2}$$

where $\mathcal{K}$ is the Kähler potential and $\mathcal{D}$ denotes the Kähler-covariant derivative. Alternatively one can introduce the explicitly holomorphic sections of a different symplectic bundle,

$$\Omega \equiv e^{-\mathcal{K}/2}\mathcal{V} \equiv \begin{pmatrix} \chi^\Lambda \\ \mathcal{F}_\Lambda \end{pmatrix}. \tag{3}$$

In appropriate symplectic frames it is possible to choose a homogeneous function of second degree $\mathcal{F}(\chi)$, called prepotential, such that $\mathcal{F}_\Lambda = \partial_\Lambda \mathcal{F}$. In terms of the sections Ω the constraint (2) becomes

$$\langle \Omega|\bar{\Omega}\rangle \equiv \bar{\chi}^\Lambda \mathcal{F}_\Lambda - \chi^\Lambda \bar{\mathcal{F}}_\Lambda = -ie^{-\mathcal{K}}. \tag{4}$$

The couplings of the vector fields to the scalars are determined by the $\bar{n} \times \bar{n}$ period matrix $\mathcal{N}$, defined by the relations

$$\mathcal{M}_\Lambda = \mathcal{N}_{\Lambda\Sigma}\,\mathcal{L}^\Sigma, \qquad \mathcal{D}_{\bar{i}}\bar{\mathcal{M}}_\Lambda = \mathcal{N}_{\Lambda\Sigma}\,\mathcal{D}_{\bar{i}}\bar{\mathcal{L}}^\Sigma. \tag{5}$$

If the theory is defined in a frame in which a prepotential exists, $\mathcal{N}$ can be obtained from

$$\mathcal{N}_{\Lambda\Sigma} = \bar{\mathcal{F}}_{\Lambda\Sigma} + 2i\frac{\left(N_{\Lambda\Gamma}\chi^\Gamma\right)\left(N_{\Sigma\Delta}\chi^\Delta\right)}{\chi^\Omega N_{\Omega\Psi}\chi^\Psi}, \tag{6}$$

where $\mathcal{F}_{\Lambda\Sigma} = \partial_\Lambda\partial_\Sigma\mathcal{F}$ and $N_{\Lambda\Sigma} \equiv \mathfrak{Im}(\mathcal{F}_{\Lambda\Sigma})$.

As an exercise, consider the prepotential $\mathcal{F} = -i\chi^0\chi^1$. This is a model with just one vector multiplet ($n_V = 1$), and thus there is only one coordinate $Z = Z^1$. We have then $\mathcal{F}_0 = -i\chi^1, \mathcal{F}_1 = -i\chi^0$. If we choose the parametrization $\chi^0 = 1, \chi^1 = Z$ (called 'special coordinates'), the holomorphic symplectic section (3) becomes

$$\Omega = \begin{pmatrix} \chi^0 \\ \chi^1 \\ \mathcal{F}_0 \\ \mathcal{F}_1 \end{pmatrix} = \begin{pmatrix} 1 \\ Z \\ -iZ \\ -i \end{pmatrix}, \tag{7}$$

and the constraint (4) gives

$$-ie^{-\mathcal{K}} = -2i(Z + \bar{Z}). \tag{8}$$

The special Kähler metric is thus

$$\mathcal{G}_{Z\bar{Z}} = \partial_Z \partial_{\bar{Z}} \mathcal{K} = \frac{1}{(Z+\bar{Z})^2}. \tag{9}$$

This is the $SU(1,1)/U(1)$ model. Using (5), determine the period matrix $\mathcal{N}$!

Further exercise: Consider the more general prepotential $\mathcal{F} = -i(\chi^0)^n(\chi^1)^{2-n}$. Determine the symplectic section, the Kähler potential, Kähler metric and period matrix. Why is n restricted to the range $0 < n < 2$? (In order to answer this last question, you need to take a look at the kinetic terms in the action (14)).

We come now to the hypermultiplet sector. The $4n_H$ real hyperscalars q^u parametrize a quaternionic Kähler manifold with metric $H_{uv}(q)$. A quaternionic Kähler manifold is a $4n$-dimensional Riemannian manifold admitting a locally defined triplet $\mathbf{K}_u^v$ of almost complex structures satisfying the quaternion relations

$$K^x K^y = \varepsilon_{xyz} K^z - \delta_{xy} \quad , \qquad (x, y, z = 1, 2, 3), \tag{10}$$

and whose Levi–Civita connection preserves $\mathbf{K}$ up to a rotation,

$$\nabla_w \mathbf{K}_u^v + \mathbf{A}_w \times \mathbf{K}_u^v = 0, \tag{11}$$

with $SU(2)$ connection $\mathbf{A} \equiv \mathbf{A}_u(q)\, dq^u$. (This distinguishes quaternionic Kähler manifolds from hyper-Kähler manifolds). A further property is that the $SU(2)$ curvature is proportional to the complex structures,

$$F^x \equiv dA^x + \frac{1}{2}\varepsilon^{xyz} A^y \wedge A^z = -2\, K^x. \tag{12}$$

Quaternionic Kähler manifolds are Einstein manifolds with holonomy group $USp(2n_H) \times SU(2)/\mathbb{Z}_2$. The $SU(2)$ factor mixes the three complex structures. Since $USp(2n_H) = Sp(n_H)$ and $SU(2) = Sp(1)$, this is sometimes written in the form $Sp(n_H) \times Sp(1)/\mathbb{Z}_2$. Notice in this context that the compact symplectic group $USp(2n)$ is the subgroup of $GL(n, \mathbb{H})$ that preserves the standard hermitian form on $\mathbb{H}^n$,

$$\langle x, y \rangle = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n.$$

$USp(2n)$ is thus just the quaternionic unitary group $U(n, \mathbb{H})$. Notice that, in general, quaternionic Kähler manifolds are not Kähler.

In what follows, we will only consider gaugings of abelian isometries of the quaternionic-Kähler metric H_{uv}. (No gauging of isometries of the vector multiplets' special Kähler manifold). These are generated by commuting Killing vectors $k_\Lambda{}^u(q)$, $[k_\Lambda, k_\Sigma] = 0$. The requirement that the quaternionic-Kähler structure is preserved implies the existence of a triplet of Killing prepotentials, or moment maps, $P_\Lambda{}^x$ for each Killing vector such that

$$P_{\Lambda}{}^{x} = \frac{1}{2n_H} K^{x}{}_{u}{}^{v} \nabla_{v} k_{\Lambda}{}^{u}, \qquad D_u P_{\Lambda}{}^{x} \equiv \partial_u P_{\Lambda}{}^{x} + \varepsilon^{xyz} A^{y}{}_{u} P_{\Lambda}{}^{z} = -2\, K^{x}{}_{uv} k_{\Lambda}{}^{v}. \quad (13)$$

Note that the moment maps are related to the generating functions of canonical transformations in classical mechanics: In that case, one requires that an infinitesimal symmetry preserves the symplectic form. This gives the canonical transformations, and implies that they are given in terms of a generating function.

The bosonic action reads

$$S = \int d^4x \sqrt{|g|} \Big[R + 2\, \mathcal{G}_{i\bar{j}}\, \partial_\mu Z^i \partial^\mu \bar{Z}^{\bar{j}} + 2\, H_{uv}\, \mathfrak{D}_\mu q^u \mathfrak{D}^\mu q^v$$
$$+ 2\, I_{\Lambda\Sigma}\, F^{\Lambda\,\mu\nu} F^{\Sigma}{}_{\mu\nu} - 2\, R_{\Lambda\Sigma}\, F^{\Lambda\,\mu\nu} \star F^{\Sigma}{}_{\mu\nu} - V(Z, \bar{Z}, q) \Big], \quad (14)$$

where the scalar potential has the form

$$V(Z, \bar{Z}, q) = g^2 \left[2 \bar{\mathcal{L}}^{\Lambda} \mathcal{L}^{\Sigma} (H_{uv} k_{\Lambda}{}^{u} k_{\Sigma}{}^{v} - P_{\Lambda}{}^{x} P_{\Sigma}{}^{x}) - \frac{1}{4} I^{\Lambda\Sigma} P_{\Lambda}{}^{x} P_{\Sigma}{}^{x} \right], \quad (15)$$

and g is the gauge coupling constant. The covariant derivatives acting on the hyper-scalars are

$$\mathfrak{D}_\mu q^u = \partial_\mu q^u + g A^{\Lambda}{}_{\mu} k_{\Lambda}{}^{u}, \quad (16)$$

and

$$I_{\Lambda\Sigma} \equiv \mathfrak{Im}(\mathcal{N}_{\Lambda\Sigma}), \qquad R_{\Lambda\Sigma} \equiv \mathfrak{Re}(\mathcal{N}_{\Lambda\Sigma}), \qquad I^{\Lambda\Sigma} I_{\Sigma\Gamma} = \delta^{\Lambda}{}_{\Gamma}. \quad (17)$$

Before we continue, some comments are in order:

- **Symplectic transformations**: Consider the case without gauging, $g = 0$, and define

$$F^{\pm\Lambda}_{\mu\nu} := \frac{1}{2} \left(F^{\Lambda}_{\mu\nu} \mp \frac{i}{2} \varepsilon_{\mu\nu\rho\sigma} F^{\Lambda\rho\sigma} \right), \qquad G^{\mu\nu}_{\pm\Lambda} := \mathcal{N}_{\Lambda\Sigma} F^{\pm\Sigma\mu\nu}. \quad (18)$$

  Then the Bianchi identities and Maxwell equations can be written in the form

$$\nabla_\mu \mathfrak{Im} F^{+\Lambda\mu\nu} = 0 \quad \text{(Bianchi)}, \qquad \nabla_\mu \mathfrak{Im} G^{\mu\nu}_{+\Lambda} = 0 \quad \text{(Maxwell)}. \quad (19)$$

  Exercise: Show this!
  The (19) are invariant under $\mathrm{GL}(2\bar{n}, \mathbb{R})$,

$$\begin{pmatrix} \tilde{F}^+ \\ \tilde{G}_+ \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} F^+ \\ G_+ \end{pmatrix}. \quad (20)$$

  Since (18) must be preserved, we need

$$\tilde{G}_+ = CF^+ + DG_+ = (C + D\mathcal{N})F^+ = (C + D\mathcal{N})(A + B\mathcal{N})^{-1}\tilde{F}^+ \stackrel{!}{=} \tilde{\mathcal{N}}\tilde{F}^+, \tag{21}$$

and thus $\mathcal{N}$ transforms as

$$\tilde{\mathcal{N}} = (C + D\mathcal{N})(A + B\mathcal{N})^{-1}. \tag{22}$$

It is an easy exercise to shew that the symmetry of $\tilde{\mathcal{N}}$ implies the relations

$$A^T C - C^T A = 0, \qquad B^T D - D^T B = 0, \qquad A^T D - C^T B = \quad , \tag{23}$$

and hence

$$S \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}(2\bar{n}, \mathbb{R}), \tag{24}$$

since (23) is equivalent to $S^T \Omega S = \Omega$, where

$$\Omega = \begin{pmatrix} \mathbb{0} & \\ - & \mathbb{0} \end{pmatrix}. \tag{25}$$

Note that in the case of gauging, the potentials A^Λ_μ appear explicitely, for instance in the covariant derivative (16). Symplectic covariance is thus broken, unless one considers an extended formalism that includes in addition also magnetic gaugings, like e.g. in [14]. Moreover, from (22) we see that the scalars in the vector multiplets transform under symplectic transformations ($\mathcal{N}$ depends on the Z^i), and thus the presence of a scalar potential typically breaks symplectic covariance. In some cases, a discrete subgroup of $\mathrm{Sp}(2\bar{n}, \mathbb{R})$ may survive after the gauging.

- **Case without hypermultiplets**: The algebra of symmetries implies the 'equivariance condition'

$$\mathbf{P}_\Lambda \times \mathbf{P}_\Sigma + \frac{1}{2}\mathbf{K}_{uv} k_\Lambda{}^u k_\Sigma{}^v = 0. \tag{26}$$

If the quaternionic Kähler manifold is nontrivial ($n_H \geq 1$), the unique solution of (26) is given by

$$\mathbf{P}_\Lambda = \frac{1}{2n_H}\mathbf{K}_u{}^v \nabla_v k_\Lambda{}^u, \tag{27}$$

cf. (13). However, for $n_H = 0$, there is still the solution

$$\mathbf{P}_\Lambda = \mathbf{e}\xi_\Lambda, \tag{28}$$

where $\mathbf{e}$ denotes an arbitrary constant vector in SU(2) space, and the ξ_Λ are constants, called Fayet-Iliopoulos (FI) parameters. The moment maps (28) are called U(1) FI terms. They will be relevant for the last part of these lectures, where we shall consider only the case without hypers.

3 Attractor Mechanism in Gauged Supergravity

In ungauged supergravity, the attractor mechanism [5–9] essentially states that, at the horizon of an extremal black hole, the scalar fields ϕ of the theory are always attracted to the same values ϕ_{hor} (fixed by the black hole charges), independently of their values ϕ_∞ at infinity. When the so-called black hole potential (to be introduced below) has flat directions, it may happen that some moduli are not stabilized, i.e., their values at the horizon are not fixed in terms of the black hole charges. Yet, the Bekenstein–Hawking entropy turns out to be independent of these unstabilized moduli. Notice that this does not hold anymore for nonextremal black holes, for which the horizon is not necessarily an attractor point.

The aim of this part of our lectures is to show how the attractor mechanism is generalized in gauged supergravity. In this case, the moduli fields have a potential, and typically approach the critical points of this potential asymptotically, where the solution approaches anti-de Sitter space. Thus, unless there are flat directions in the scalar potential, the values of the moduli at infinity are completely fixed (for instance in terms of FI parameters or other constants appearing in the potential), and therefore a more suitable formulation of the attractor mechanism in the gauged case would be to say that the black hole entropy is determined entirely by the charges, and is independent of the values of the moduli on the horizon that are not fixed by the charges. First steps towards a systematic analysis of the attractor flow in gauged supergravity were made in [10–15, 19, 20]. In particular, in [12] the authors presented a generalization of the attractor mechanism to extremal static black holes in $N = 2$, $d = 4$ gauged supergravity coupled to abelian vector multiplets. In this section, which is mainly based on the results of [21], we closely follow their argument, generalizing it by taking into account the presence of gauged hypermultiplets. As in [12], we make no assumption on the form of the scalar potential, of the vectors' kinetic matrix $\mathcal{N}$ or on the scalar manifolds, so that our results are valid not only for $N = 2$ supergravity, but for any theory described by an action of the form (14).

The equations of motion obtained from the variation of (14) are

$$R_{\mu\nu} + T_{\mu\nu} + 2\mathcal{G}_{i\bar{j}}\, \partial_{(\mu} Z^i \partial_{\nu)} \bar{Z}^{\bar{j}} + 2H_{uv}\mathfrak{D}_\mu q^u \mathfrak{D}_\nu q^v - \frac{1}{2} g_{\mu\nu} V = 0, \tag{29}$$

$$\nabla_\nu \left(\star F_\Lambda{}^{\nu\mu}\right) + \frac{g}{2} k_{\Lambda u} \mathfrak{D}^\mu q^u = 0, \tag{30}$$

$$\mathfrak{D}^2 Z^i + \partial^i F_\Lambda{}^{\mu\nu} \star F^\Lambda{}_{\mu\nu} + \frac{1}{2}\partial^i V = 0, \tag{31}$$

$$\mathfrak{D}^2 q^u + \frac{1}{4}\partial^u V = 0, \tag{32}$$

where

$$T_{\mu\nu} \equiv I_{\Lambda\Sigma}\left(4F^\Lambda{}_\mu{}^\rho F^\Sigma{}_{\nu\rho} - g_{\mu\nu} F^\Lambda{}_{\rho\sigma} F^{\Sigma\rho\sigma}\right), \tag{33}$$

the dual field strengths are given by

$$F_{\Lambda\mu\nu} \equiv -\frac{1}{4\sqrt{|g|}}\frac{\delta S}{\delta \star F^{\Lambda\mu\nu}} = R_{\Lambda\Sigma}F^{\Sigma}{}_{\mu\nu} + I_{\Lambda\Sigma}\star F^{\Sigma}{}_{\mu\nu}, \tag{34}$$

and the second covariant derivatives on the scalars act as

$$\mathfrak{D}^2 Z^i = \nabla_\mu \partial^\mu Z^i + \Gamma^i_{jk}\partial_\mu Z^j \partial^\mu Z^k, \tag{35}$$

$$\mathfrak{D}^2 q^u = \nabla_\mu \mathfrak{D}^\mu q^u + \Gamma^u_{vw}\mathfrak{D}_\mu q^v \mathfrak{D}^\mu q^w + gA^\Lambda{}_\mu \partial_v k_\Lambda{}^u \mathfrak{D}^\mu q^v. \tag{36}$$

The metric for the most general static extremal black hole background with flat, spherical or hyperbolic horizon can be written in the form

$$ds^2 = e^{2U(r)}dt^2 - e^{-2U(r)}\left[dr^2 + e^{2W(r)}\left(d\vartheta^2 + f_\kappa(\vartheta)^2 d\varphi^2\right)\right], \tag{37}$$

with

$$f_\kappa(\vartheta) = \begin{cases} \sin\vartheta, & \kappa = 1, \\ \vartheta, & \kappa = 0, \\ \sinh\vartheta, & \kappa = -1. \end{cases} \tag{38}$$

We require that all the fields are invariant under the symmetries of the metric, namely the time translation isometry generated by ∂_t and the spatial isometries generated by the Killing vectors

$$\partial_\varphi, \qquad \cos\varphi\,\partial_\vartheta - \frac{f_\kappa'}{f_\kappa}\sin\varphi\,\partial_\varphi, \qquad \sin\varphi\,\partial_\vartheta + \frac{f_\kappa'}{f_\kappa}\cos\varphi\,\partial_\varphi. \tag{39}$$

The scalar fields can then only depend on the radial coordinate r, and the request of invariance of the field strength 2-forms F^Λ leads to

$$F^\Lambda = \frac{1}{2}F^\Lambda{}_{\mu\nu}(x)dx^\mu dx^\nu = F^\Lambda{}_{tr}(r)dt\wedge dr + F^\Lambda{}_{\vartheta\varphi}(r,\vartheta)d\vartheta\wedge d\varphi, \tag{40}$$

with

$$F^\Lambda{}_{\vartheta\varphi}(r,\vartheta) = 4\pi p^\Lambda(r) f_\kappa(\vartheta), \tag{41}$$

where $p^\Lambda(r)$ is a generic function of r. The Bianchi identities

$$\nabla_\nu\left(\star F^{\Lambda\nu\mu}\right) = 0 \qquad \Longleftrightarrow \qquad \partial_{[\mu}F^\Lambda{}_{\nu\rho]} = 0 \tag{42}$$

imply that p^Λ must be constant. With field strengths of this form, it is always possible to choose a gauge in which the gauge potential 1-forms can be written as

$$A^\Lambda = A^\Lambda{}_t(r)dt + A^\Lambda{}_\varphi(\vartheta)d\varphi. \tag{43}$$

The r-component of the Maxwell equations (30) reduces then to the condition

$$k_{\Lambda u}(q)\partial_r q^u = 0, \tag{44}$$

while the ϑ-component is automatically satisfied and the φ-component gives

$$A^\Sigma{}_\varphi k_\Sigma{}^u k_{\Lambda u} = 0 \tag{45}$$

for every value of Λ, or equivalently

$$k_\Lambda{}^u(q)\, p^\Lambda = 0. \tag{46}$$

Finally if we define a function $e_\Lambda(r)$ such that

$$F^\Lambda{}_{tr}(r) = 4\pi I^{\Lambda\Sigma}\left(e_\Sigma(r) - R_{\Sigma\Gamma}p^\Gamma\right)e^{2(U-W)}, \tag{47}$$

we have $F_{\Lambda\vartheta\varphi} = 4\pi e_\Lambda(r) f_\kappa(\vartheta)$ and the t-component of the Maxwell equations becomes

$$4\pi e^{2(U-W)}\partial_r e_\Lambda = \frac{g^2}{2}e^{-2U}A^\Sigma{}_t k_\Sigma{}^u k_{\Lambda u}. \tag{48}$$

The quantities p^Λ and $e_\Lambda(r)$ are the magnetic and electric charge densities inside the 2-surfaces S_r of constant r and t,

$$p^\Lambda = \frac{1}{4\pi \mathbf{V}}\int_{S_r} F^\Lambda, \qquad e_\Lambda(r) = \frac{1}{4\pi\mathbf{V}}\int_{S_r} F_\Lambda, \qquad \mathbf{V} = \int_{S_r} f_\kappa(\vartheta)d\vartheta \wedge d\varphi. \tag{49}$$

The r-dependence of e_Λ can be easily understood: Due to (16), the hyperscalars are charged, and thus they contribute to the electric charge densities inside the surfaces S_r.

The non-vanishing components of $T_{\mu\nu}$ are given by

$$T^t_t = T^r_r = -T^\theta_\theta = -T^\varphi_\varphi = (8\pi)^2 e^{4(U-W)}\tilde{V}_{\text{BH}}, \tag{50}$$

where $\tilde{V}_{\text{BH}}$ is the so-called black hole potential,

$$\tilde{V}_{\text{BH}} = -\frac{1}{2}\left(p^\Lambda,\, e_\Lambda(r)\right)\begin{pmatrix} I_{\Lambda\Sigma} + R_{\Lambda\Gamma}I^{\Gamma\Omega}R_{\Omega\Sigma} & -R_{\Lambda\Gamma}I^{\Gamma\Sigma} \\ -I^{\Lambda\Gamma}R_{\Gamma\Sigma} & I^{\Lambda\Sigma}\end{pmatrix}\begin{pmatrix} p^\Sigma \\ e_\Sigma(r)\end{pmatrix}, \tag{51}$$

which however, unlike the usual definition, has an explicit dependence on r through the varying electric charges e_Λ. It is also straightforward, using the expressions (41), (47) and the definition (34), to verify that

$$\partial^i F_\Lambda{}^{\mu\nu} \star F^\Lambda{}_{\mu\nu} = (8\pi)^2 e^{4(U-W)}\partial^i \tilde{V}_{\text{BH}}, \tag{52}$$

where on the left-hand side only the dual field strengths F_Λ are taken to depend on the complex scalars Z^i and only through the matrices $R_{\Lambda\Sigma}$ and $I_{\Lambda\Sigma}$ appearing in (34), while on the right-hand side the charges $e_\Lambda(r)$ are considered to be independent of the Z^i. Equations (29), (31) and (32) then reduce to

$$e^{2U}\left(2U'W' + U''\right) - (8\pi)^2 e^{4(U-W)}\tilde{V}_{\text{BH}} - 2g^2 e^{-2U} A^\Lambda{}_t k_{\Lambda u} A^\Sigma{}_t k_\Sigma{}^u + \frac{V}{2} = 0, \tag{53}$$

$$e^{2U}\left(U'^2 + W'^2 + W''\right) - (8\pi)^2 e^{4(U-W)}\tilde{V}_{\text{BH}} + e^{2U}\mathcal{G}_{i\bar{j}} Z^{i\prime}\bar{Z}^{\bar{j}\prime} + e^{2U} H_{uv} q^{u\prime} q^{v\prime} - g^2 e^{-2U} A^\Lambda{}_t k_{\Lambda u} A^\Sigma{}_t k_\Sigma{}^u + \frac{V}{2} = 0, \tag{54}$$

$$e^{2U}\left(-\kappa e^{-2W} + 2W'^2 + W''\right) - 2g^2 e^{-2U} A^\Lambda{}_t k_{\Lambda u} A^\Sigma{}_t k_\Sigma{}^u + V = 0, \tag{55}$$

$$e^{2U}\left(Z^{i\prime\prime} + 2W' Z^{i\prime} + \mathcal{G}^{i\bar{j}}\partial_l \mathcal{G}_{k\bar{j}} Z^{l\prime} Z^{k\prime}\right) - (8\pi)^2 e^{4(U-W)}\partial^i \tilde{V}_{\text{BH}} - \frac{1}{2}\partial^i V = 0, \tag{56}$$

$$e^{2U}\left(q^{u\prime\prime} + 2W' q^{u\prime} + \Gamma^u_{vz} q^{v\prime} q^{z\prime}\right) - g^2 e^{-2U} A^\Lambda{}_t k_\Lambda{}^v A^\Sigma{}_t \nabla_v k_\Sigma{}^u - \frac{1}{4}\partial^u V = 0, \tag{57}$$

where a prime denotes a derivative with respect to r.

Suppose now to have an extremal black hole, with horizon at $r = 0$, where the geometry becomes $\text{AdS}_2 \times \Sigma$, with $\Sigma = \mathbb{E}^2, \mathbb{H}^2$ or S^2 for $\kappa = 0, -1, 1$ respectively. In the near horizon limit ($r \to 0$) one has

$$U \sim \log \frac{r}{r_{\text{AdS}}}, \qquad W \sim \log\left(\frac{r_H}{r_{\text{AdS}}} r\right), \tag{58}$$

where r_{AdS} is the AdS_2 curvature radius. We require all the fields, their derivatives, the scalar potential and the couplings to be regular on the horizon. Then we can choose a gauge such that

$$A^\Lambda{}_t\big|_{r=0} = 0 \qquad \Longrightarrow \qquad A^\Lambda{}_t \overset{r\to 0}{\sim} F^\Lambda{}_{rt}\big|_{r=0}\, r. \tag{59}$$

It is also reasonable to assume that the derivative of the electric charges $\partial_r e_\Lambda$ remains finite on the horizon. In this case, (48) implies that in the near-horizon limit the quantity $A^\Sigma{}_t k_{\Sigma u} k_\Lambda{}^u$ is at least of order r^2. If we expand in powers of r, in the gauge (59) the order zero term automatically vanishes, while for the order one term we have

$$0 = \partial_r\left(A^\Sigma{}_t k_{\Sigma u} k_\Lambda{}^u\right)\big|_{r=0} = -F^\Sigma{}_{tr} k_{\Sigma u} k_\Lambda{}^u\big|_{r=0} \qquad \Longrightarrow \qquad F^\Lambda{}_{tr} k_\Lambda{}^u\big|_{r=0} = 0. \tag{60}$$

Using (59) and (60) one can see that the terms with $A^\Lambda{}_t$ in the equations of motion, $e^{-2U} A^\Lambda{}_t k_{\Lambda u} A^\Sigma{}_t k_\Sigma{}^u$ and $e^{-2U} A^\Lambda{}_t k_\Lambda{}^v A^\Sigma{}_t \nabla_v k_\Sigma{}^u$, go to zero in the near-horizon limit. In this limit the equations of motion (53)–(57) thus reduce to

$$\frac{1}{r_{\text{AdS}}^2} = (8\pi)^2 \frac{V_{\text{BH}}}{r_H^4} - \frac{V}{2}, \tag{61}$$

$$\frac{\kappa}{r_H^2} = \frac{1}{r_{\mathrm{AdS}}^2} + V, \tag{62}$$

$$\partial_i \left[(8\pi)^2 \frac{V_{\mathrm{BH}}}{r_H^4} + \frac{V}{2} \right] = 0, \tag{63}$$

$$\partial_u V = 0, \tag{64}$$

where $V_{\mathrm{BH}} \equiv \tilde{V}_{\mathrm{BH}}|_{e_\Lambda(r)\to e_\Lambda(0)}$. Solving the first two equations for r_H^2 and r_{AdS}^2 one gets

$$r_H^2 = \left. \frac{\kappa \pm \sqrt{\kappa^2 - 2(8\pi)^2 V_{\mathrm{BH}} V}}{V} \right|_{r=0}, \tag{65}$$

$$r_{\mathrm{AdS}}^2 = \mp \left. \frac{r_H^2}{\sqrt{\kappa^2 - 2(8\pi)^2 V_{\mathrm{BH}} V}} \right|_{r=0}, \tag{66}$$

and since of course $r_{\mathrm{AdS}}^2 > 0$ we have to choose the lower sign. We also have to require $r_H^2 > 0$, which means that flat or hyperbolic geometries, $\kappa = 0, -1$, are only possible if the scalar potential takes negative values on the horizon, $V|_{r=0} < 0$. Spherical geometry ($\kappa = 1$), on the other hand, is compatible with both positive or negative values of V on the horizon, but for $V|_{r=0} > 0$ there is the restriction $V_{\mathrm{BH}} V|_{r=0} < \frac{1}{2(8\pi)^2}$, since V_{BH} is always positive.

We can introduce an effective potential as a function of the scalars,

$$V_{\mathrm{eff}}(Z, \bar{Z}, q) \equiv \frac{\kappa - \sqrt{\kappa^2 - 2(8\pi)^2 V_{\mathrm{BH}} V}}{V}, \tag{67}$$

defined for $V_{\mathrm{BH}} V < \frac{1}{2(8\pi)^2}$, and write

$$r_H^2 = V_{\mathrm{eff}}|_{Z_H, q_H}, \tag{68}$$

$$r_{\mathrm{AdS}}^2 = \left. \frac{V_{\mathrm{eff}}}{\sqrt{\kappa^2 - 2(8\pi)^2 V_{\mathrm{BH}} V}} \right|_{Z_H, q_H}, \tag{69}$$

with $Z_H^i \equiv \lim_{r\to 0} Z^i$, $q_H^u \equiv \lim_{r\to 0} q^u$. Because of (62)–(63), V_{eff} is extremized on the horizon by all the scalar fields of the theory,

$$\partial_i V_{\mathrm{eff}}|_{Z_H, q_H} = 0, \qquad \partial_u V_{\mathrm{eff}}|_{Z_H, q_H} = 0. \tag{70}$$

The values Z_H^i, q_H^u of the scalars on the horizon are then determined by the extremization conditions (70), and the Bekenstein–Hawking entropy density is given by the critical value of V_{eff},

$$s = \frac{S}{\mathbf{V}} = \frac{A}{4\mathbf{V}} = \frac{r_H^2}{4} = \frac{V_{\text{eff}}(Z_H, \bar{Z}_H, q_H)}{4}. \tag{71}$$

For a given theory this critical value, and thus also the entropy, depend only on the charges (on the horizon) p^Λ and $e_\Lambda(0)$, so that the attractor mechanism still works. On the other hand Z_H^i and q_H^u may not be uniquely determined, since in general V_{eff} may have flat directions.

The limit for $V \to 0$ of V_{eff} only exists for $\kappa = 1$, in which case $V_{\text{eff}} \to (8\pi)^2 V_{\text{BH}}$ and one recovers the attractor mechanism for ungauged supergravity. The fact that this limit does not exist for $\kappa = 0, -1$ is not surprising since flat or hyperbolic horizon geometries are incompatible with vanishing cosmological constant.

Note finally that in the case without hypers and U(1) FI gauging, the effective potential (67) was obtained in [12].

4 Supersymmetric and Nonextremal Black Holes in Gauged Supergravity

In this last part of the lectures we shall give some details on how to explicitely construct black hole solutions in gauged supergravity. Similar to the ungauged case, where black holes are typically determined by harmonic functions on a flat base space, we will see that there appears a quite generic structure.

In this section, only the case without hypers ($n_H = 0$) and U(1) FI gauging is considered. Let us first choose the model with prepotential

$$\mathcal{F}(\chi) = -2i(\chi^0\chi^1\chi^2\chi^3)^{1/2}, \tag{72}$$

which has three vector multiplets ($n_V = 3$). Notice that

- For zero axions (which means essentially real Z^i, $i = 1, 2, 3$) and equal FI parameters $\xi\Lambda$, this model can be obtained by compactifying $d = 11$ supergravity on S^7 and truncating to the Cartan subgroup $U(1)^4$ of SO(8) [22].
- In absence of gauging this model is related to

$$\mathcal{F}(\chi) = -\chi^1\chi^2\chi^3/\chi^0 \tag{73}$$

  (cubic prepotential, that appears in what is called 'very special geometry') by a symplectic transformation. As we said in Sect. 2, the gauging breaks symplectic covariance, and thus the physics of (72) and (73) is different!

If we choose the parametrization

$$\chi^0 = 1, \qquad \chi^1 = Z^2 Z^3, \qquad \chi^2 = Z^1 Z^3, \qquad \chi^3 = Z^1 Z^2, \tag{74}$$

the holomorphic symplectic vector (3) becomes

$$\Omega = (1, Z^2 Z^3, Z^1 Z^3, Z^1 Z^2, -iZ^1 Z^2 Z^3, -iZ^1, -iZ^2, -iZ^3)^T, \tag{75}$$

and the Kähler potential and nonvanishing components of the Kähler metric are given respectively by

$$e^{-\mathcal{K}} = 8\, \mathfrak{Re} Z^1\, \mathfrak{Re} Z^2\, \mathfrak{Re} Z^3, \qquad \mathcal{G}_{i\bar{\imath}} = \mathcal{G}_{\bar{\imath}i} = (Z^i + \bar{Z}^{\bar{\imath}})^{-2}. \tag{76}$$

In what follows, we assume the Z^i to be real (this is a consistent truncation) and positive. The latter requirement comes from the positivity of the kinetic terms in the action. Then the kinetic matrix for the vectors reads

$$\mathcal{N} = -i \operatorname{diag}\left(Z^1 Z^2 Z^3, \frac{Z^1}{Z^2 Z^3}, \frac{Z^2}{Z^1 Z^3}, \frac{Z^3}{Z^1 Z^2} \right), \tag{77}$$

and thus

$$R_{\Lambda\Sigma} = 0, \qquad (I^{\Lambda\Sigma}) = -8 \operatorname{diag}\left(\mathcal{L}^{0^2}, \mathcal{L}^{1^2}, \mathcal{L}^{2^2}, \mathcal{L}^{3^2} \right). \tag{78}$$

The scalar potential (15) becomes

$$V = -4g^2 \left(\frac{\xi_0 \xi_1}{Z^1} + \xi_2 \xi_3 Z^1 + \frac{\xi_0 \xi_2}{Z^2} + \xi_1 \xi_3 Z^2 + \frac{\xi_0 \xi_3}{Z^3} + \xi_1 \xi_2 Z^3 \right), \tag{79}$$

which has an extremum at

$$Z^1 = \left(\frac{\xi_0 \xi_1}{\xi_2 \xi_3} \right)^{1/2}, \qquad Z^2 = \left(\frac{\xi_0 \xi_2}{\xi_1 \xi_3} \right)^{1/2}, \qquad Z^3 = \left(\frac{\xi_0 \xi_3}{\xi_1 \xi_2} \right)^{1/2}. \tag{80}$$

Exercise: Verify (75)–(80).

We have thus a model with four U(1) vector fields and three real scalars. A class of black hole solutions to this model was constructed in [23], using the ansatz

$$ds^2 = e^{2(\psi(r) - V(r))} dt^2 - e^{-2(\psi(r) - V(r))} dr^2 - e^{2V(r)} (d\theta^2 + f_\kappa(\theta)^2 d\phi^2), \tag{81}$$

where

$$e^{2\psi(r)} = \sum_{n=0}^{4} a_n r^n \tag{82}$$

is a *quartic* polynomial. Without loss of generality one can take $a_4 = 1$ by using the scaling symmetry

$$t \to t/\lambda, \qquad r \to \lambda r, \qquad a_n \to a_n \lambda^{2-n}$$

of the solution, and $a_3 = 0$ by shifting r. The function $V(r)$ turns out to be given by

$$e^{2V(r)} = (f_0 f_1 f_2 f_3)^{1/2}, \tag{83}$$

where the f_Λ are *linear* functions,

$$f_\Lambda = \alpha_\Lambda r + \beta_\Lambda, \tag{84}$$

with α_Λ and β_Λ constants. The upper part of the symplectic section $\mathcal{V}$ reads

$$\mathcal{L}^\Lambda = \frac{1}{2\sqrt{2}} e^{-V} f_\Lambda. \tag{85}$$

Form this one can read off the scalars. The gauge field strengths are

$$F^\Lambda = p^\Lambda f_\kappa(\theta) d\theta \wedge d\phi, \tag{86}$$

so we have only magnetic charges p^Λ. Dyonic generalizations have been constructed in [24–26].[2]

The equations of motion are then satisfied if and only if the parameters satisfy the equations

$$\alpha_\Lambda = \frac{1}{2\sqrt{2} g \xi_\Lambda}, \qquad \sum_{\Lambda=0}^{3} \xi_\Lambda \beta_\Lambda = 0, \qquad a_2 = \kappa - 4 \sum_{\Lambda=0}^{3} g^2 \xi_\Lambda^2 \beta_\Lambda^2,$$
$$p^{\Lambda 2} = \frac{a_2}{2} \beta_\Lambda^2 + \frac{a_0}{16 g^2 \xi_\Lambda^2} - \frac{a_1 \beta_\Lambda}{4\sqrt{2} g \xi_\Lambda} + 4 g^2 \xi_\Lambda^2 \beta_\Lambda^4. \tag{87}$$

This leaves a five-parameter family of solutions, labeled e.g. by $(\beta_0, \beta_1, \beta_2, a_0, a_1)$, or by four magnetic charges and the mass (that is related to the coefficient a_1).

In what follows, we shall discuss some physical properties of the solution. First of all, there is an event horizon at the largest root r_h of $e^{2\psi} = 0$. The Bekenstein–Hawking entropy is given by

$$S_{\text{BH}} = \frac{e^{2V(r_h)}}{4G} \mathbf{V}, \tag{88}$$

[2] Reference [26] considers only the BPS case, but has a more general class of prepotentials, defining symmetric very special Kähler manifolds.

where the volume $\mathbf{V}$ was defined in (49). Moreover, a so-called area product formula [27] holds: If we decompose

$$e^{2\psi} = \prod_{\alpha=1}^{4} (r - r_\alpha),$$

then the product of all horizon areas (including also possible complex roots r_α) becomes

$$\prod_{\alpha=1}^{4} A(r_\alpha) = \frac{36\mathbf{V}^4}{\Lambda^2} p^0 p^1 p^2 p^3. \tag{89}$$

Here, $A(r_\alpha)$ denotes the area of the α-th horizon, while

$$\Lambda = -24g^2(\xi_0\xi_1\xi_2\xi_3)^{1/2} \tag{90}$$

is the asymptotic value of the cosmological constant. The area product (89) depends thus only on the charges and the asymptotic cosmological constant, in agreement with the analysis in [27]. Notice also that (89) reflects the form of the prepotential. The deeper reason for this fact, which was first observed in [28], remains to be understood.

The BPS limit of the above solution is obtained for [13] $a_1 = 0$ and

$$2g\xi_\Lambda p^\Lambda = -\kappa. \tag{91}$$

This is a Dirac-type quantization condition, due to the minimal coupling of the gravitinos to the linear combination $\xi_\Lambda A^\Lambda$, that is used to gauge a U(1) subgroup of the SU(2) R-symmetry group. Equation (91) can also be viewed as a twisting condition [29] that expresses the cancellation in the gravitino variation $\delta\psi_\mu$ of a piece coming from the spin connection on S^2, $\mathbb{H}^2$ or $\mathbb{E}^2$ with a piece coming from the U(1) connection $\xi_\Lambda A^\Lambda$. If (91) (together with $a_1 = 0$) is satisfied, the black hole preserves two real supercharges. It interpolates between $AdS_2 \times \Sigma$ near the horizon and AdS_4 at infinity. The infrared (near-horizon) geometry is 1/2 BPS (4 real supercharges preserved), while the AdS_4 in the UV is fully supersymmetric (8 real supercharges). This is in contrast to the asymptotically flat case, where one has 1/2 BPS black holes, and a maximally supersymmetric near-horizon geometry.

Notice also that, by allowing for running scalars (as we did here), one can have supersymmetric genuine black holes with spherical horizon [13], which are not possible in minimal gauged supergravity [30] (where there are no scalars). In the latter case, static BPS black holes must have hyperbolic horizons [31].

Let us compare the solution presented here with the one constructed by Duff and Liu [32]. They considered a $U(1)^4$ truncation of SO(8) $N = 8$ gauged supergravity, which is exactly the $N = 2$ model considered in this section, with all FI parameters ξ_Λ equal. The metric, moduli and gauge fields of their solution read (cf. (6.2) of [32])

$$
\begin{aligned}
ds^2 &= -(H_0H_1H_2H_3)^{-1/2} f dt^2 + (H_0H_1H_2H_3)^{1/2}\left(\frac{dr^2}{f} + r^2 d\Omega^2\right),\\
e^{2\phi^{(12)}} &= \frac{H_2H_3}{H_0H_1}, \qquad e^{2\phi^{(13)}} = \frac{H_1H_3}{H_0H_2}, \qquad e^{2\phi^{(14)}} = \frac{H_1H_2}{H_0H_3},\\
H_\Lambda &= 1 + \frac{k\sinh^2\mu_\Lambda}{r}, \qquad f = 1 - \frac{k}{r} + 2g^2r^2H_0H_1H_2H_3,\\
F^\Lambda_{\vartheta\varphi} &= \frac{\eta_\Lambda}{2\sqrt{2}} k\cosh\mu_\Lambda \sinh\mu_\Lambda \sin\vartheta. \qquad (92)
\end{aligned}
$$

Here, $\eta_\Lambda = \pm 1$ are arbitrary signs, the μ_Λ determine the magnetic charges, and k is a sort of nonextremality parameter (although the solution with $k = 0$ is not an extremal black hole, but a naked singularity). After the coordinate transformation

$$
r = \frac{r'}{\sqrt{2}g} - \frac{k}{4}\sum_{\Lambda=0}^{3}\sinh^2\mu_\Lambda,
$$

(and then dropping the prime), the solution (92) takes the form (81), (85) and (86),[3] with FI parameters $\xi_\Lambda = 1/2$, and

$$
\begin{aligned}
\beta_\Lambda &= k\left(\sinh^2\mu_\Lambda - \frac{1}{4}\sum_\Sigma \sinh^2\mu_\Sigma\right),\\
a_0 &= \frac{g^2k^2}{2}\left[\left(\frac{1}{2}\sum_\Lambda \sinh^2\mu_\Lambda\right)^2 + \sum_\Lambda \sinh^2\mu_\Lambda\right] + 4g^4\beta_0\beta_1\beta_2\beta_3,\\
a_1 &= -\sqrt{2}gk\left(1 + \frac{1}{2}\sum_\Lambda \sinh^2\mu_\Lambda\right) - \sqrt{2}\frac{g^3k^3}{4}\left(\sinh^2\mu_0 + \sinh^2\mu_1 - \sinh^2\mu_2\right.\\
&\quad \left.- \sinh^2\mu_3\right)\left[(\sinh^2\mu_2 - \sinh^2\mu_3)^2 - (\sinh^2\mu_0 - \sinh^2\mu_1)^2\right],\\
a_2 &= 1 - g^2\sum_\Lambda \beta_\Lambda^2. \qquad (93)
\end{aligned}
$$

Since both (92) and the solution (81), (85), (86) are labeled by five continuous parameters (k, μ_Λ for (92)), one might wonder if they are not equivalent (if all ξ_Λ are equal; for generic FI parameters the solution (81), (85), (86) is clearly more general). This is however not the case: Suppose for instance that all charges are equal in (92). Then, the 'harmonic' functions H_Λ coincide as well, and thus the scalar fields are constant. In the solution considered here instead, one can have equal charges and yet nontrivial profiles for the moduli (take e.g. $a_1 = 0$, $\beta_0 = \beta_1 = -\beta_2 = -\beta_3$). Moreover, (81)–(87) contains a subclass of black holes that are BPS, while it was shown in [32] that (92) can never be supersymmetric. To understand better what

[3] As an exercise, work out the relation between the scalars Z^i and $\phi^{(12)}$, $\phi^{(13)}$, $\phi^{(14)}$.

happens, let us consider the subcase $\beta_2 = \beta_0$, $\beta_3 = \beta_1$, such that $\mathcal{L}^2 = \mathcal{L}^0$, $\mathcal{L}^3 = \mathcal{L}^1$ and $p^2 = p^0$, $p^3 = p^1$, which amounts to taking the model with prepotential $\mathcal{F} = -i\chi^0\chi^1$ considered in [33]. Since all g_I are equal ($2g_I = g$), the second equation of (87) boils down to $\beta_1 = -\beta_0$. From the last equation of (87) one obtains then

$$p^{1^2} - p^{0^2} = \frac{a_1\beta_0}{g\sqrt{2}}.$$

If the charges are equal (up to a sign), $p^{1^2} = p^{0^2}$, and we have therefore $a_1 = 0$ or $\beta_0 = 0$. The former case is (for $p^1 = p^0$) the supersymmetric black hole found in [13] (with running scalar), whereas the latter corresponds to the Duff–Liu solution (92), with constant scalar profiles. In this context, notice also that in the parametrization (93), for $\mu_2 = \mu_0$, $\mu_3 = \mu_1$, we get

$$a_1 = -\sqrt{2}gk(1 + \sinh^2\mu_0 + \sinh^2\mu_1),$$

which is always nonvanishing (if $k \neq 0$), and thus the supersymmetric case cannot appear.

In conclusion, the solution (81)–(87) contains more than one branch. One of them is the Duff/Liu solution (92), while the other contains the BPS black holes constructed in [13].

Let us conclude this section with some final comments:
First of all, the appearance of a *quartic polynomial* $e^{2\psi}$ and *linear rescaled sections* $e^V \cdot (\mathcal{L}^\Lambda, \mathcal{M}_\Lambda)$ (actually in a different duality frame that has both electric and magnetic gaugings, cf. [34]) is a quite generic feature, that was shown to hold in a lot of models, at least for symmetric very special Kähler manifolds, characterized by a cubic prepotential

$$\mathcal{F} \sim \frac{d_{ijk}\chi^i\chi^j\chi^k}{\chi^0},$$

where the tensor d_{ijk} must satisfy certain properties in order for the special Kähler manifold to be symmetric.
Note that quartic polynomials appear also in the Plebański–Demiański solution [35].[4] This is the complete family of type-D spacetimes with a non-null electromagnetic field, whose two principal null congruences are aligned with the two repeated principal null congruences of the Weyl tensor. It solves the field equations of Einstein–Maxwell–(A)dS gravity and describes a rotating, charged and uniformly accelerating mass. The metric and field strength read respectively

$$ds^2 = \frac{1}{(1-pq)^2}\left\{-\frac{Q(q)}{p^2+q^2}(d\tau - p^2 d\sigma)^2 + \frac{p^2+q^2}{Q(q)}dq^2\right.$$

[4] For a more recent review cf. [36].

$$+\frac{p^2+q^2}{P(p)}dp^2+\frac{P(p)}{p^2+q^2}(d\tau+q^2d\sigma)^2\Big\}\,, \tag{94}$$

$$F=\frac{\mathsf{Q}(p^2-q^2)+2\mathsf{P}pq}{(p^2+q^2)^2}dq\wedge(d\tau-p^2d\sigma)+\frac{\mathsf{P}(p^2-q^2)-2\mathsf{Q}pq}{(p^2+q^2)^2}dp\wedge(d\tau+q^2d\sigma), \tag{95}$$

where the structure functions are given by

$$\begin{aligned}P(p)&=(-\Lambda/6-\mathsf{P}^2+\alpha)+2np-\varepsilon p^2+2mp^3+(-\Lambda/6-\mathsf{Q}^2-\alpha)p^4,\\ Q(q)&=(-\Lambda/6+\mathsf{Q}^2+\alpha)-2mq+\varepsilon q^2-2nq^3+(-\Lambda/6+\mathsf{P}^2-\alpha)q^4.\end{aligned} \tag{96}$$

Here, m and n are the mass and NUT parameters respectively, P and Q represent the magnetic and electric charges, while α, ε are additional non-dynamical parameters.

A subclass of solutions can be obtained by scaling the coordinates according to

$$p\to l^{-1}p,\qquad q\to l^{-1}q,\qquad \tau\to l\tau,\qquad \sigma\to l^3\sigma, \tag{97}$$

and simultaneously adjusting the constants

$$\mathsf{P}\to l^{-2}\mathsf{P},\quad \mathsf{Q}\to l^{-2}\mathsf{Q},\quad m\to l^{-3}m,\quad n\to l^{-3}n,\quad \varepsilon\to l^{-2}\varepsilon,\quad \alpha\to l^{-4}\alpha+\Lambda/6, \tag{98}$$

and taking the limit $l\to\infty$. This removes the acceleration parameter[5] and leads to [35]

$$ds^2=-\frac{Q(q)}{p^2+q^2}(d\tau-p^2d\sigma)^2+\frac{p^2+q^2}{Q(q)}dq^2+\frac{p^2+q^2}{P(p)}dp^2+\frac{P(p)}{p^2+q^2}(d\tau+q^2d\sigma)^2, \tag{99}$$

$$\begin{aligned}P(p)&=\alpha-\mathsf{P}^2+2np-\varepsilon p^2+(-\Lambda/3)p^4,\\ Q(q)&=\alpha+\mathsf{Q}^2-2mq+\varepsilon q^2+(-\Lambda/3)q^4.\end{aligned} \tag{100}$$

The electromagnetic field is still given by (95). Equation (99) is called the Carter–Plebański solution, since it was derived and studied already by Carter [37] and later by Plebański [38]. Notice that one can take a different scaling limit (after the inversion $q\to-1/q$), leading to the cosmological C-metric, that describes either a pair of accelerated black holes (with the acceleration provided by the pressure exerted by a strut), or a single accelerated black hole, depending on the value of the acceleration parameter, cf. the discussion in [39].

The appearance of quartic structure functions raises the question if we can generalize the Plebański–Demiański, Carter–Plebański or C-metric to gauged supergavity with running scalars. For the simple $\mathcal{F}=-i\chi^0\chi^1$ model, this can indeed be done,

[5] The acceleration parameter is essentially given by l^{-2}, as can be seen by comparing (97) and (98) with (3) and (4) of [36].

cf. [24, 40, 41]. We expect the construction of such solutions to have a wide range of applications in AdS/CFT, AdS/cond-mat, black hole microstate counting etc.

References

1. S.A. Hartnoll, Lectures on holographic methods for condensed matter physics. Class. Quantum Gravity **26**, 224002 (2009). arXiv:0903.3246 [hep-th]
2. C. Charmousis, B. Gouteraux, B.S. Kim, E. Kiritsis, R. Meyer, Effective holographic theories for low-temperature condensed matter systems. JHEP **1011**, 151 (2010). arXiv:1005.4690 [hep-th]
3. N. Iizuka, N. Kundu, P. Narayan, S.P. Trivedi, Holographic Fermi and non-Fermi liquids with transitions in dilaton gravity. JHEP **1201**, 094 (2012). arXiv:1105.1162 [hep-th]
4. S.A. Hartnoll, C.P. Herzog, G.T. Horowitz, Building a holographic superconductor. Phys. Rev. Lett. **101**, 031601 (2008). arXiv:0803.3295 [hep-th]
5. S. Ferrara, R. Kallosh, A. Strominger, $N = 2$ extremal black holes. Phys. Rev. D **52**, 5412 (1995). arXiv:hep-th/9508072
6. A. Strominger, Macroscopic entropy of $N = 2$ extremal black holes. Phys. Lett. B **383**, 39 (1996). arXiv:hep-th/9602111
7. S. Ferrara, R. Kallosh, Supersymmetry and attractors. Phys. Rev. D **54**, 1514 (1996). arXiv:hep-th/9602136
8. S. Ferrara, R. Kallosh, Universality of supersymmetric attractors. Phys. Rev. D **54**, 1525 (1996). arXiv:hep-th/9603090
9. S. Ferrara, G.W. Gibbons, R. Kallosh, Black holes and critical points in moduli space. Nucl. Phys. B **500**, 75 (1997). arXiv:hep-th/9702103
10. J.F. Morales, H. Samtleben, Entropy function and attractors for AdS black holes. JHEP **0610**, 074 (2006). arXiv:hep-th/0608044
11. M. Hübscher, P. Meessen, T. Ortín, S. Vaulà, Supersymmetric $N = 2$ Einstein-Yang-Mills monopoles and covariant attractors. Phys. Rev. D **78**, 065031 (2008). arXiv:0712.1530 [hep-th]
12. S. Bellucci, S. Ferrara, A. Marrani, A. Yeranyan, $d = 4$ black hole attractors in $N = 2$ supergravity with Fayet-Iliopoulos terms. Phys. Rev. D **77**, 085027 (2008). arXiv:0802.0141 [hep-th]
13. S.L. Cacciatori, D. Klemm, Supersymmetric AdS_4 black holes and attractors. JHEP **1001**, 085 (2010). arXiv:0911.4926 [hep-th]
14. G. Dall'Agata, A. Gnecchi, Flow equations and attractors for black holes in $N = 2$ U(1) gauged supergravity. JHEP **1103**, 037 (2011). arXiv:1012.3756 [hep-th]
15. S. Kachru, R. Kallosh, M. Shmakova, Generalized attractor points in gauged supergravity. Phys. Rev. D **84**, 046003 (2011). arXiv:1104.2884 [hep-th]
16. L. Andrianopoli, M. Bertolini, A. Ceresole, R. D'Auria, S. Ferrara, P. Fré, T. Magri, $N = 2$ supergravity and $N = 2$ super-Yang-Mills theory on general scalar manifolds: symplectic covariance, gaugings and the momentum map. J. Geom. Phys. **23**, 111 (1997). arXiv:hep-th/9605032
17. D.Z. Freedman, A. Van Proeyen, *Supergravity* (Cambridge University Press, Cambridge, 2012)
18. A. Van Proeyen, "$N = 2$ supergravity in $d = 4, 5, 6$ and its matter couplings," extended version of lectures given during the semester "Supergravity, superstrings and M-theory" at Institut Henri Poincaré, Paris, November 2000, http://itf.fys.kuleuven.ac.be/~toine/home.htm#B
19. N. Halmagyi, BPS black hole horizons in $N = 2$ gauged supergravity. JHEP **1402**, 051 (2014). arXiv:1308.1439 [hep-th]
20. A. Anabalón, D. Astefanesei, On attractor mechanism of AdS_4 black holes. Phys. Lett. B **727**, 568 (2013). arXiv:1309.5863 [hep-th]
21. S. Chimento, D. Klemm, N. Petri, Supersymmetric black holes and attractors in gauged supergravity with hypermultiplets. JHEP **1506**, 150 (2015). arXiv:1503.09055 [hep-th]

22. M. Cvetič et al., Embedding AdS black holes in ten- and eleven dimensions. Nucl. Phys. B **558**, 96 (1999). arXiv:hep-th/9903214
23. D. Klemm, O. Vaughan, Nonextremal black holes in gauged supergravity and the real formulation of special geometry II. Class. Quantum Gravity **30**, 065003 (2013). arXiv:1211.1618 [hep-th]
24. D.D.K. Chow, G. Compère, Dyonic AdS black holes in maximal gauged supergravity. Phys. Rev. D **89**6, 065003 (2014). arXiv:1311.1204 [hep-th]
25. N. Halmagyi, T. Vanel, AdS black holes from duality in gauged supergravity. JHEP **1404**, 130 (2014). arXiv:1312.5430 [hep-th]
26. N. Halmagyi, Static BPS black holes in AdS_4 with general dyonic charges. JHEP **1503**, 032 (2015). arXiv:1408.2831 [hep-th]
27. M. Cvetič, G.W. Gibbons, C.N. Pope, Universal area product formulae for rotating and charged black holes in four and higher dimensions. Phys. Rev. Lett. **106**, 121301 (2011). arXiv:1011.0008 [hep-th]
28. C. Toldo, S. Vandoren, Static nonextremal AdS_4 black hole solutions. JHEP **1209**, 048 (2012). arXiv:1207.3014 [hep-th]
29. J.M. Maldacena, C. Nuñez, Supergravity description of field theories on curved manifolds and a no go theorem. Int. J. Mod. Phys. A **16**, 822 (2001). arXiv:hep-th/0007018
30. L.J. Romans, Supersymmetric, cold and lukewarm black holes in cosmological Einstein-Maxwell theory. Nucl. Phys. B **383**, 395 (1992). arXiv:hep-th/9203018
31. M.M. Caldarelli, D. Klemm, Supersymmetry of anti-de Sitter black holes. Nucl. Phys. B **545**, 434 (1999). arXiv:hep-th/9808097
32. M.J. Duff, J.T. Liu, Anti-de Sitter black holes in gauged $N = 8$ supergravity. Nucl. Phys. B **554**, 237 (1999). arXiv:hep-th/9901149
33. D. Klemm, O. Vaughan, Nonextremal black holes in gauged supergravity and the real formulation of special geometry. JHEP **1301**, 053 (2013). arXiv:1207.2679 [hep-th]
34. A. Gnecchi, N. Halmagyi, Supersymmetric black holes in AdS_4 from very special geometry. JHEP **1404**, 173 (2014). arXiv:1312.2766 [hep-th]
35. J.F. Plebański, M. Demiański, Rotating, charged, and uniformly accelerating mass in general relativity. Ann. Phys. **98**, 98 (1976)
36. J.B. Griffiths, J. Podolsky, A new look at the Plebański-Demiański family of solutions. Int. J. Mod. Phys. D **15**, 335 (2006). arXiv:gr-qc/0511091
37. B. Carter, Hamilton-Jacobi and Schrödinger separable solutions of Einstein's equations. Commun. Math. Phys. **10**, 280 (1968)
38. J.F. Plebański, A class of solutions of Einstein-Maxwell equations. Ann. Phys. **90**, 196 (1975)
39. O.J.C. Dias, J.P.S. Lemos, Pair of accelerated black holes in anti-de Sitter background: AdS C metric. Phys. Rev. D **67**, 064001 (2003). arXiv:hep-th/0210065
40. Z.-W. Chong, M. Cvetič, H. Lu, C.N. Pope, Charged rotating black holes in four-dimensional gauged and ungauged supergravities. Nucl. Phys. B **717**, 246 (2005). arXiv:hep-th/0411045
41. A. Gnecchi, K. Hristov, D. Klemm, C. Toldo, O. Vaughan, Rotating black holes in 4d gauged supergravity. JHEP **1401**, 127 (2014). arXiv:1311.1795 [hep-th]

Lectures on Holographic Renormalization

Ioannis Papadimitriou

Abstract We provide a pedagogical introduction to the method of holographic renormalization, in its Hamiltonian incarnation. We begin by reviewing the description of local observables, global symmetries, and ultraviolet divergences in local quantum field theories, in a language that does not require a weak coupling Lagrangian description. In particular, we review the formulation of the Renormalization Group as a Hamiltonian flow, which allows us to present the holographic dictionary in a precise and suggestive language. The method of holographic renormalization is then introduced by first computing the renormalized two-point function of a scalar operator in conformal field theory and comparing with the holographic computation. We then proceed with the general method, formulating the bulk theory in a radial Hamiltonian language and deriving the Hamilton–Jacobi equation. Two methods for solving recursively the Hamilton–Jacobi equation are then presented, based on covariant expansions in eigenfunctions of certain functional operators on the space of field theory couplings. These algorithms constitute the core of the method of holographic renormalization and allow us to obtain the holographic Ward identities and the asymptotic expansions of the bulk fields.

1 Introduction

The gauge/gravity duality [1] stipulates a mathematical equivalence between a theory of (quantum) gravity and a local quantum field theory (QFT), without gravity, on a lower dimensional space. The best studied examples of such holographic dualities typically involve gravity in an asymptotically anti de Sitter (AdS) space and a dual QFT 'living' on the boundary of AdS. This mathematical equivalence is reflected in a precise map between physical observables on the two sides of the duality. For local observables, this map is summarized in the prescription for computing QFT correlation functions from the gravity dual, originally proposed in [2, 3]. Namely,

I. Papadimitriou (✉)
SISSA and INFN - Sezione di Trieste, Via Bonomea 265, 34136 Trieste, Italy
e-mail: ioannis.papadimitriou@sissa.it

R. Kallosh and E. Orazi (eds.), *Theoretical Frontiers in Black Holes and Cosmology*, Springer Proceedings in Physics 176,
DOI 10.1007/978-3-319-31352-8_4

for every local, single-trace and gauge-invariant operator $\mathcal{O}(x)$ there is a field, Φ, in the dual 'bulk' gravity theory. The generating functional of connected correlation functions of $\mathcal{O}(x)$, $W[J]$, is then identified with the bulk on-shell action

$$W[J] \sim S_{\text{on-shell}}[\Phi]|_{\Phi\sim J}\,, \tag{1}$$

evaluated on solutions of the bulk equations of motion subject to Dirichlet boundary conditions on the AdS boundary. The arbitrary function that is kept fixed at the boundary is identified with the source $J(x)$. This statement is an operational definition of the holographic dictionary, allowing one to compute, in principle, any local QFT observable from the bulk theory. However, there are a number of practical and conceptual obstacles.

The most obvious technical difficulty is that both sides of (1) actually involve infinite quantities. On the QFT side, we know that the generating functional of composite operators generically possesses ultraviolet (UV) divergences, even in a conformal field theory (CFT). We will see an explicit example of this phenomenon later on. On the gravity side, the on-shell action is also generically divergent, due to the infinite volume of AdS space. In order to make sense of (1), therefore, one must somehow remove the divergences from both sides and identify the remaining finite expressions. On the QFT side the procedure for systematically and consistently removing the UV divergences is known as *renormalization*. *Holographic renormalization* [4–13] is the analogous procedure for the gravity side of the duality.

A more conceptual drawback of the identification (1) is that it only maps certain objects on the two sides of the duality, such as the on-shell action and the generating function. However, the bulk fields, or indeed the equations of motion in the bulk are not given any concrete meaning on the QFT side, except from the indirect role in evaluating the on-shell action. As we shall see, both the Renormalization Group (RG) of local QFTs and the dual gravitational theories admit a Hamiltonian description that allows us to formulate the holographic dictionary more precisely.

These lecture notes are organized as follows. In Sect. 2 we discuss local QFT observables and global symmetries in a language that does not assume a weak coupling or Lagrangian description. Moreover, we put forward a Hamiltonian formulation of the Renormalization Group of local QFTs that directly parallels the description of the holographic dual bulk theory later on. We end Sect. 2 with a concrete example of UV divergences in the two-point function of a scalar operator in a CFT. In Sect. 3 we carry out explicitly the holographic computation for the two-point function on a fixed AdS background and reproduce the renormalized result obtained from the CFT calculation. The Hamiltonian formulation of the holographic dictionary is presented in Sect. 3.2. Section 4 discusses at length the radial Hamiltonian formulation of the bulk dynamics for Einstein–Hilbert gravity coupled to a self interacting scalar. In Sect. 5 we present two algorithms for recursively solving the radial Hamilton–Jacobi equation, which constitutes the core of holographic renormalization. Given the solution of the Hamilton–Jacobi equation derived in Sect. 5, in Sect. 6 we provide general expressions for the renormalized one-point functions in the presence of sources and derive the holographic Ward identities. Finally, in Sect. 7 we show how the asymptotic expansions of the bulk fields can be obtained systematically from the solution of the Hamilton–Jacobi equation. Some background material is presented in the appendices. In particular, Appendix "Hamilton–Jacobi primer" is a self contained review of Hamilton–Jacobi theory in classical mechanics.

2 Local QFT Observables and the Local Renormalization Group

Before we delve into the details of the holographic dictionary and the computation of QFT observables from the bulk gravitational theory, it is instructive to review some basic aspects of QFTs and to put them in a language that will later help us make contact with the holographic dual bulk theory. In particular, since the gauge/gravity duality relates the strongly coupled regime of local QFTs to the bulk gravity theory, it is crucial to describe the local QFT observables and their properties in a way that is valid at strong coupling. Ideally we would like to discuss local QFT observables without assuming the existence of a microscopic Lagrangian description of the QFT.

2.1 QFT Correlation Functions and the Generating Functional

The basic objects of a local QFT are correlation functions of local operators, $\mathcal{O}(x)$, namely

$$\langle \mathcal{O}_1(x_1)\mathcal{O}_2(x_2)\ldots\mathcal{O}_n(x_n)\rangle. \tag{2}$$

In particular, if we know all correlation functions of all local operators of a local QFT, then in most cases we know all there is to know about this theory.[1] In a generic theory, even if there is only a finite number of local operators present in a given QFT, the number of correlation functions that we need to know can be infinite. So, instead of having to deal with an infinite number of correlation functions, it is useful to introduce the generating function of correlation functions, $Z[J]$, as a book keeping device. For a single local operator $\mathcal{O}(x)$, the generating function takes the form

$$Z[J]=\sum_{k=0}^{\infty}\frac{1}{k!}\int d^dx_1\int d^dx_2\ldots\int d^dx_k J(x_1)J(x_2)\ldots J(x_k)\langle\mathcal{O}(x_1)\mathcal{O}(x_2)\ldots\mathcal{O}(x_k)\rangle, \tag{3}$$

[1] Sometimes, additional global observables must be specified to uniquely identify a theory [14].

where d is the spacetime dimension. Given $Z[J]$, any correlation function of the operator $\mathcal{O}(x)$ can be extracted by multiple functional differentiation:

$$\langle \mathcal{O}(x_1)\mathcal{O}(x_2)\ldots\mathcal{O}(x_k)\rangle = \left.\frac{\delta^k Z[J]}{\delta J(x_1)\delta J(x_2)\ldots\delta J(x_k)}\right|_{J=0}. \tag{4}$$

These definitions straightforwardly generalize to a set of local operators $\{\mathcal{O}_1(x), \mathcal{O}_2(x), \ldots\}$, with the corresponding generating functional $Z[J_1, J_2, \ldots]$ depending on the sources $J_1(x)$, $J_2(x)$, ... Moreover, the definition of the generating functional through (3) is completely general and it does not assume a Lagrangian description of the theory. Of course, if the theory admits a Lagrangian description, then the generating functional $Z[J]$ has the standard path integral representation

$$Z[J] = \int \mathcal{D}\phi e^{i\int d^dx\mathcal{L}(\phi)+\int d^dx J(x)\mathcal{O}(x)}, \tag{5}$$

where ϕ here stand for the elementary Lagrangian fields.

An alternative but equivalent way to encode all local observables is in terms of the generating function of *connected* correlation functions

$$W[J] = \log Z[J], \tag{6}$$

or

$$W[J] = \sum_{k=0}^{\infty}\frac{1}{k!}\int d^dx_1\int d^dx_2\ldots\int d^dx_k J(x_1)J(x_2)\ldots J(x_k)\langle\mathcal{O}(x_1)\mathcal{O}(x_2)\ldots\mathcal{O}(x_k)\rangle_c, \tag{7}$$

where $\langle\mathcal{O}(x_1)\mathcal{O}(x_2)\ldots\mathcal{O}(x_k)\rangle_c$ are now connected correlation functions. The first derivative of the generating function (7) corresponds to the one-point function of the dual operator *in the presence of an arbitrary source*, namely

$$\langle\mathcal{O}(x)\rangle_J = \frac{\delta W[J]}{\delta J(x)}. \tag{8}$$

Taking further derivatives with respect to the source we can obtain any desired correlation function of the operator $\mathcal{O}(x)$. In particular, the one-point function in the presence of sources (8) encodes the same local information as the generating function (7). This fact will be crucial for the discussion of the holographic dictionary later on.

Another important aspect of (8) is that it amounts to a prescription for the insertion of the local operator $\mathcal{O}(x)$ in any correlation function and so, in effect, it provides a *definition* of the local operator $\mathcal{O}(x)$. This is indeed the point of view adopted in the so called *local Renormalization Group* formulation of QFT [15], where local operators are defined as derivatives of the generating function with respect to the corresponding local coupling. For example, the stress tensor, a $U(1)$ current and a scalar operator are *defined* through the relations

$$\mathcal{T}_{ij}(x) = -\frac{2}{\sqrt{g}}\frac{\delta W}{\delta g^{ij}(x)}, \tag{9a}$$

$$\mathcal{J}^{i}(x) = -\frac{1}{\sqrt{g}}\frac{\delta W}{\delta A_i(x)}, \tag{9b}$$

$$\mathcal{O}(x) = -\frac{1}{\sqrt{g}}\frac{\delta W}{\delta \varphi(x)}, \tag{9c}$$

where g_{ij} is a general background metric on the space where the QFT is defined, and A_i is an Abelian background gauge field. The indices $i, j = 1, 2, \ldots, d$ run over all coordinates parameterizing the space where the QFT is defined.

2.2 The Local Renormalization Group as a Hamiltonian Flow

The expressions (9) for the one-point functions in the presence of sources bare striking resemblance to the expression for the canonical momenta in classical Hamilton–Jacobi (HJ) theory. In particular, the one-point functions (9) look mathematically identical to the expressions (199) or (205) for the canonical momenta in Appendix "Hamilton–Jacobi primer", where we review some basic aspects of HJ theory that we will use repeatedly throughout these lectures.

This analogy turns out to be particularly useful for developing the holographic dictionary and can be formalized as follows [16]. Let $\mathcal{Q}$ be the space of functions (more generally tensors) on the spacetime, S, where the QFT resides (e.g. $\mathbb{R}^d$). The sources $J^\alpha(x)$ are coordinates on $\mathcal{Q}$, which is the analogue of the configuration space in classical mechanics. Let us extend this configuration space to $\mathcal{Q}_{\text{ext}} = \mathcal{Q} \times \mathbb{R}$, by appending an abstract "time" τ to the generalized coordinates $J^\alpha(x)$ as in Appendix "Hamilton–Jacobi primer" in the case of a time-dependent Hamiltonian. Accordingly, an abstract Hamiltonian operator, $\mathbb{H}$, must be introduced as conjugate momentum to τ. Note that $\mathbb{H}$ is a global operator, i.e. it does not depend on x.[2] The extended phase space is then parameterized by the variables

$$\{\mathcal{O}_\alpha(x), \mathbb{H}; J^\alpha(x), \tau\}, \tag{10}$$

and it is isomorphic to the cotangent bundle $T^*\mathcal{Q}_{\text{ext}}$, which is endowed with the pre-symplectic form

$$\Theta = \int d^d x \mathcal{O}_\alpha(x)\delta J^\alpha(x) - \mathbb{H} d\tau, \tag{11}$$

[2] To make contact with [16] one can introduce a Hamiltonian density, $\mathbb{h}(x)$, through $\mathbb{H} = \int d^d x\ \mathbb{h}(x)$.

and the canonical symplectic closed 2-form

$$\Omega = \int d^d x \delta \mathcal{O}_\alpha(x) \wedge \delta J^\alpha(x) - d\mathbb{H} \wedge d\tau, \tag{12}$$

that can be written locally as $\Omega = \delta\Theta$.

Any functional, $F[J;\tau]$, provides a closed section of the cotangent bundle, s: $\mathcal{Q}_{\text{ext}} \longrightarrow T^*\mathcal{Q}_{\text{ext}}$, given locally by

$$s = \delta F[J;\tau]. \tag{13}$$

It follows that

$$\Theta \circ s = \delta F[J;\tau], \tag{14}$$

or equivalently

$$\mathcal{O}_\alpha = \frac{\delta F[J;\tau]}{\delta J^\alpha}, \quad \mathbb{H} = -\frac{\partial F[J;\tau]}{\partial \tau}, \tag{15}$$

while

$$\Omega \circ s = \int d^d x \int d^d x' \frac{\delta^2 F[J;\tau]}{\delta J^\beta(x')\delta J^\alpha(x)} \delta J^\beta(x') \wedge \delta J^\alpha(x) - \frac{\partial^2 F[J;\tau]}{\partial \tau^2} d\tau \wedge d\tau = 0. \tag{16}$$

As follows from the Hamilton–Jacobi theorem (see Appendix "Hamilton–Jacobi primer"), the τ-evolution of all the variables is then governed by Hamilton's equations

$$\dot{J}^\alpha = \frac{\delta \mathbb{H}}{\delta \mathcal{O}_\alpha}, \quad \dot{\mathcal{O}}_\alpha = -\frac{\delta \mathbb{H}}{\delta J^\alpha}, \quad \dot{\mathbb{H}} = \frac{\partial \mathbb{H}}{\partial \tau}. \tag{17}$$

Note that the functional derivatives in (15) and (17) are partial derivatives.

There are two different closed sections of the cotangent bundle $T^*\mathcal{Q}_{\text{ext}}$ one can naturally define for any local QFT. Taking τ to be related to some generic energy scale μ via $\tau = \log(\mu/\mu_o)$, where μ_o is some constant reference scale, the bare and renormalized generating functions, respectively $W[J]$ and $W_{\text{ren}}[J;\tau]$, provide two distinct closed sections of the cotangent bundle $T^*\mathcal{Q}_{\text{ext}}$. The difference between these two functionals is that $W_{\text{ren}}[J;\tau]$ is RG invariant, i.e. given $\sigma : \mathbb{R} \longrightarrow \mathcal{Q}$, its total derivative with respect to τ vanishes, $\dot{W}_{\text{ren}}[\sigma(\tau);\tau] = 0$, while $W[J]$ is *not* an RG invariant. The total derivative of $W[J]$ with respect to τ gives, by construction, the Legendre transform of the Hamiltonian $\mathbb{H}$, i.e. the associated Lagrangian[3]

[3]Note that in [16] only the RG invariant $W_{\text{ren}}[J;\tau]$ is considered, written in terms of the bare and renormalized couplings. $W[J]$ is not discussed at all in that reference.

$$\dot{W}[J] = \mathbb{L} = \int d^dx\, \dot{J}^\alpha \mathcal{O}_\alpha - \mathbb{H} = \int d^dx \beta^\alpha \mathcal{O}_\alpha - \mathbb{H}, \tag{18}$$

where $\beta^\alpha = \dot{J}^\alpha$ are the beta functions of the couplings J^α. Moreover, $W[J]$ depends on τ only through the couplings J^α, while $W_{\text{ren}}[J;\tau]$ can also depend *explicitly* on τ through the conformal anomaly. Through (15), these two sections define different local operators and Hamiltonians, which are related through a canonical transformation [17].

Renormalized RG Hamiltonian

Taking $F[J;\tau] = W_{\text{ren}}[J;\tau]$, the first equation in (15) is just the renormalized version of the local RG definition of local operators that we saw above in (9), namely[4]

$$\mathcal{O}_\alpha^{\text{ren}} = \frac{\delta W_{\text{ren}}[J;\tau]}{\delta J^\alpha}. \tag{19}$$

The second equation in (15), with $F[J;\tau] = W_{\text{ren}}[J;\tau]$, can be viewed as a *definition* of the Hamiltonian $\mathbb{H}_{\text{ren}}$ in QFT. In particular, we conclude that $\mathbb{H}_{\text{ren}}$ is numerically equal to the conformal anomaly,

$$\boxed{\mathbb{H}_{\text{ren}} = -\frac{\partial W_{\text{ren}}[J;\tau]}{\partial \tau} = -\int d^dx \sqrt{g}\mathcal{A},} \tag{20}$$

where $\mathcal{A}$ is the conformal anomaly.

Bare RG Hamiltonian

Taking $F[J;\tau] = W[J]$, on the other hand, provides a section of $T^*\mathcal{Q}$. The first equation in (15) is then identical to the local RG expressions (9), while the second equation in (15) implies that the bare RG Hamiltonian vanishes identically

$$\boxed{\mathbb{H} = -\frac{\partial W[J]}{\partial \tau} = 0.} \tag{21}$$

As we mentioned above, the bare and renormalized Hamiltonians, as well as the corresponding local operators, are related by a canonical transformation whose generating function (in the sense of canonical transformations) is given by the local counterterms, $W_{\text{ct}}[J;\tau]$, [17]. Note that the explicit τ-dependence of $W_{\text{ren}}[J;\tau]$ is entirely due to the local counterterms and, in particular, the conformal anomaly. Under this canonical transformation

$$W[J] \longrightarrow W_{\text{ren}}[J;\tau] = W[J] + W_{\text{ct}}[J;\tau]. \tag{22}$$

[4] The way we have defined the operators $\mathcal{O}_\alpha$ and $\mathbb{H}$ in this subsection, they are in fact densities with respect to the background metric g_{ij}, i.e. we have not divided by $\sqrt{g}$ as in (9). Moreover, $\mathcal{O}_\alpha$ include the stress tensor.

RG Equations

The RG equations for the generating functions $W[J]$ and $W_{\rm ren}[J;\tau]$ are respectively

$$\mathbb{L} = \dot{W} = \int d^dx \beta^\alpha \mathcal{O}_\alpha \Leftrightarrow \mathbb{H} = 0, \tag{23a}$$

$$0 = \dot{W}_{\rm ren} = \int d^dx \beta^\alpha \mathcal{O}_\alpha^{\rm ren} + \frac{\partial W_{\rm ren}}{\partial \tau} = \int d^dx \beta^\alpha \mathcal{O}_\alpha^{\rm ren} + \int d^dx \sqrt{g}\mathcal{A}. \tag{23b}$$

The first of these equations is just the HJ equation (21). Comparing the second equation with the HJ equation (20) we conclude that the renormalized Hamiltonian takes the form

$$\boxed{\mathbb{H}_{\rm ren} = \int d^dx \beta^\alpha \mathcal{O}_\alpha^{\rm ren},} \tag{24}$$

where the sum in this expression is over all operators in the theory, including the stress tensor. Given the beta functions as functions of the local running couplings J^α, this Hamiltonian is linear in the canonical momenta, i.e. in $\mathcal{O}_\alpha^{\rm ren}$ [16]. The standard renormalization procedure in QFT is equivalent to determining the beta functions as functions of the local running couplings and $W_{\rm ren}[J;\tau]$ through the HJ equation (20), i.e.

$$\boxed{\left(\int d^dx \beta^\alpha[J] \frac{\delta}{\delta J^\alpha} + \frac{\partial}{\partial \tau}\right) W_{\rm ren}[J;\tau] = 0.} \tag{25}$$

This is the standard RG equation.

Given $\beta^\alpha[J]$ one can integrate the first Hamilton equation in (17) to obtain

$$\mathbb{H} = \int d^dx \beta^\alpha[J]\mathcal{O}_\alpha + \mathcal{F}[J;\tau], \tag{26}$$

for some unspecified $\mathcal{F}[J;\tau]$. Combining this relation with the fact that $\mathbb{H}$ and $\mathbb{H}_{\rm ren}$ are related by a canonical transformation generated by $W_{\rm ct}[J;\tau]$, namely

$$\mathbb{H} - \mathbb{H}_{\rm ren} + \frac{\partial W_{\rm ct}}{\partial \tau} = 0, \tag{27}$$

we deduce that

$$\mathcal{F}[J;\tau] = \left(\int d^dx \beta^\alpha[J] \frac{\delta}{\delta J^\alpha} - \frac{\partial}{\partial \tau}\right) W_{\rm ct}[J;\tau], \tag{28}$$

and hence

$$\boxed{\mathbb{H}[O_\alpha, J^\beta] = \int d^d x \beta^\alpha[J]\mathcal{O}_\alpha + \left(\int d^d x \beta^\alpha[J]\frac{\delta}{\delta J^\alpha} - \frac{\partial}{\partial \tau}\right) W_{\text{ct}}[J;\tau].} \tag{29}$$

However, if the beta functions are not just functions of the running couplings, but depend linearly on the local operators $\mathcal{O}_\alpha$, i.e.

$$\beta^\alpha[O, J] = \mathcal{G}^{\alpha\beta}[J]\mathcal{O}_\beta, \tag{30}$$

then the first of Hamilton's equations in (17) gives

$$\mathbb{H} = \frac{1}{2}\int d^d x\; \mathcal{G}^{\alpha\beta}[J]\mathcal{O}_\alpha\mathcal{O}_\beta + \widetilde{\mathcal{F}}[J;\tau], \tag{31}$$

for some unspecified $\widetilde{\mathcal{F}}[J;\tau]$. Notice that if the beta functions take the form (30), then the RG flow is a *gradient flow*, since $\beta^\alpha = \mathcal{G}^{\alpha\beta}\delta W/\delta J^\beta$. As we shall see, this form of the beta functions and of the Hamiltonian $\mathbb{H}$ are directly related to the bulk holographic description of the theory.

2.3 Global Symmetries and Ward Identities

A general property of QFTs is that they typically possess a number of global symmetries. For example, a relativistic QFT on flat Minkowski space possesses Poincaré symmetry. If the theory is additionally scale invariant, then it will generically possess conformal symmetry. Such theories are known as conformal field theories (CFTs) and the fact that they are conformally invariant allows us to make sense of them on curved backgrounds that are conformally related to flat Minkowski space. Other examples of global symmetries include internal symmetries such as $SU(2)$ isospin (for massless up and down quarks) or supersymmetry.

In QFTs that admit a classical Lagrangian description, global symmetries manifest themselves as invariances of the classical action and lead via Noether's theorem to conserved currents. For example, Poincaré invariance of the classical action implies that the stress-energy tensor, $\mathcal{T}_{ij}$, is conserved, i.e.

$$\partial^i \mathcal{T}_{ij} = 0. \tag{32}$$

Similarly, global internal symmetries lead to conserved currents $\mathcal{J}^i$,

$$\partial_i \mathcal{J}^i = 0. \tag{33}$$

At the quantum level these currents become quantum operators and their classical conservation laws imply relations among certain correlation functions that involve these currents. These identities, relating various correlation functions as a result of the classical Noether theorem, are known as *Ward identities*.

It is often the case, however, that some of the classical symmetries are broken at the quantum level. This happens because in a QFT various quantities contain ultraviolet divergences which must be regulated and renormalized to yield a well defined quantity. However, there may not exist a regulator that preserves all of the classical symmetries of the theory, which leads to the breaking of some symmetries at the quantum level. This breaking of the classical symmetries at the quantum level leads to the so-called *quantum anomalies* in the Ward identities.

A particularly elegant way to derive the Ward identities of a quantum field theory, without relying on a classical Lagrangian description of the theory, is to work with the generating functional of correlation functions and gauge the global symmetries by promoting the sources of the corresponding conserved currents to gauge fields. Among all operators in any QFT there is always the stress tensor, $\mathcal{T}_{ij}$, and let us assume that there is in addition an internal $U(1)$ symmetry giving rise to a current, $\mathcal{J}^i$, in the spectrum of operators. Moreover, to be generic, let us suppose that there is also a scalar operator, $\mathcal{O}$, transforming trivially both under the Poincaré group and the $U(1)$ symmetry, but has definite scaling dimension Δ. The generating functional of connected correlation functions will then be a function of the sources, g_{ij}, A_i, φ, respectively for the stress tensor, the current of the internal symmetry, and for the scalar operator, as well as for all other operators in the theory which we will not need to consider:

$$W[g, A, \varphi, \ldots]. \tag{34}$$

As we would now do in a classical Lagrangian description of the theory to derive Noether's theorem, we gauge the global symmetries by promoting the Poincaré transformations to diffeomorphisms and the internal global symmetry to a local gauge symmetry, while promoting the sources[5] ${g_{(0)}}^{ij}$ and $A_{(0)i}$ to gauge fields of the corresponding local symmetries. In a classical Lagrangian description this would amount to introducing 'minimal couplings' in the Lagrangian. Under infinitesimal diffeomorphisms, parameterized by the vector $\xi^i(x)$, the sources then transform as

$$\begin{aligned}&\delta_\xi g^{ij}_{(0)} = -(D^i_{(0)}\xi^j + D^j_{(0)}\xi^i), \quad \delta_\xi A_{(0)i} = A_{(0)j} D_{(0)i}\xi^j + \xi^j D_{(0)j} A_{(0)i},\\ &\delta_\xi \varphi_{(0)} = \xi^j D_{(0)j}\varphi_{(0)},\end{aligned} \tag{35}$$

while under infinitesimal $U(1)$ gauge transformations, parameterized by the gauge function $\alpha(x)$, they transform as

$$\delta_\alpha g_{(0)ij} = 0, \quad \delta_\alpha A_{(0)i} = D_{(0)i}\alpha(x), \quad \delta_\alpha \varphi_{(0)} = 0, \tag{36}$$

[5]The subscript (0) here is intended to help make contact with the holographic computation later.

where $D_{(0)i}$ denotes the covariant derivative with respect to the metric $g_{(0)ij}$. The Ward identities now can be stated very simply and generally as

$$\delta_\xi W = 0, \quad \delta_\alpha W = 0, \quad \forall \quad \xi^i, \alpha, \tag{37}$$

respectively following from the Poincaré and $U(1)$ symmetries. We can manipulate these expressions a bit further to bring the Ward identities in a more familiar form. Starting with the $U(1)$ Ward identity we have

$$\begin{aligned} \delta_\alpha W = 0 \Leftrightarrow \int d^d x \left(\delta_\alpha g_{(0)}{}^{ij} \frac{\delta W}{\delta g_{(0)}{}^{ij}} + \delta_\alpha A_{(0)i} \frac{\delta W}{\delta A_{(0)i}} + \delta_\alpha \varphi_{(0)} \frac{\delta W}{\delta \varphi_{(0)}} \right) = 0 \\ \Leftrightarrow \int d^d x D_{(0)i} \alpha(x) \frac{\delta W}{\delta A_{(0)i}} = 0 \Leftrightarrow \int d^d x \alpha(x) D_{(0)i} \left(\frac{\delta W}{\delta A_{(0)i}} \right) = 0, \end{aligned} \tag{38}$$

where we have integrated by parts in the last step and have dropped the boundary term. Since $\alpha(x)$ is arbitrary, it follows that the $U(1)$ Ward identity is equivalent to the identity

$$D_{(0)i} \left(\frac{\delta W}{\delta A_{(0)i}} \right) = 0. \tag{39}$$

We can now repeat this exercise for diffeomorphisms to obtain

$$\delta_\xi W = 0 \Leftrightarrow D^i_{(0)} \left(2 \frac{\delta W}{\delta g_{(0)}{}^{ij}} \right) - F_{(0)ij} \frac{\delta W}{\delta A_{(0)i}} + \frac{\delta W}{\delta \varphi_{(0)}} D_{(0)j} \varphi_{(0)}(x) = 0, \tag{40}$$

where $F_{(0)ij} = \partial_i A_{(0)j} - \partial_j A_{(0)i}$ is the field strength of the gauge field $A_{(0)i}$.

In terms of the one-point functions in the presence of sources the above Ward identities take the simple form

$$\boxed{D_{(0)i} \langle \mathcal{J}^i(x) \rangle = 0,} \tag{41}$$

$$\boxed{D^i_{(0)} \langle \mathcal{T}_{ij}(x) \rangle - \langle \mathcal{J}^i(x) \rangle_s F_{(0)ij} + \langle \mathcal{O}(x) \rangle D_{(0)j} \varphi_{(0)}(x) = 0,} \tag{42}$$

following respectively from $U(1)$ and Poincaré invariance.

Finally, let us consider Weyl transformations, i.e. *local* scale transformations, parameterized by the Weyl factor $\sigma(x)$. Under infinitesimal Weyl transformations the sources transform as

$$\delta_\sigma g_{(0)}{}^{ij} = -2\delta\sigma(x) g_{(0)}{}^{ij}, \quad \delta_\sigma A_{(0)i} = 0, \quad \delta_\sigma \varphi_{(0)} = -(d - \Delta)\delta\sigma(x)\varphi_{(0)}, \tag{43}$$

where Δ is the conformal dimension of the operator $\mathcal{O}(x)$ and we focus here on a CFT since scale invariance is not a symmetry of a generic QFT. As we have seen, even if our theory is a conformal field theory, the generating functional of renormalized correlation functions will not be in general invariant under such a Weyl transformation. The variation of the generating functional with respect to Weyl transformations defines the *conformal anomaly*

$$\delta_\sigma W = \int d^d x \sqrt{g_{(0)}} \delta\sigma(x) \mathcal{A}, \tag{44}$$

where the anomaly density, $\mathcal{A}$ is a local function of the sources. Using the above transformation of the sources, this then leads to the trace Ward identity

$$\boxed{\langle \mathcal{T}_i^i(x) \rangle = -(d-\Delta)\varphi_{(0)} \langle \mathcal{O}(x) \rangle + \mathcal{A}.} \tag{45}$$

We recognize this Ward identity as the local version of the RG equation (25), at a fixed point of the renormalization group.

2.4 UV Divergences and Renormalization of Composite Operators

Let us now address in more detail the question of renormalization in QFT with a simple example. This will allow us to directly compare with a holographic calculation in the next subsection in order to get a first idea of the holographic dictionary.

Consider a CFT with a scalar operator $\mathcal{O}_\Delta(x)$ of conformal dimension Δ. Conformal symmetry determines the two-point function up to an overall constant, namely

$$\langle \mathcal{O}_\Delta(x) \mathcal{O}_\Delta(y) \rangle = \frac{c(g, \Delta)}{|x-y|^{2\Delta}}, \tag{46}$$

where c is an arbitrary constant, depending on the dimension Δ and possibly any coupling constants, g, of the CFT, that we could absorb into the normalization of the operator $\mathcal{O}_\Delta$, but we will not. Depending on the conformal dimension, Δ, this correlator may suffer from short distance singularities. Consider the case $\Delta = d/2 + k + \varepsilon$, where ε is an infinitesimal parameter and k is a non-negative integer. Iterating the identity

$$\frac{1}{|x-y|^{2\Delta}} = \frac{1}{2(\Delta-1)(2\Delta-d)} \square \frac{1}{|x-y|^{2\Delta-2}}, \qquad |x-y| \neq 0, \tag{47}$$

where $\square = \delta^{ij}\partial_i\partial_j$, $k+1$ times, we find

$$\begin{aligned} \frac{1}{|x-y|^{2\Delta}} &= \frac{1}{2\varepsilon} \frac{\Gamma(1+\varepsilon)\Gamma(d/2+\varepsilon)}{2^{2k}\Gamma(k+1+\varepsilon)\Gamma(d/2+k+\varepsilon)} \frac{1}{d-2+2\varepsilon} \square^{k+1} \frac{1}{|x-y|^{d-2+2\varepsilon}} \\ &\sim \frac{-1}{2\varepsilon} \frac{\omega_{d-1}\Gamma(d/2)}{2^{2k}\Gamma(k+1)\Gamma(d/2+k)} \square^k \delta^{(d)}(x-y), \end{aligned} \tag{48}$$

where $\omega_{d-1} = 2\pi^{d/2}/\Gamma(d/2)$ is the volume of the unit $(d-1)$-sphere and we have used the identity $\Box(x^2)^{-d/2+1} = -(d-2)\omega_{d-1}\delta^{(d)}(x)$. We thus find that there is a pole at $\Delta = d/2 + k$, or $\varepsilon = 0$. To produce a well defined distribution we subtract the pole and define [18]

$$\langle \mathcal{O}_\Delta(x)\mathcal{O}_\Delta(0)\rangle_{\text{ren}} = c(g, \Delta) \lim_{\varepsilon\to 0} \left\{ \frac{1}{2\varepsilon} \frac{\Gamma(1+\varepsilon)\Gamma(d/2+\varepsilon)}{2^{2k}\Gamma(k+1+\varepsilon)\Gamma(d/2+k+\varepsilon)} \frac{1}{d-2+2\varepsilon}\Box^{k+1}\frac{1}{|x|^{d-2}}\left(\frac{1}{|x|^{2\varepsilon}} - \mu^{2\varepsilon}\right)\right\}$$
$$= \frac{-c_k}{2(d-2)}\Box^{k+1}\frac{1}{|x|^{d-2}}\left\{\log\left(\mu^2 x^2\right) + a(k)\right\}, \tag{49}$$

where

$$c_k \equiv c(g, \Delta)\frac{\Gamma(d/2)}{2^{2k}\Gamma(k+1)\Gamma(d/2+k)}. \tag{50}$$

The constant $a(k)$ reflects the scheme dependence in the subtraction of the pole. Here we have defined the subtraction in such a way so that $a = 0$, but other subtraction schemes, such as minimal subtraction, lead to a non-zero a. The renormalized correlator agrees with the bare one away from coincident points but is also well-defined at $x^2 = 0$. To allow a direct comparison of the renormalized two-point function with the result we will obtain below from the bulk calculation, it is useful to write down its Fourier transform. Using the identity

$$\int d^d x e^{ip\cdot x}\frac{1}{|x|^{d-2}}\log\left(\mu^2 x^2\right) = -\frac{4\pi^{d/2}}{\Gamma(d/2-1)}\frac{1}{p^2}\log(p^2/\bar{\mu}^2), \tag{51}$$

where $\bar{\mu} = 2\mu/\gamma$ and $\gamma = 1.781072\ldots$ is the Euler constant, we obtain

$$\langle \mathcal{O}_\Delta(p)\mathcal{O}_\Delta(-p)\rangle_{\text{ren}} = c_k\frac{(-1)^{k+1}}{2(d-2)}\frac{4\pi^{d/2}}{\Gamma(d/2-1)}p^{2k}\log(p^2/\bar{\mu}^2). \tag{52}$$

3 The Holographic Dictionary

3.1 A First Look at the and Holographic Renormalization

In order to compute the above scalar two-point function holographically, we consider a self interacting scalar field in a fixed Euclidean background with the action

$$S = \int d^{d+1}x\sqrt{g}\left(\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + V(\phi)\right). \tag{53}$$

We will take the metric to be of the form

$$ds^2 = dr^2 + \gamma_{ij}(r,x)dx^i dx^j, \tag{54}$$

where $i, j = 1, 2, \ldots, d$ run over the field theory directions, and the induced metric on the constant r slices is given by

$$\gamma_{ij}(r,x) = e^{2A(r)}\hat{g}_{ij}(x), \tag{55}$$

with

$$A(r) = r, \quad \hat{g}_{ij}(x) = \delta_{ij}, \tag{56}$$

for AdS. This metric is diffeomorphic to the upper-half plane or Poincaré coordinates metric

$$ds^2 = \frac{dz_0^2 + d\vec{z}^2}{z_0^2}. \tag{57}$$

Our first task is to obtain the radial Hamiltonian for this model, interpreting the radial coordinate r as Hamiltonian 'time'. The action can be written in the form

$$S = \int^r dr' L = \int^r dr' d^d x \sqrt{\gamma}\left(\frac{1}{2}\dot{\phi}^2 + \frac{1}{2}\gamma^{ij}\partial_i\phi\partial_j\phi + V(\phi)\right). \tag{58}$$

The canonical momentum conjugate to ϕ then is

$$\pi = \frac{\delta L}{\delta\dot{\phi}} = \sqrt{\gamma}\dot{\phi}. \tag{59}$$

The HJ equation can be derived from the relation

$$\dot{\mathcal{S}} = L = \int d^d x \left(\dot{\phi}\frac{\delta\mathcal{S}}{\delta\phi} + \dot{\gamma}_{ij}\frac{\delta\mathcal{S}}{\delta\gamma_{ij}}\right), \tag{60}$$

where Hamilton's principal function (see Appendix "Hamilton–Jacobi primer"), $\mathcal{S}$, has no explicit r dependence since the Lagrangian is diffeomorphism covariant. Writing

$$\pi = \sqrt{\gamma}\dot{\phi} = \frac{\delta\mathcal{S}}{\delta\phi}, \tag{61}$$

this equation becomes

$$\int d^d x \left[\sqrt{\gamma}\left(\frac{1}{2}\left(\frac{1}{\sqrt{\gamma}}\frac{\delta\mathcal{S}}{\delta\phi}\right)^2 - \frac{1}{2}\gamma^{ij}\partial_i\phi\partial_j\phi - V(\phi)\right) + 2\dot{A}\gamma_{ij}\frac{\delta\mathcal{S}}{\delta\gamma_{ij}}\right] = 0. \tag{62}$$

This is the HJ equation for the scalar field in a fixed gravitational background, which can be rewritten in the more useful form

$$\sqrt{\gamma}\left(\frac{1}{2}\left(\frac{1}{\sqrt{\gamma}}\frac{\delta\mathcal{S}}{\delta\phi}\right)^2-\frac{1}{2}\gamma^{ij}\partial_i\phi\partial_j\phi-V(\phi)\right)+2\dot{A}\delta_\gamma\mathcal{L}=\partial_i v^i, \tag{63}$$

where

$$\mathcal{S}=\int d^dx\mathcal{L}, \tag{64}$$

and

$$\delta_\gamma=\int d^dx\gamma_{ij}\frac{\delta}{\delta\gamma_{ij}}. \tag{65}$$

The term $\partial_i v^i$ on the RHS is a total derivative that can be arbitrary, but which generically needs to be taken into account when trying to solve (63). It is not difficult to solve this equation iteratively, for example in a derivative expansion, for a general potential $V(\phi)$. However, for the present discussion it suffices to consider the simple—yet far from trivial—case of a free scalar field with the potential

$$V(\phi)=\frac{1}{2}m^2\phi^2. \tag{66}$$

The great simplification that results from this potential is that we can solve the corresponding HJ equation exactly, to all orders in transverse derivatives.

The HJ equation (63) in this case becomes

$$\sqrt{\gamma}\left(\frac{1}{2}\left(\frac{1}{\sqrt{\gamma}}\frac{\delta\mathcal{S}}{\delta\phi}\right)^2-\frac{1}{2}\gamma^{ij}\partial_i\phi\partial_j\phi-\frac{1}{2}m^2\phi^2\right)+2\delta_\gamma\mathcal{L}=\partial_i v^i. \tag{67}$$

Inserting an ansatz of the form

$$\mathcal{S}=\frac{1}{2}\int d^dx\sqrt{\gamma}\phi f(-\Box_\gamma)\phi, \tag{68}$$

we find that it solves the HJ equation, provided the function $f(x)$ satisfies [19]

$$f^2(x)+df(x)-m^2-x-2xf'(x)=0. \tag{69}$$

The general solution of this equation is

$$f(x)=-\frac{d}{2}-\frac{\sqrt{x}\left(K_k'(\sqrt{x})+cI_k'(\sqrt{x})\right)}{K_k(\sqrt{x})+cI_k(\sqrt{x})}, \tag{70}$$

where $k = \Delta - d/2 > 0$, c is an arbitrary constant, and $I_k(x)$ and $K_k(x)$ denote the modified Bessel function of the first and second kind respectively. Using the asymptotic behaviors as $x \to 0$

$$K_0(x) \sim -\log x, \quad K_k(x) \sim \frac{\Gamma(k)}{2}\left(\frac{x}{2}\right)^{-k}, \quad k > 0, \qquad I_k(x) \sim \frac{1}{\Gamma(k+1)}\left(\frac{x}{2}\right)^{k}, \tag{71}$$

we see that $K_k(x)$ dominates in $f(x)$ as $x \to 0$, unless $|c| \to \infty$. In particular, we find

$$f(x) \overset{x\to 0}{\sim} \begin{cases} -\frac{d}{2} + k = -(d-\Delta), & |c| < \infty, \\ -\frac{d}{2} - k = -\Delta, & |c| \to \infty. \end{cases} \tag{72}$$

Since,

$$\dot{\phi} = \frac{1}{\sqrt{\gamma}}\frac{\delta \mathcal{S}}{\delta \phi}, \tag{73}$$

we see that the two asymptotic solutions for $f(x)$ correspond to $\phi \sim e^{-(d-\Delta)r}$ and $\phi \sim e^{-\Delta r}$ respectively, which are precisely the asymptotic behaviors of the two linearly independent solutions of the equation of motion. The solution for $f(x)$ with $|c| < \infty$ corresponds to the asymptotically dominant mode. Hence, in order to make the variational problem well defined for generic solutions of the equation of motion we have no choice but demand that $|c| < \infty$.

Expanding the solution for $f(x)$ with $|c| < \infty$ for small x and taking k to be an integer we obtain,

$$\begin{aligned} f(x) = &-(d-\Delta) + \frac{x}{(2\Delta - d - 2)} - \frac{x^2}{(2\Delta - d - 2)(2\Delta - d - 4)} \\ &+ \cdots + \frac{(-1)^k}{2^{2k-1}\Gamma(k)^2} x^k \log x + \left(a(k) - \frac{c}{2^{2k-2}\Gamma(k)^2}\right) x^k + \cdots, \end{aligned} \tag{74}$$

where $a(k)$ is a known function of k, whose explicit form we will not need, and the dots denote asymptotically subleading terms. A number of comments are in order here. Firstly, this solution depends explicitly on the undetermined constant $|c| < \infty$. Secondly, this solution seems to lead to a non-local boundary term due to the logarithmic term. And finally, one may worry that higher terms in this asymptotic expansion need to be considered. Fortunately, all these issues can be addressed by noticing that the contribution of the last term to the boundary term is proportional to

$$\int d^d x \sqrt{\gamma} \phi (-\Box_\gamma)^k \phi, \tag{75}$$

which, taking into account the asymptotic behavior of the scalar and of the induced metric, can be easily seen to have a finite limit as $r \to \infty$. Such terms, therefore, correspond to adding finite local contributions to the boundary term S_b. We conclude that higher order terms in the asymptotic expansion of $f(x)$ need not be considered since they would give rise to a vanishing contribution to S_b in the limit $r \to \infty$. Moreover, the arbitrariness in the value of c is not a problem because different values of c lead to boundary terms S_b which differ by a finite local term. Any value of $|c| < \infty$, therefore, is equally acceptable since the corresponding boundary term makes the variational problem well defined. Finally, coming to the apparent non-locality of the boundary term we have deduced above, we notice that the logarithmic term can be written as

$$(-\Box_\gamma)^k \log(-\Box_\gamma) = (-\Box_\gamma)^k \left(\log(\mu^2 e^{-2r}) + \log(-\Box_\delta/\mu^2)\right), \tag{76}$$

where μ^2 is an arbitrary scale and $\Box_\delta = \partial_i \partial_i$ denotes the Laplacian in the flat transverse space. Crucially, the non-local part gives rise to a finite contribution in Hamilton's principal function and so it can be omitted from counterterms. The most general local boundary term that makes the variational problem well defined is therefore [12, 19]

$$S_{\text{ct}}[\gamma, \phi, r] = -\frac{1}{2}\int d^d x \sqrt{\gamma}\phi \Bigg(-(d-\Delta) + \frac{-\Box_\gamma}{(2\Delta - d - 2)} - \frac{(-\Box_\gamma)^2}{(2\Delta - d - 2)(2\Delta - d - 4)} + \cdots + \frac{(-1)^k}{2^{2k-1}\Gamma(k)^2}(-\Box_\gamma)^k \log(\mu^2 e^{-2r}) + \xi(-\Box_\gamma)^k \Bigg)\phi, \tag{77}$$

where we have allowed for a local finite boundary term with arbitrary coefficient ξ. Notice that although it is possible to find counterterms that remove the UV divergences and are also local in transverse derivatives, this is only at the cost of introducing explicit dependence in the radial coordinate, r. This is precisely the origin of the holographic conformal anomaly [4].

The renormalized action on the UV cut-off r_o is defined as

$$S_{\text{ren}} := S_{\text{reg}} + S_{\text{ct}}, \tag{78}$$

and it admits a finite limit, $\hat{S}_{\text{ren}}$, as the cut-off is removed:

$$\hat{S}_{\text{ren}} = \lim_{r_o \to \infty} S_{\text{ren}}. \tag{79}$$

In this case, ignoring the scheme dependent contact terms, we obtain

$$S_{\text{ren}} = \frac{(-1)^k}{2^{2k}\Gamma(k)^2}\int d^d x \phi_{(0)}(-\Box)^k \log(-\Box/\bar{\mu}^2)\phi_{(0)}. \tag{80}$$

The holographic dictionary identifies S_{ren} with the renormalized generating function of connected correlators, $W_{\text{ren}}[J]$, and $\phi_{(0)}$ with the source J. We therefore deduce that the renormalized two-point function of the dual scalar operator takes the form

$$\langle \mathcal{O}_\Delta(p)\mathcal{O}_\Delta(-p)\rangle_{\text{ren}} = \frac{(-1)^{k+1}}{2^{2k-1}\Gamma(k)^2} p^{2k}\log(p^2/\bar{\mu}^2), \tag{81}$$

which agrees with the CFT calculation in (52). Comparing the coefficients, we determine

$$c(g,\Delta) = \frac{2k\Gamma(d/2+k)}{\pi^{d/2}\Gamma(k)}, \tag{82}$$

which turns out to be precisely the correct coefficient consistent with the Ward identities.

3.2 The Holographic Dictionary in Hamiltonian Language

The local RG description of QFTs that we discussed above allows us to formulate the holographic dictionary in a more precise language, identifying all quantities in the bulk theory with QFT quantities. In particular, we identify the following objects on the two sides of the gauge/gravity duality.

Radial coordinate	$r \leftrightarrow \tau = \log\mu$	RG "time"
Induced fields	$\phi \leftrightarrow J$	Running local couplings (sources)
Regularized action	$S_{\text{reg}}[\phi] \leftrightarrow W[J]$	Generating function
Renormalized action	$S_{\text{ren}}[\phi] \leftrightarrow W_{\text{ren}}[J]$	Renormalized generating function
Radial Hamiltonian	$H \leftrightarrow \mathbb{H}$	RG Hamiltonian
Radial momenta	$\pi_\phi \leftrightarrow \langle\mathcal{O}\rangle$	Running local operators
Non-normalizable modes	$\phi_{(0)} \leftrightarrow J_R\|_\infty$	Renormalized couplings at ∞
Renormalized momenta	$\hat{\pi}_{(\Delta)} \leftrightarrow \langle\mathcal{O}\rangle\|_\infty$	Bare operators

The above Table should serve as a guide in order to interpret all calculations in the bulk theory that we are going to describe in the next sections.

4 Radial Hamiltonian Formulation of Gravity Theories

The holographic dictionary consists in a precise map between observables on the two sides of the duality. From the point of view of the bulk gravitational theory, the physical observables correspond to the symplectic space of asymptotic data, which is the key to formulating a well posed variational problem [17]. As we will now review, a general systematic construction of the symplectic space of asymptotic data proceeds by formulating the bulk dynamics in Hamiltonian language, with the radial coordinate identified with the Hamiltonian "time". As we saw in the previous section, this formulation of the bulk dynamics parallels the real space renormalization group of the dual QFT.

For concreteness, let us consider Einstein–Hilbert gravity in a $d+1$-dimensional non-compact manifold $\mathcal{M}$ coupled to a scalar field described by the action

$$S=-\frac{1}{2\kappa^2}\left(\int_{\mathcal{M}} d^{d+1}x\sqrt{g}\left(R[g]-\frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi-V(\varphi)\right)+\int_{\partial\mathcal{M}} d^dx\sqrt{\gamma}2K\right). \tag{83}$$

Here, $\kappa^2=8\pi G_{d+1}$ is the gravitational constant in $d+1$ dimensions and the boundary term is the standard Gibbons–Hawking term for Einstein–Hilbert gravity [20], which, as we shall see, is required in order to formulate the dynamics in a Hamiltonian language.[6] Moreover, throughout these lectures we will work in Euclidean signature, but the entire analysis can be straightforwardly adapted to Lorentzian signature.

The radial Hamiltonian formulation of the bulk dynamics starts with picking a radial coordinate r such that $r\to\infty$ corresponds to the location of the boundary $\partial\mathcal{M}$ of $\mathcal{M}$. This radial coordinate need not be a Gaussian normal coordinate, nor should it be a good coordinate throughout $\mathcal{M}$. Instead, r need only cover an open chart $\mathcal{M}_\varepsilon$ in the vicinity of $\partial\mathcal{M}$ in $\mathcal{M}$. Moreover, if $\partial\mathcal{M}$ consists of multiple disconnected components then a different radial coordinate must be used in the vicinity of each boundary component and different Hamiltonian descriptions must be applied to describe the various asymptotic regimes, as is illustrated in Fig. 1.

Having picked a radial coordinate r emanating from (a component of) the boundary $\mathcal{M}$, the radial Hamiltonian formulation of the dynamics proceeds as in the standard ADM formalism [22], except that the Hamiltonian "time" r is a spacelike coordinate instead of a timelike one. All tensor fields are decomposed in components along and transverse to the radial coordinate r. In particular, the metric is parameterized in terms of the lapse function N, the shift vector N_i, and the induced metric γ_{ij} on the hypersurfaces Σ_r of constant radial coordinate r as

$$ds^2=(N^2+N_iN^i)dr^2+2N_idrdx^i+\gamma_{ij}dx^idx^j, \tag{84}$$

where $i,j=1,\ldots,d$. The metric $g_{\mu\nu}$ is therefore replaced in the Hamiltonian description by the three fields $\{N,N_i,\gamma_{ij}\}$ on Σ_r. Moreover, the curvature tensors of the metric $g_{\mu\nu}$ can be expressed in terms of the (intrinsic) curvature tensors of the hypersurfaces Σ_r and the extrinsic curvature, K_{ij}, describing the embedding of $\Sigma_r\hookrightarrow\mathcal{M}$. The latter is defined as

$$K_{ij}=\frac{1}{2}(\mathcal{L}_ng)_{ij}=\frac{1}{2N}\left(\dot\gamma_{ij}-D_iN_j-D_jN_i\right), \tag{85}$$

[6] We emphasize that, contrary to what is often claimed, the Gibbons–Hawking term *does not* render the variational problem well posed in a non-compact manifold. It does so in a *compact* space, but in a non-compact manifold additional boundary terms are required [17, 21].

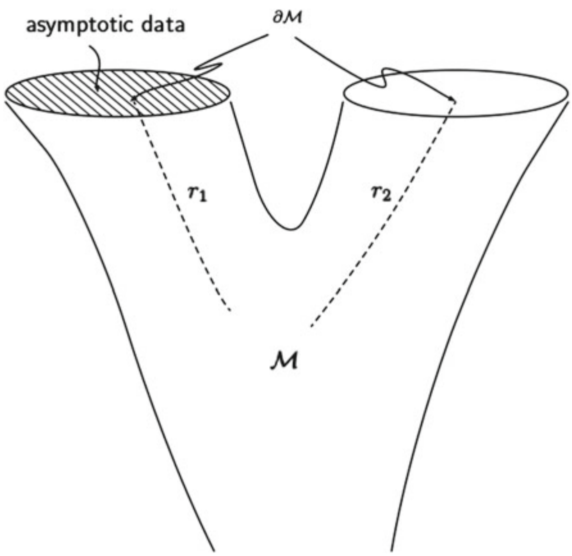


Fig. 1 A non-compact manifold $\mathcal{M}$ with a boundary $\partial\mathcal{M}$ consisting of two disconnected components. The Hamiltonian formulation of the bulk dynamics in the vicinity of the two disconnected components must be done separately, using two different radial coordinates, r_1 and r_2, emanating respectively from each disconnected component of the boundary. The Hamiltonian analysis need only be applicable in an open neighborhood of each boundary component, which is sufficient in order to construct the symplectic space of asymptotic data on each component, as well as the appropriate boundary terms required to render the variational problem well posed

where the dot $\dot{}$ denotes a derivative w.r.t. the radial coordinate r, D_i denotes the covariant derivative w.r.t. the induced metric γ_{ij}, and the unit normal to Σ_r, n^μ, is given by $n^\mu = \left(1/N, -N^i/N\right)$. Using the expressions for the inverse metric and the Christoffel symbols given in Appendix "ADM identities" one finds that the Ricci scalar takes the form

$$R[g] = R[\gamma] + K^2 - K_{ij}K^{ij} + \nabla_\mu \zeta^\mu, \tag{86}$$

where $R[\gamma]$ is the Ricci scalar of the induced metric γ_{ij}, $K = \gamma^{ij}K_{ij}$ denotes the trace of the extrinsic curvature, and $\zeta^\mu = -2Kn^\mu + 2n^\rho\nabla_\rho n^\mu$. From the identities in Appendix "ADM identities" follows that $\zeta^r = -2K/N$ and, hence, the Gibbons–Hawking term in (83) precisely cancels the total derivative term in Ricci curvature (86). This allows us to write the action as an integral over a radial Lagrangian as

$$S = \int dr L, \tag{87}$$

where

$$L = -\frac{1}{2\kappa^2}\int_{\Sigma_r} d^d x\sqrt{\gamma}N\left(R[\gamma]+K^2-K^i_jK^j_i-\frac{1}{2N^2}\left(\dot{\varphi}-N^i\partial_i\varphi\right)^2-\frac{1}{2}\gamma^{ij}\partial_i\varphi\partial_j\varphi-V(\varphi)\right). \tag{88}$$

Note that, as we anticipated earlier, the Gibbons–Hawking term is required for the radial Hamiltonian formulation of the bulk dynamics. This observation can be utilized in order to derive the correct Gibbons–Hawking term for general bulk Lagrangians, such as, for example, that describing a scalar field *conformally* coupled to Einstein–Hilbert gravity [23].

From the radial Lagrangian (88) we read-off the canonical momenta conjugate to the induced metric γ_{ij} and the scalar φ

$$\pi^{ij} = \frac{\delta L}{\delta\dot{\gamma}_{ij}} = -\frac{1}{2\kappa^2}\sqrt{\gamma}(K\gamma^{ij}-K^{ij}), \tag{89a}$$

$$\pi_\varphi = \frac{\delta L}{\delta\dot{\varphi}} = \frac{1}{2\kappa^2}\sqrt{\gamma}\; N^{-1}\left(\dot{\varphi}-N^i\partial_i\varphi\right). \tag{89b}$$

However, the Lagrangian (88) does not depend on the radial derivatives (generalized velocities), $\dot{N}$ and $\dot{N}_i$, of the shift function and lapse vector and so their conjugate momenta vanish identically. This means that the lapse function and the shift vector are not dynamical fields, but rather Lagrange multipliers, whose equations of motion lead to *constraints*. The separation of variables into dynamical fields and Lagrange multipliers through the ADM decomposition (84) is one of the main advantages of the Hamiltonian formulation of the bulk dynamics.

The Legendre transform of the Lagrangian (88) gives the Hamiltonian

$$H = \int_{\Sigma_r} d^d x\left(\pi^{ij}\dot{\gamma}_{ij}+\pi_\varphi\dot{\varphi}\right)-L = \int_{\Sigma_r} d^d x\left(N\mathcal{H}+N_i\mathcal{H}^i\right), \tag{90}$$

where

$$\mathcal{H} = 2\kappa^2\gamma^{-\frac{1}{2}}\left(\pi^i_j\pi^j_i-\frac{1}{d-1}\pi^2+\frac{1}{2}\pi^2_\varphi\right)+\frac{1}{2\kappa^2}\sqrt{\gamma}\left(R[\gamma]-\frac{1}{2}\partial_i\varphi\partial^i\varphi-V(\varphi)\right), \tag{91a}$$

$$\mathcal{H}^i = -2D_j\pi^{ij}+\pi_\varphi\partial^i\varphi. \tag{91b}$$

It follows that Hamilton's equations for the Lagrange multipliers N and N_i impose the constraints

$$\mathcal{H} = \mathcal{H}^i = 0, \tag{92}$$

and, hence, the Hamiltonian vanishes identically on the constraint surface. This is a direct consequence of the diffeomorphism invariance of the bulk theory [24]. In particular, the constraints $\mathcal{H} = 0$ and $\mathcal{H}^i = 0$ are first class constraints that, through the Poisson bracket, generate diffeomorphisms along the radial direction and along Σ_r, respectively.

4.1 Hamilton–Jacobi Formalism

From the expressions (90) and (91) we observe that the Hamiltonian does not depend explicitly on the radial coordinate r, but only through the induced fields on S_r. This is a consequence of the diffeomorphism invariance of the action (83) and it implies that the HJ equation takes the form

$$H = 0, \tag{93}$$

which is equivalent to the two constraints (92), where the canonical momenta are expressed as gradients of Hamilton's principal function $\mathcal{S}$ (see Appendix "Hamilton–Jacobi Primer")

$$\pi^{ij} = \frac{\delta \mathcal{S}}{\delta \gamma_{ij}}, \quad \pi_\varphi = \frac{\delta \mathcal{S}}{\delta \varphi}. \tag{94}$$

This form of the canonical momenta turns the constraints (92) into functional partial differential equations for $\mathcal{S}$. The momentum constraint, $\mathcal{H}^i = 0$, implies that $\mathcal{S}[\gamma, \varphi]$ is invariant with respect to diffeomorphims on the radial slice Σ_r. The Hamiltonian constraint, $\mathcal{H} = 0$, takes the form

$$\frac{2\kappa^2}{\sqrt{\gamma}}\left(\left(\gamma_{ik}\gamma_{jl} - \frac{1}{d-1}\gamma_{ij}\gamma_{kl}\right)\frac{\delta \mathcal{S}}{\delta \gamma_{ij}}\frac{\delta \mathcal{S}}{\delta \gamma_{kl}} + \frac{1}{2}\left(\frac{\delta \mathcal{S}}{\delta \varphi}\right)^2\right) + \frac{\sqrt{\gamma}}{2\kappa^2}\left(R[\gamma] - \frac{1}{2}\partial_i\varphi\partial^i\varphi - V\right) = 0, \tag{95}$$

and dictates the radial evolution of the induced fields on Σ_r.

As is reviewed in Appendix "Hamilton–Jacobi primer", a solution $\mathcal{S}[\gamma, \varphi]$ of the HJ equation leads to a solution of Hamilton's equations, and hence of the second order equations of motion. In particular, given a solution $\mathcal{S}[\gamma, \varphi]$ of the HJ equation, equating the expressions (89) and (94) for the canonical momenta (this corresponds to the first of Hamilton's equations) leads to the first order flow equations

$$\dot{\gamma}_{ij} = 4\kappa^2\left(\gamma_{ik}\gamma_{jl} - \frac{1}{d-1}\gamma_{kl}\gamma_{ij}\right)\frac{1}{\sqrt{\gamma}}\frac{\delta \mathcal{S}}{\delta \gamma_{kl}}, \tag{96a}$$

$$\dot{\varphi} = \frac{2\kappa^2}{\sqrt{\gamma}}\frac{\delta \mathcal{S}}{\delta \varphi}. \tag{96b}$$

Integrating these first order equations one obtains the corresponding solution of the second order equations of motion. Crucially, to determine the most general solution of the equations of motion one need *not* find the most general solution of the HJ equation. The HJ equation is a (functional) partial differential equation and so its general solution contains arbitrary integration functions of the induced fields. However, the general solution of the equations of motion is parameterized by $2n$ integration constants,[7] where n is the number of generalized coordinates, i.e. of induced fields on Σ_r. The general solution of the equations of motion, therefore, can be obtained from a *complete integral* of the HJ equation, which is a principal function $\mathcal{S}[\gamma, \varphi]$ containing n integration *constants* (functions of the transverse coordinates only) [24]. Another n integration constants are obtained by integrating the first order equations (96), which leads to a solution of the equations of motion with $2n$ integration constants, i.e. the general solution.

Another important aspect of HJ theory reviewed in Appendix "Hamilton–Jacobi primer" is that the regularized action, defined as the on-shell action evaluated with the radial cut-off Σ_r, i.e.

$$S_{\text{reg}}[\gamma(r,x), \varphi(r,x)] = \int^{r} dr'\, L|_{\text{on-shell}}\,, \tag{97}$$

is naturally a functional of the induced fields γ_{ij} and φ on Σ_r and satisfies the HJ equation (95). If the regularized action is evaluated on the general solution of the equations of motion, then S_{reg} contains n integration constants and so it corresponds to a complete integral of the HJ equation. If, however, S_{reg} is evaluated on solutions of the equations of motion that satisfy certain conditions in the deep interior of $\mathcal{M}$, such as regularity conditions, then it will generically contain less than n integration constants and so it will not correspond to a complete integral of the HJ equation.

Recapitulating the last two paragraphs, we have seen that the $2n$ integration constants parameterizing the general solution of the equations of motion are divided into two distinct sets of integration constants in the HJ formalism: n integration constants parameterize a complete integral of the HJ equation, while the remaining n arise as integration constants of the first order equations (96). As we shall see later, the integration constants parameterizing a complete integral of the HJ equation correspond generically to the *normalizable* modes of the asymptotic solutions of the equations of motion, while the integration constants coming from the flow equations correspond to the *non-normalizable modes*.[8] Moreover, we have argued that the regularized action (97), evaluated on the general solution of the equations of motion gives rise to a complete integral of the HJ equation. Combining these two

[7] In order to distinguish them from arbitrary integration functions of the HJ partial differential equation, we refer to arbitrary functions of the transverse coordinates arising from the integration of the radial equations of motion as "integration constants".

[8] Since under certain conditions both modes can be normalizable, more generally the distinction is between asymptotically subleading and dominant modes, respectively.

facts leads to an observation that is fundamental to holographic renormalization and its relation to HJ theory. In order for a theory to be (holographically) renormalizable, the near-boundary divergences of the regularized action (97) must be the same for all solutions of the equations of motion and should not depend on the details of the solutions in the deep interior of $\mathcal{M}$. This means that the near-boundary divergences of any complete integral of the HJ equation must be the same, and hence independent of the n integration constants parameterizing a complete integral of the HJ equation. We therefore arrive at the following definition:

Definition (*Holographic Renormalizability*)
A gravity theory in a non-compact manifold that admits a radial Hamiltonian description is holographically renormalizable if:

(i) The near boundary divergences of any complete integral of the radial HJ equation are the same, so the difference between any two complete integrals is free of divergences.
(ii) The common divergent terms of all complete integrals are local functionals of the induced fields on the radial cut-off S_r, i.e. analytic functions of the induced fields and polynomial in transverse derivatives.

The first of these conditions is equivalent with the existence of a well defined symplectic space of asymptotic solutions of the equations of motion and it is required in order to render the variational problem in $\mathcal{M}$ well posed [17, 21]. The second condition, however, is necessary only due to the holographic interpretation of the near boundary divergences of the regularized action as the UV divergences of the generating functional of a *local* quantum field theory. As is discussed in [17], a free scalar field in $\mathbb{R}^{d+1}$ is an example of a system that satisfies condition (i), but not (ii). In cases when condition (i) is not met, there are two possibilities for making progress. One option is to treat the mode(s) that causes condition (i) to be violated perturbatively, and proceed as one would in conformal perturbation theory in the presence of an irrelevant operator. This approach was discussed in general in [25, 26] and explicit examples can be found in [27–32]. Such an analysis is often sufficient, but it is also possible to treat the modes that violate condition (i) non-perturbatively. This requires constructing a well-defined symplectic space of asymptotic solutions of the equations of motion and generically involves some rearrangement of the bulk degrees of freedom, such as a Kaluza–Klein reduction. This approach, which is discussed in [17], is the holographic dual of following the RG flow in the presence of the irrelevant operator in reverse until a new UV “fixed point” is found. The new “fixed point” in this case is defined in terms of the symplectic space of asymptotic solutions of the bulk equation of motion, and almost in all cases it involves asymptotically non-AdS backgrounds.

Assuming that both conditions of definition of holographic renormalizability hold, as we will assume from now on, the UV divergences of any complete integral of the HJ equation, and hence of the regularized action, can be removed by adding the negative of the divergent part of any solution of the HJ equation as a boundary term in the original action (83). Namely, we define the counterterms as

$$S_{\text{ct}} = -\mathcal{S}_{\text{local}}, \tag{98}$$

where $\mathcal{S}_{\text{local}}$ is the divergent part of any complete integral of the HJ equation, which, by condition (ii) of the above definition, is a local functional of the induced fields on the radial slice Σ_r. In the next section we will give a precise definition of $\mathcal{S}_{\text{local}}$, and discuss procedures for systematically determining these terms by solving the HJ equation. Before we turn to the systematic construction of $\mathcal{S}_{\text{local}}$, however, we should emphasize one last important point. Although the local and divergent part of the HJ solution is unique, the above discussion suggests that it is possible to add further *finite* and local boundary terms to the bulk action (83), corresponding to (a very special choice of) the integration constants of a complete integral of the HJ equation. More generally, therefore, the counterterms will be defined as

$$S_{\text{ct}} = -\left(\mathcal{S}_{\text{local}} + \mathcal{S}_{\text{scheme}}\right), \tag{99}$$

where $\mathcal{S}_{\text{scheme}}$ denotes these extra finite terms, which we will discuss in more detail in the next sections. These terms, an example of which is the term proportional to ξ in (77), do not cancel divergences, but they correspond to choosing a renormalization scheme [8]. Once the local counterterms, $\mathcal{S}_{\text{ct}}$, have been determined, the renormalized action on the radial cut-off is given by

$$S_{\text{ren}} := S_{\text{reg}} + S_{\text{ct}} = \int d^d x \left(\gamma_{ij}\Pi^{ij} + \varphi\Pi_\varphi\right), \tag{100}$$

where the renormalized canonical momenta Π^{ij} and Π_φ are arbitrary functions that correspond to the integration constants parameterizing an asymptotic complete integral of the HJ equation. As we shall see explicitly later, the holographic dictionary relates Π^{ij} and Π_φ with the renormalized one-point functions of the dual operators.

5 Recursive Solution of the Hamilton–Jacobi Equation

The main task in carrying out the procedure of holographic renormalization is determining the local functional $\mathcal{S}_{\text{local}}$, as well as the asymptotic expansions for the induced fields on Σ_r. There is a number of methods to obtain these, differing in generality and efficiency. The approach of [4, 8–10] does not rely on the HJ equation and its first objective is to obtain the asymptotic expansions for the induced fields by solving asymptotically the second order equations of motion. Evaluating the regularized action on these asymptotic solutions and then inverting the asymptotic expansions in order to express the result in terms of induced fields on the cut-off Σ_r leads to an explicit expression for $\mathcal{S}_{\text{local}}$. This method is general but it is unnecessarily complicated. In particular, as we shall see, it is much more efficient to first obtain $\mathcal{S}_{\text{local}}$ by solving the HJ equation, and only then derive the asymptotic expansions of the induced fields by integrating the first order equations (96), instead of the second order equations. Moreover, deriving the holographic Ward identities is much simpler in the radial Hamiltonian language since they follow directly from the first class constraints (92).

The method of [6, 11] does use the HJ equation to obtain $\mathcal{S}_{\text{local}}$, but it does so by postulating an ansatz consisting of all possible local and covariant terms that can potentially contribute to the UV divergences with arbitrary coefficients. Inserting this ansatz in the HJ equation leads to equations for the coefficients that can be solved to determine $\mathcal{S}_{\text{local}}$. For simple cases this approach is practical since the possible terms in $\mathcal{S}_{\text{local}}$ can be easily guessed. However, this method becomes impractical for more complicated systems where $\mathcal{S}_{\text{local}}$ contains more than a couple of terms, or when it is not easy to guess all terms (e.g. for asymptotically Lifshitz backgrounds). In particular, if n is the number of independent terms in the ansatz for $\mathcal{S}_{\text{local}}$, the number of equations for the arbitrary coefficients in the ansatz one obtains from the HJ equation is generically of order $n(n+1)/2$, which grows much larger than n very fast. The system of equations determining the coefficients in the ansatz is therefore overdetermined, but all equations need to be checked to ensure that the solution is consistent.

A systematic algorithm for solving the HJ equation recursively, without relying on an ansatz, was developed in [13]. This method is based on a formal expansion of the principal function $\mathcal{S}$ in eigenfunctions of the dilatation operator of the dual theory at the UV, and can be applied to any background that possesses some kind of asymptotic scaling symmetry. Besides asymptotically locally AdS backgrounds, this includes backgrounds with non-relativistic Lifshitz symmetry [33, 34]. This method was generalized to relativistic backgrounds that do not necessarily possess an asymptotic scaling symmetry in [35], while a further generalization to include non-relativistic backgrounds was carried out in [31, 36]. This latter generalization involves an expansion of $\mathcal{S}$ in simultaneous eigenfunctions of *two* commuting operators. However, here we will focus on the simpler cases discussed in [13, 35], which involve an expansion in eigenfunctions of a single operator.

The initial steps in the recursive algorithms of [13, 35] are common, and they just rely on the fact that we seek a solution $\mathcal{S}$ of the HJ equation in the form of a covariant expansion in eigenfunctions of a—yet unspecified—functional operator $\mathfrak{F}$. Namely, we formally write

$$\mathcal{S} = \mathcal{S}_{(\alpha_0)} + \mathcal{S}_{(\alpha_1)} + \mathcal{S}_{(\alpha_2)} + \cdots , \tag{101}$$

where each term is an eigenfunction of $\mathfrak{F}$, i.e.

$$\mathfrak{F}\mathcal{S}_{(\alpha_k)} = \lambda_k \mathcal{S}_{(\alpha_k)}, \tag{102}$$

with an eigenvalue λ_k. α_k denotes a convenient label that counts the order of the expansion. In order to obtain a recursive algorithm for determining $\mathcal{S}_{(\alpha_k)}$ it is necessary to also introduce a density $\mathcal{L}$ such that

$$\mathcal{S} = \int_{\Sigma_r} d^d x \mathcal{L}[\gamma, \varphi]. \tag{103}$$

We then have

$$\mathcal{L} = \mathcal{L}_{(\alpha_0)} + \mathcal{L}_{(\alpha_1)} + \mathcal{L}_{(\alpha_2)} + \cdots , \tag{104}$$

where $\mathcal{S}_{(\alpha_k)} = \int_{\Sigma_r} d^d x \mathcal{L}_{(\alpha_k)}$. Note that the densities $\mathcal{L}_{(\alpha_k)}$ are only defined up to total derivative terms and they are not necessarily eigenfunctions of the operator $\hat{\delta}$. They are eigenfunctions up to total derivatives.

An important identity that is crucial in the construction of the recursion algorithm follows from the expressions (94) for the canonical momenta. Namely, for arbitrary variations we have

$$\pi^{ij}\delta\gamma_{ij} + \pi_\varphi \delta\varphi = \delta\mathcal{L} + \partial_i v^i(\delta\gamma, \delta\varphi), \tag{105}$$

for some vector field $v^i(\delta\gamma, \delta\varphi)$. Specializing this to the operator $\hat{\delta}$ gives

$$\pi^{ij}_{(\alpha_k)}\hat{\delta}\gamma_{ij} + \pi_{\varphi(\alpha_k)}\hat{\delta}\varphi = \hat{\delta}\mathcal{L}_{(\alpha_k)} + \partial_i v^i_{(\alpha_k)}(\hat{\delta}\gamma, \hat{\delta}\varphi) = \lambda_k \mathcal{L}_{(\alpha_k)} + \partial_i \widetilde{v}^i_{(\alpha_k)}(\hat{\delta}\gamma, \hat{\delta}\varphi), \tag{106}$$

where

$$\pi^{ij}_{(\alpha_k)} = \frac{\delta\mathcal{S}_{(\alpha_k)}}{\delta\gamma_{ij}}, \quad \pi_{\varphi(\alpha_k)} = \frac{\delta\mathcal{S}_{(\alpha_k)}}{\delta\varphi}, \tag{107}$$

and $\widetilde{v}^i_{(\alpha_k)}$ is a vector field, generically different from $v^i_{(\alpha_k)}$ due to the fact that the action of $\hat{\delta}$ on $\mathcal{L}_{(\alpha_k)}$ may involve a total derivative. Since $\mathcal{L}$ is defined only up to a total derivative, however, without loss of generality we can choose the total derivatives in such a way so that

$$\boxed{\pi^{ij}_{(\alpha_k)}\hat{\delta}\gamma_{ij} + \pi_{\varphi(\alpha_k)}\hat{\delta}\varphi = \lambda_k \mathcal{L}_{(\alpha_k)}.} \tag{108}$$

This identity will be crucial in the construction of the recursion algorithm.

5.1 *The Induced Metric Expansion*

To proceed with the recursion algorithm we need to pick a suitable operator $\hat{\delta}$. The choice of such an operator is not unique, but it has to satisfy certain consistency criteria. Here we will discuss two specific choices. The first one is the operator

$$\boxed{\delta_\gamma = \int 2\gamma_{ij} \frac{\delta}{\delta\gamma_{ij}},} \tag{109}$$

which was introduced in [35]. The covariant expansion in eigenfunctions of this operator treats the scalar field non-perturbatively. In particular the resulting asymptotic solution of the HJ equation is expressed in terms of a generic scalar potential $V(\varphi)$, without the need to explicitly specify $V(\varphi)$. As a result, this expansion of the solution of principal function $\mathcal{S}$ is valid even for (relativistic) asymptotically non-AdS backgrounds, such as non-conformal branes [37].

It is easy to see that the covariant expansion in eigenfunctions of the operator (109) is a derivative expansion.[9] Choosing the label $\alpha_k = 2k$ to count derivatives, the corresponding eigenvalue is $\lambda_k = d - 2k$, where d is the contribution of the volume element. The zero order solution, therefore takes the form

$$\mathcal{S}_{(0)} = \frac{1}{\kappa^2} \int_{\Sigma_r} d^d x \sqrt{\gamma} U(\varphi), \tag{110}$$

for some "superpotential" $U(\varphi)$. Inserting this ansatz into the Hamiltonian constraint we find that $U(\varphi)$ satisfies the equation

$$\boxed{2(U')^2 - \frac{d}{d-1} U^2 - V(\varphi) = 0.} \tag{111}$$

As for the full HJ equation, we only need to obtain an asymptotic solution of this equation, around the value of φ near the boundary. As we have emphasized already, the recursive algorithm for solving the HJ equation we are describing here applies equally to asymptotically AdS and non-AdS backgrounds. The form of the scalar potential, therefore, is largely unrestricted, and we will keep both $V(\varphi)$ and $U(\varphi)$ general in the subsequent discussion. However, before we proceed it is instructive to have a closer look at the explicit form of $V(\varphi)$ and $U(\varphi)$ in the case of asymptotically AdS backgrounds.

In order for the theory (83) to admit an AdS solution, corresponding to $\varphi = 0$, the scalar potential must admit a Taylor expansion of the form

$$V(\varphi) = -\frac{d(d-1)}{\ell^2} + \frac{1}{2} m^2 \varphi^2 + \cdots, \tag{112}$$

where ℓ is the AdS radius of curvature and the scalar mass must satisfy the Breitenlohner–Freedman (BF) bound [38]

$$m^2 \ell^2 \geq -(d/2)^2, \tag{113}$$

[9] For the gravity-scalar system the expansion in eigenfunctions of (109) is indeed a derivative expansion. However, in general this is not the case. A counterexample is a Maxwell field.

in order for the AdS vacuum to be stable with respect to scalar perturbations. Moreover, the mass is related to the dimension Δ of the dual operator through the quadratic equation

$$m^2\ell^2 = -\Delta(d-\Delta). \tag{114}$$

Seeking a solution of (111) in the form of a Taylor expansion in φ, one finds two distinct solutions of the form[10]

$$U(\varphi) = -\frac{d-1}{\ell} - \frac{1}{4\ell}\mu\varphi^2 + \cdots, \tag{115}$$

where μ takes the two possible values Δ or $d-\Delta$. However, only a solution of the form

$$U(\varphi) = -\frac{d-1}{\ell} - \frac{1}{4\ell}(d-\Delta)\varphi^2 + \cdots, \tag{116}$$

can be used as a counterterm since only this solution removes the divergences from all possible solutions involving a non-trivial scalar [39].

Given the superpotential $U(\varphi)$ that determines the zero order solution in the covariant expansion of the HJ equation, we insert the formal expansion in eigenfunctions of the operator (109) in the HJ equation and match terms of equal eigenvalue using the identity (108), which leads to the linear recursion equations

$$\boxed{2U'(\varphi)\frac{\delta}{\delta\varphi}\int d^dx\mathcal{L}_{(2n)} - \left(\frac{d-2n}{d-1}\right)U(\varphi)\mathcal{L}_{(2n)} = \mathcal{R}_{(2n)}, \quad n>0,} \tag{117}$$

where

$$\mathcal{R}_{(2)} = -\frac{\sqrt{\gamma}}{2\kappa^2}\left(R[\gamma] - \frac{1}{2}\partial_i\varphi\partial^i\varphi\right), \tag{118}$$

$$\mathcal{R}_{(2n)} = -\frac{2\kappa^2}{\sqrt{\gamma}}\sum_{m=1}^{n-1}\left(\pi_{(2m)}{}^i_j\pi_{(2(n-m))}{}^j_i - \frac{1}{d-1}\pi_{(2m)}\pi_{(2(n-m))} + \frac{1}{2}\pi_{\varphi(2m)}\pi_{\varphi(2(n-m))}\right), \quad n>1.$$

Note that if $U'(\varphi)=0$, i.e. $U(\varphi)$ is a constant, then these recursion equations become algebraic. When $U'(\varphi)\neq 0$, these equations are first order linear inhomogeneous functional differential equations. The general solution, therefore, is the sum of the

[10]The overall sign of U is determined by requiring that the first order equations (96) imply the correct leading asymptotic behavior for the scalar, namely $\varphi \sim e^{-(d-\Delta)r}$. Moreover, when the scalar mass saturates the BF bound, one of the two asymptotic solutions for $U(\varphi)$ contains logarithms. We refer to [39] for the explicit form of the function $U(\varphi)$ in that case.

homogeneous solution and a unique inhomogeneous solution. The homogeneous solution takes the form

$$\mathcal{L}^{hom}_{(2n)} = \mathcal{F}_{(2n)}[\gamma] \exp\left(\frac{1}{2}\left(\frac{d-2n}{d-1}\right)\int^{\varphi}\frac{d\bar{\varphi}}{U'(\bar{\varphi})}U(\bar{\varphi})\right), \tag{119}$$

where $\mathcal{F}_{(2n)}[\gamma]$ is a local covariant functional of the induced metric of weight $d-2n$. It can be easily shown that these homogeneous solutions contribute only to the finite part of the on-shell action, and so we are not interested in them [35]. We are, therefore, only interested in the *inhomogeneous* solution of (117), which formally takes the form

$$\boxed{\mathcal{L}_{(2n)} = \frac{1}{2}e^{-(d-2n)A(\varphi)}\int^{\varphi}\frac{d\bar{\varphi}}{U'(\bar{\varphi})}e^{(d-2n)A(\bar{\varphi})}\mathcal{R}_{(2n)}(\bar{\varphi}),} \tag{120}$$

where

$$A = -\frac{1}{2(d-1)}\int^{\varphi}\frac{d\bar{\varphi}}{U'(\bar{\varphi})}U(\bar{\varphi}). \tag{121}$$

If $\mathcal{R}_{(2n)}$ does not involve derivatives of the scalar field with respect to the transverse coordinates, then evaluating the integral (120) is straightforward since it reduces to an ordinary integral. When $\mathcal{R}_{(2n)}$ does contain derivatives of the scalar field, however, some care is required in evaluating this integral. Table 1 in [35] provides general integration identities for up to and including four transverse derivatives, in all possible tensor combinations. This allows one to determine $\mathcal{L}_{(2n)}$ for $n \leq 2$, which suffices for $d \leq 4$.

The recursive procedure to successively determine $\mathcal{L}_{(2n)}$ proceeds as follows. For $n = 1$, $\mathcal{R}_{(2)}$ is given explicitly in (118) and so $\mathcal{L}_{(2)}$ can be immediately obtained from (120). The result is given in Table 2 of [35]. Having obtained the solution for $\mathcal{L}_{(2)}$, the relations (107) give the corresponding canonical momenta, which allow one to evaluate the next $\mathcal{R}_{(2n)}$ using (118). Inserting this back in (120) and performing the integral gives the next order solution for $\mathcal{L}_{(2n)}$. For $n = 2$ the general result is given in Table 3 of [35].

The order at which the recursive procedure stops depends on the leading asymptotic behavior of the fields. For asymptotically locally AdS backgrounds the recursion stops at order $n = [d/2]$, i.e. the integer part of $d/2$, since higher order terms are UV finite and arbitrary integration constants, parameterizing a complete integral of the HJ equation, enter in the solution. In that case, therefore, the counterterms are defined as

$$S_{\text{ct}} := -\sum_{n=0}^{[d/2]}\mathcal{S}_{(2n)}. \tag{122}$$

For even d, the last term in this sum gives rise to explicit cut-off dependence through a logarithmic divergence. The way this arises in this approach is as follows. The recursive procedure described above must be done keeping d as an arbitrary parameter. Denoting by $2k$ the final value of d, the recursion is carried out up to order $n = k$, where one finds that the solution $\mathcal{L}_{(2k)}$ contains a factor of $1/(d-2k)$, which is singular when we set d to its integer value $2k$. This singularity is then removed by the replacement

$$\frac{1}{d-2k} \to r_o, \tag{123}$$

where r_o is the radial cut-off [13, 35]. After this replacement one sets $d = 2k$ in the counterterms, which now contain a term which explicitly depends on r_o. This term is identified with the holographic conformal anomaly [4].

5.2 Dilatation Operator Expansion

We next turn to the covariant expansion developed in [39], which is an expansion in eigenfunctions of the dilatation operator

$$\boxed{\delta_D = \int d^d x \left(2\gamma_{ij}\frac{\delta}{\delta\gamma_{ij}} + (\Delta - d)\varphi\frac{\delta}{\delta\varphi}\right),} \tag{124}$$

where Δ is the conformal dimension of the scalar operator dual to φ. As we pointed out earlier, this expansion is less general than the expansion in eigenfunctions of δ_γ that we just discussed, since it is applicable only to backgrounds with an asymptotic scaling symmetry, but for such backgrounds it is technically simpler than the induced metric expansion. For an application of this expansion to backgrounds with asymptotic Lifshitz symmetry we refer the interested reader to [33, 34].

The dilatation operator (124) can be motivated as follows. Since the bulk theory is diffeomorphism invariant, the Hamiltonian does not explicitly depend on the radial coordinate r. It follows that the solution $\mathcal{S}$ of the HJ equation also only depends on the radial coordinate through the induced fields, i.e. $\mathcal{S} = \mathcal{S}[\gamma, \varphi]$. Hence, the radial derivative can be represented by the functional operator

$$\partial_r = \int d^d x \left(\dot{\gamma}_{ij}[\gamma,\varphi]\frac{\delta}{\delta\gamma_{ij}} + \dot{\varphi}[\gamma,\varphi]\frac{\delta}{\delta\varphi}\right). \tag{125}$$

Using the leading asymptotic form of the induced fields appropriate for asymptotically locally AdS backgrounds, namely (setting the AdS radius of curvature, ℓ, to 1)

$$\gamma_{ij} \sim e^{2r} g_{(0)ij}(x), \qquad \varphi \sim e^{-(d-\Delta)r}\varphi_{(0)}(x), \tag{126}$$

where $g_{(0)ij}(x)$ and $\varphi_{(0)}(x)$ are arbitrary sources, implies that

$$\dot{\gamma}_{ij} \sim 2\gamma_{ij}, \qquad \dot{\varphi} \sim -(d-\Delta)\varphi. \tag{127}$$

Inserting these expressions in the covariant representation (125) of the radial derivative we obtain

$$\partial_r \sim \int d^d x \left(2\gamma_{ij} \frac{\delta}{\delta \gamma_{ij}} + (\Delta - d)\varphi \frac{\delta}{\delta \varphi} \right) \equiv \delta_D, \tag{128}$$

where δ_D is the dilatation operator. This operator is ideally suited for asymptotically locally AdS backgrounds, but in order to construct the corresponding covariant expansion one must fix the dimension Δ from the beginning. Hence, contrary to the expansion in eigenfunctions of δ_γ, one must repeat the whole procedure for every different value of Δ.

As above, we start by writing the principal function as[11]

$$\mathcal{S} = \int_{\Sigma_r} d^d x \sqrt{\gamma} \mathcal{L}, \tag{129}$$

and formally expand $\mathcal{L}[\gamma, \varphi]$ in an expansion in eigenfunctions of the dilatation operator as

$$\mathcal{L} = \mathcal{L}_{(0)} + \mathcal{L}_{(2)} + \cdots + \tilde{\mathcal{L}}_{(d)} \log \mathrm{e}^{-2r} + \mathcal{L}_{(d)} + \cdots, \tag{130}$$

where

$$\delta_D \mathcal{L}_{(n)} = -n \mathcal{L}_{(n)}, \ \forall n < d, \quad \delta_D \tilde{\mathcal{L}}_{(d)} = -d \tilde{\mathcal{L}}_{(d)}. \tag{131}$$

A number of comments are in order here. Firstly, note that here we have defined $\mathcal{L}_{(n)}$ as eigenfunctions of δ_D, while earlier we only required $\mathcal{S}_{(\alpha_k)}$ to be eigenfunctions of the operator $\mathfrak{S}$. This implied that $\mathcal{L}_{(\alpha_k)}$ is an eigenfunction of $\mathfrak{S}$ up to a total derivative term. In order to derive (106), however, we argued that, since $\mathcal{L}_{(\alpha_k)}$ is defined only up to a total derivative, one can always choose the total derivatives terms in $\mathcal{L}_{(\alpha_k)}$ such that it is an eigenfunction of $\mathfrak{S}$. In (130) we have applied this argument already so that $\mathcal{L}_{(n)}$ are eigenfunctions of δ_D. A second comment concerns the eigenvalue of $\mathcal{L}_{(n)}$ under δ_D, and the corresponding subscript labeling $\mathcal{L}_{(n)}$. In general, these eigenvalues depend on the value of the conformal dimension Δ of the scalar operator and need not be integer. However, the terms of weight 0 and d are universal and are always there. What changes depending on the value of Δ is the intermediate terms. Finally, notice that we have included the logarithmic term already in the expansion (130), introducing explicit cut-off dependence. We could have proceeded instead

[11] To keep in line with the original notation in [39], we define the density $\mathcal{L}$ without $\sqrt{\gamma}$ here, in contrast to the earlier definition (103).

using dimensional regularization as in the expansion in eigenfunctions of δ_γ above, but it is instructive to discuss this alternative argument as well.

In particular, the explicit cut-off dependence introduced in the expansion (130) implies that the term $\mathcal{L}_{(d)}$ transforms inhomogeneously under δ_D. In order to derive the action of the dilatation operator on the coefficient $\mathcal{L}_{(d)}$ we recall that the full on-shell action must not depend explicitly on the radial coordinate r, as a consequence of the diffeomorphism invariance of the bulk action. Hence, requiring that ∂_r gives asymptotically the same result as δ_D we must have

$$\partial_r\left(\sqrt{\gamma}(\tilde{\mathcal{L}}_{(d)}\log e^{-2r}+\mathcal{L}_{(d)})\right)\sim\delta_D\left(\sqrt{\gamma}(\tilde{\mathcal{L}}_{(d)}\log e^{-2r}+\mathcal{L}_{(d)})\right), \tag{132}$$

which determines, using $\delta_D\sqrt{\gamma}=d\sqrt{\gamma}$, that

$$\delta_D\mathcal{L}_{(d)}=-d\mathcal{L}_{(d)}-2\tilde{\mathcal{L}}_{(d)}. \tag{133}$$

This transformation of the finite part of the on-shell action implies that $\mathcal{L}_{(d)}$ cannot be a local function of the fields γ_{ij} and φ, unless $\tilde{\mathcal{L}}_{(d)}$ vanishes identically. This is summarized in the following lemma:

Lemma *If $\tilde{\mathcal{L}}_{(d)}$ is not identically zero, then the transformation $\delta_D\mathcal{L}_{(d)}=-d\mathcal{L}_{(d)}-2\tilde{\mathcal{L}}_{(d)}$ implies that $\mathcal{L}_{(d)}$ cannot be a local functional of the induced fields γ_{ij} and φ*

Proof What we need to show is that $\mathcal{L}_{(d)}$ cannot be a polynomial in derivatives. Suppose $\mathcal{L}_{(d)}$ is a polynomial in derivatives. Since $\mathcal{L}_{(d)}$ is scalar, derivatives must come in pairs and must be contracted with an inverse metric γ^{ij}. It follows that every polynomial in derivatives can be decomposed as a finite sum of eigenfunctions of the dilatation operator, namely,

$$\mathcal{L}_{(d)}=\mathcal{F}_{(0)}+\mathcal{F}_{(1)}+\cdots+\mathcal{F}_{(N)}, \tag{134}$$

for some positive integer N, where $\delta_D\mathcal{F}_{(n)}=-n\mathcal{F}_{(n)}$. Hence,

$$\delta_D\mathcal{L}_{(d)}=-\left(\mathcal{F}_{(1)}+2\mathcal{F}_{(2)}+\cdots+N\mathcal{F}_{(N)}\right)=-d(\mathcal{F}_{(0)}+\mathcal{F}_{(1)}+\cdots+\mathcal{F}_{(N)})-2\tilde{\mathcal{L}}_{(d)}. \tag{135}$$

Identifying terms of equal dilatation weight then gives

$$\mathcal{F}_{(n)}=0,\ n\neq d,\quad 2\tilde{\mathcal{L}}_{(d)}=(n-d)\mathcal{F}_{(n)}=0,\ n=d. \tag{136}$$

This implies that $\tilde{\mathcal{L}}_{(d)}=0$, contradicting the original hypothesis. □

In fact this is no accident. As we shall see, the term $\mathcal{L}_{(d)}$ corresponds to the renormalized on-shell action, while $\tilde{\mathcal{L}}_{(d)}$ is the conformal anomaly. The fact that $\tilde{\mathcal{L}}_{(d)}$ is the conformal anomaly we will see more explicitly below when we derive the trace Ward identity. However, the fact that $\mathcal{L}_{(d)}$ corresponds to the renormalized on-shell action can be deduced directly from the dilatation weight of the various terms in the covariant expansion. Note that $\mathcal{L}_{(n)}$ with $n < d$, as well as $\tilde{\mathcal{L}}_{(d)}$ all lead to divergences as $r \to \infty$. This is because $\mathcal{L}_{(n)} \sim e^{-nr}$ as $r \to \infty$ and $\sqrt{\gamma} \sim e^{dr}$. We therefore *define* the counterterms as

$$\boxed{S_{\rm ct} := -\int_{\Sigma_r} d^d x \sqrt{\gamma} \left(\mathcal{L}_{(0)} + \mathcal{L}_{(2)} + \cdots + \tilde{\mathcal{L}}_{(d)} \log \mathrm{e}^{-2r} \right).} \tag{137}$$

It follows that the renormalized on-shell action on the radial cut-off is

$$S_{\rm ren} := S_{\rm reg} + S_{\rm ct} = \int_{\Sigma_r} d^d x \sqrt{\gamma} \mathcal{L}_{(d)} + \cdots, \tag{138}$$

where the dots stand for terms of higher dilatation weight that vanish as $r \to \infty$. By construction, $S_{\rm ren}$ admits a finite limit as $r \to \infty$, namely

$$\boxed{\hat{S}_{\rm ren} := \lim_{r\to\infty} S_{\rm ren} = \lim_{r\to\infty} \int_{\Sigma_r} d^d x \sqrt{\gamma} \mathcal{L}_{(d)}.} \tag{139}$$

As we anticipated, the term $\mathcal{L}_{(d)}$, which is a non-local function of the induced fields, determines the renormalized on-shell action.

Let us now proceed to determine the divergent coefficients $\mathcal{L}_{(n)}$ with $n < d$ and $\tilde{\mathcal{L}}_{(d)}$. Since the canonical momenta are related to the on-shell action via the relations (94), it follows that the momenta also admit an expansion of the form

$$\pi^{ij} = \frac{\delta}{\delta \gamma_{ij}} \int_{\Sigma_r} d^d x \sqrt{\gamma} \mathcal{L} = \sqrt{\gamma} \left(\pi_{(0)}{}^{ij} + \pi_{(2)}{}^{ij} + \cdots + \tilde{\pi}_{(d)}{}^{ij} \log e^{-2r} + \pi_{(d)}{}^{ij} + \cdots \right), \tag{140a}$$

$$\pi_\varphi = \frac{\delta}{\delta \varphi} \int_{\Sigma_r} d^d x \sqrt{\gamma} \mathcal{L} = \sqrt{\gamma} (\pi_{\varphi(d-\Delta)} + \cdots + \tilde{\pi}_{\varphi(\Delta)} \log e^{-2r} + \pi_{\varphi(\Delta)} + \cdots). \tag{140b}$$

Note that $\delta_D \pi^i_{j(n)} = -n \pi^i_{j(n)}$ and $\delta_D \pi^{ij}{}_{(n)} = -(n+2) \pi^{ij}{}_{(n)}$. With these expansions at hand, we are ready to develop the recursive algorithm. Before we discuss the general algorithm, however, let us point out that the first two of the $\mathcal{L}_{(n)}$ coefficients can be obtained easily, without relying on the algorithm. From the asymptotic relations (127) and the expressions (89) for the canonical momenta we deduce that

$$\pi^{ij} \sim -\frac{1}{2\kappa^2} (d-1) \sqrt{\gamma} \gamma^{ij}, \quad \pi_\varphi \sim -\frac{1}{\kappa^2} (d-\Delta) \sqrt{\gamma} \varphi, \tag{141}$$

and hence

$$\pi_{(0)}{}^{ij} = -\frac{1}{2\kappa^2} (d-1) \gamma^{ij}, \quad \pi_{\varphi(d-\Delta)} = -\frac{1}{\kappa^2} (d-\Delta) \varphi. \tag{142}$$

Integrating ${\pi_{(0)}}^{ij}$ with respect to γ_{ij} determines $\mathcal{L}_{(0)}$, whereas integrating $\pi_{(d-\Delta)}$ with respect to φ (assuming $\Delta < d$) determines $\mathcal{L}_{(2(d-\Delta))}$. Namely,

$$\mathcal{L}_{(0)} = -\frac{1}{\kappa^2}(d-1), \quad \mathcal{L}_{(2(d-\Delta))} = -\frac{1}{2\kappa^2}(d-\Delta)\varphi^2. \tag{143}$$

As we shall see below, these results are reproduced by the general algorithm.

The first step in the algorithm it to relate the coefficients $\mathcal{L}_{(n)}$ with $n < d$ and $\tilde{\mathcal{L}}_{(d)}$ to the corresponding canonical momenta using the identity (108). Since

$$\delta_D \gamma_{ij} = 2\gamma_{ij}, \quad \delta_D \varphi = -(d-\Delta)\varphi, \tag{144}$$

applied to the dilatation operator this identity reads

$$2\pi^i_i - (d-\Delta)\pi_\varphi \varphi = \delta_D \left(\sqrt{\gamma}\mathcal{L}\right), \tag{145}$$

or, inserting the expansions (130) and (140),

$$\begin{aligned} &2\sqrt{\gamma}\left(\pi_{(0)} + \pi_{(2)} + \cdots + \tilde{\pi}_{(d)} \log e^{-2r} + \pi_{(d)} + \cdots\right) \\ &\quad - (d-\Delta)\sqrt{\gamma}\varphi(\pi_{\varphi(d-\Delta)} + \cdots + \tilde{\pi}_{\varphi(\Delta)} \log e^{-2r} + \pi_{\varphi(\Delta)} + \cdots) \\ &= \sqrt{\gamma}\left(d\mathcal{L}_{(0)} + (d-2)\mathcal{L}_{(2)} + \cdots + 0 \cdot \tilde{\mathcal{L}}_{(d)} \log \mathrm{e}^{-2r} - 2\tilde{\mathcal{L}}_{(d)} + 0 \cdot \mathcal{L}_{(d)} + \cdots\right). \end{aligned} \tag{146}$$

In order to equate terms of the same dilatation weight, i.e. to obtain the exact analogue of (108), we need to know the precise value of the scalar dimension Δ. However, this identity shows that the coefficients $\mathcal{L}_{(n)}$ of the on-shell action can always be expressed in terms of the coefficients in the expansion of the canonical momenta.

As an example, we can use (146) to determine $\mathcal{L}_{(0)}$. Provided $\Delta < d$, identifying terms of dilatation weight zero gives

$$\mathcal{L}_{(0)} = \frac{2}{d}\pi_{(0)} = \frac{2}{d}\left(-\frac{1}{2\kappa^2}d(d-1)\right) = -\frac{1}{\kappa^2}(d-1), \tag{147}$$

where we have used the trace of ${\pi_{(0)}}^{ij}$ given in (142) in the second equality. This is in agreement with the result (143) we found above. Similarly we deduce that

$$\tilde{\mathcal{L}}_{(d)} = -\pi_{(d)} + \frac{1}{2}(d-\Delta)\varphi\pi_{\varphi(\Delta)}. \tag{148}$$

As we shall see shortly, this relation is in fact the trace Ward identity. The general algorithm using the dilatation operator expansion can be summarized as follows:

The algorithm:

1. The first step is to use the identity (146) to express $\mathcal{L}_{(n)}$, for $n < d$, and $\tilde{\mathcal{L}}_{(d)}$, in terms of the canonical momenta by matching terms of equal dilatation weight. Note that the on-shell action, $\mathcal{L}$, depends only on the trace of π^{ij}.
2. The second step is to insert the expansions (140) into the Hamiltonian constraint (91a) and match terms of equal dilatation weight. This gives an iterative relation for the trace $\pi_{(n)}$ and $\pi_{\varphi(\Delta-d+n)}$ in terms of the momentum terms of lower dilatation weight.
3. Having determined $\pi_{(n)}$ and $\pi_{\varphi(\Delta-d+n)}$ at order n, we can use the relations we found in the first step to determine $\mathcal{L}_{(n)}$. The *full* momentum ${\pi_{(n)}}^{ij}$—i.e. not just its trace—is then obtained via the relations (107).
4. Steps 2 and 3 are iterated until all local terms are determined.

5.3 An Example

It is instructive to work out the counterterms explicitly in a concrete example. To this end, let us apply the dilatation operator expansion to asymptotically AdS gravity in five dimensions ($d = 4$) coupled to a scalar field, φ, dual to an operator of conformal dimension $\Delta = 3$, and with a general scalar potential. The action takes the form[12]

$$S = \int d^5x\sqrt{g}\left(-\frac{1}{2\kappa^2}R[g] + \frac{1}{2}g^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi + V(\varphi)\right), \tag{149}$$

where

$$V(\varphi) = \kappa^{-2}V_0 + \kappa^{-1}V_1\varphi + V_2\varphi^2 + \kappa V_3\varphi^3 + \kappa^2 V_4\varphi^4 + \cdots, \tag{150}$$

with

$$V_0 = \Lambda = -6, \quad V_1 = 0, \quad V_2 = \frac{1}{2}m^2 = -3/2. \tag{151}$$

Let us now implement step by step the algorithm we described above. The first step is to use (146) to express all local terms of the expansion of the on-shell action, i.e. $\mathcal{L}_{(n)}$, $n < d$, and $\tilde{\mathcal{L}}_{(d)}$, in terms of the canonical momenta by matching terms of equal dilatation weight. For the system at hand, and dropping the total divergence term, (146) becomes

[12]Note that the scalar field here is rescaled by a factor of $\sqrt{2\kappa^2}$ relative to the scalar in (83).

$$2\left(\pi_{(0)}+\pi_{(1)}+\pi_{(2)}+\pi_{(3)}+\tilde{\pi}_{(4)}\log e^{-2r}+\pi_{(4)}+\cdots\right)$$
$$-\varphi(\pi^{\varphi}{}_{(1)}+\pi^{\varphi}{}_{(2)}+\tilde{\pi}^{\varphi}{}_{(3)}\log e^{-2r}+\pi^{\varphi}{}_{(3)}+\cdots) \tag{152}$$
$$=\left(4\mathcal{L}_{(0)}+3\mathcal{L}_{(1)}+2\mathcal{L}_{(2)}+\mathcal{L}_{(3)}+0\cdot\tilde{\mathcal{L}}_{(4)}\log \mathrm{e}^{-2r}-2\tilde{\mathcal{L}}_{(4)}+0\cdot\mathcal{L}_{(4)}+\cdots\right). \tag{153}$$

Matching terms of equal dilatation weight we obtain

$$\begin{aligned}
\mathcal{L}_{(0)}&=\frac{1}{2}\pi_{(0)}=-3/\kappa^2,\\
\mathcal{L}_{(1)}&=\frac{2}{3}\pi_{(1)},\\
\mathcal{L}_{(2)}&=\pi_{(2)}-\frac{1}{2}\varphi\pi^{\varphi}{}_{(1)}=\pi_{(2)}+\frac{1}{2}\varphi^2,\\
\mathcal{L}_{(3)}&=2\pi_{(3)}-\varphi\pi^{\varphi}{}_{(2)},\\
\tilde{\mathcal{L}}_{(4)}&=-\pi_{(4)}+\frac{1}{2}\varphi\pi^{\varphi}{}_{(3)},
\end{aligned} \tag{154}$$

as well as the constraint on the momenta

$$\tilde{\pi}_{(4)}-\frac{1}{2}\varphi\tilde{\pi}^{\varphi}{}_{(3)}=0. \tag{155}$$

Note that $\mathcal{L}_{(4)}$ is not determined, but it does not contribute to the divergences of the on-shell action. As we saw in (139), it is the renormalized part of the on-shell action. At this point we have determined all divergent terms of the on-shell action in terms of the canonical momenta.

The second step is to insert the covariant expansions for the momenta into the Hamiltonian constraint (91a), which in this case takes the form

$$\mathcal{H}=\sqrt{\gamma}\left\{\frac{1}{2\kappa^2}R[\gamma]+2\kappa^2\gamma^{-1}\left(\pi^{ij}\pi_{ij}-\frac{1}{3}\pi^2\right)+\frac{1}{2}\gamma^{-1}(\pi^{\varphi})^2-\frac{1}{2}\gamma^{ij}\partial_i\varphi\partial_j\varphi-V(\varphi)\right\}=0. \tag{156}$$

Inserting the covariant expansions for the momenta and equating terms of equal dilatation weight we obtain

$$2\kappa^2\left(\pi_{(0)}{}^{ij}\pi_{(0)ij}-\frac{1}{3}\pi_{(0)}{}^2\right)-\kappa^{-2}V_0=0,$$
$$4\kappa^2\left(\pi_{(0)}{}^{ij}\pi_{(1)ij}-\frac{1}{3}\pi_{(0)}\pi_{(1)}\right)-\kappa^{-1}V_1\varphi=0,$$
$$\frac{1}{2\kappa^2}R[\gamma]+2\kappa^2\left(2\pi_{(0)}{}^{ij}\pi_{(2)ij}+\pi_{(1)}{}^{ij}\pi_{(1)ij}-\frac{2}{3}\pi_{(0)}\pi_{(2)}-\frac{1}{3}\pi_{(1)}{}^2\right)+\frac{1}{2}(\pi^{\varphi})^2-V_2\varphi^2=0,$$
$$4\kappa^2\left(\pi_{(0)}{}^{ij}\pi_{(3)ij}+\pi_{(1)}{}^{ij}\pi_{(2)ij}-\frac{1}{3}\pi_{(0)}\pi_{(3)}-\frac{1}{3}\pi_{(1)}\pi_{(2)}\right)+\pi^{\varphi}{}_{(1)}\pi^{\varphi}{}_{(2)}-\kappa V_3\varphi^3=0,$$

$$
\begin{aligned}
&2\kappa^2\left(2\pi_{(0)}{}^{ij}\pi_{(4)ij}+2\pi_{(1)}{}^{ij}\pi_{(3)ij}+\pi_{(2)}{}^{ij}\pi_{(2)ij}-\frac{2}{3}\pi_{(0)}\pi_{(4)}-\frac{2}{3}\pi_{(1)}\pi_{(3)}-\frac{1}{3}\pi_{(2)}{}^2\right)\\
&\quad+\pi^{\varphi}{}_{(1)}\pi^{\varphi}{}_{(3)}+\frac{1}{2}(\pi^{\varphi}{}_{(2)})^2-\frac{1}{2}\gamma^{ij}\partial_i\varphi\partial_j\varphi-\kappa^2V_4\varphi^4=0,\\
&4\kappa^2\left(\pi_{(0)}{}^{ij}\tilde{\pi}_{(4)ij}-\pi_{(0)}\tilde{\pi}_{(4)}\right)+\pi^{\varphi}{}_{(1)}\tilde{\pi}^{\varphi}{}_{(3)}=0.
\end{aligned}
\tag{157}
$$

The first of these equations is trivially satisfied, while the second equation determines $\pi_{(1)}=0$ and hence from above $\mathcal{L}_{(1)}=0$. Next we must use the third step in the algorithm, namely the relations

$$
\pi_{(n)}{}^{ij}=\frac{1}{\sqrt{\gamma}}\frac{\delta}{\delta\gamma_{ij}}\int d^dx\sqrt{\gamma}\mathcal{L}_{(n)},\quad \tilde{\pi}_{(d)}{}^{ij}=\frac{1}{\sqrt{\gamma}}\frac{\delta}{\delta\gamma_{ij}}\int d^dx\sqrt{\gamma}\tilde{\mathcal{L}}_{(d)}. \tag{158}
$$

This allows us to determine the full momentum $\pi_{(n)}{}^{ij}$ from its trace $\pi_{(n)}$ for $n<d$. In particular, we conclude $\pi_{(1)}{}^{ij}=0$. The third equation in (157) gives

$$
\pi_{(2)}-\frac{1}{4\kappa^2}R[\gamma]-\varphi^2, \tag{159}
$$

and hence,

$$
\mathcal{L}_{(2)}=-\frac{1}{4\kappa^2}R[\gamma]-\frac{1}{2}\varphi^2. \tag{160}
$$

It follows that

$$
\pi_{(2)}{}^{ij}=\frac{1}{\sqrt{\gamma}}\frac{\delta}{\delta\gamma_{ij}}\int d^dx\sqrt{\gamma}\mathcal{L}_{(2)}=\frac{1}{4\kappa^2}\left(R^{ij}-\frac{1}{2}R\gamma^{ij}\right)-\frac{1}{4}\varphi^2\gamma^{ij}. \tag{161}
$$

Continuing this recursive procedure we determine

$$
\begin{aligned}
\mathcal{L}_{(3)}&=\kappa V_3\varphi^3,\\
\tilde{\mathcal{L}}_{(4)}&=\frac{1}{16\kappa^2}\left(R^{ij}R_{ij}-\frac{1}{3}R^2\right)-\frac{1}{24}R\varphi^2-\frac{1}{4}\gamma^{ij}\partial_i\varphi\partial_j\varphi-\frac{\kappa^2}{2}\left(V_4-\frac{9}{2}V_3^2+\frac{1}{6}\right)\varphi^4,\\
\pi_{(3)}{}^{ij}&=\frac{\kappa}{2}V_3\varphi^3\gamma^{ij},\\
\pi^{\varphi}{}_{(2)}&=3\kappa V_3\varphi^2,\\
\tilde{\pi}^{\varphi}{}_{(3)}&=\frac{1}{12}R\varphi+\frac{1}{2}\Box_\gamma\varphi-2\kappa^2\left(V_4-\frac{9}{2}V_3^2+\frac{1}{6}\right)\varphi^3,\\
\tilde{\pi}_{(4)}{}^{ij}&=\frac{1}{16\kappa^2}\left[-2R^{kl}R_k{}^i{}_l{}^j+\frac{1}{3}D^iD^jR-\Box_\gamma R^{ij}+\frac{2}{3}RR^{ij}\right.\\
&\qquad\left.+\frac{1}{2}\gamma^{ij}\left(R^{kl}R_{kl}+\frac{1}{3}\Box_\gamma R-\frac{1}{3}R^2\right)\right]\\
&\quad+\frac{1}{24}\left(R^{ij}-\frac{1}{2}R\gamma^{ij}\right)\varphi^2-\frac{1}{24}\left(D^iD^j-\gamma^{ij}\Box_\gamma\right)\varphi^2+\frac{1}{4}\partial^i\varphi\partial^j\varphi-\frac{1}{8}\gamma^{ij}\partial^k\varphi\partial_k\varphi\\
&\quad-\frac{\kappa^2}{4}\left(V_4-\frac{9}{2}V_3^2+\frac{1}{6}\right)\varphi^4\gamma^{ij}.
\end{aligned}
\tag{162}
$$

Note that these satisfy the identity

$$\tilde{\pi}_{(4)} - \frac{1}{2}\varphi\tilde{\pi}^{\varphi}{}_{(3)} = 0, \tag{163}$$

as required.

6 Renormalized One-Point Functions and Ward Identities

We found above that the renormalized action (139) admits a finite limit, $\hat{S}_{\text{ren}}$, as $r \to \infty$. The AdS/CFT dictionary identifies this with the generating functional of renormalized connected correlation functions in the dual quantum field theory. In particular, the first derivatives of the renormalized action with respect to the sources correspond to the one-point functions of the dual operators. This implies that we can identify the renormalized one-point functions with certain terms in the covariant expansion of the canonical momenta in eigenfunctions of the dilatation operator. Namely, we define

$$\langle \mathcal{T}^{ij} \rangle_{\text{ren}} = -2|\gamma|^{-1/2} \frac{\delta S_{\text{ren}}}{\delta \gamma_{ij}} = -2\pi_{(d)}{}^{ij}, \tag{164a}$$

$$\langle \mathcal{O} \rangle_{\text{ren}} = |\gamma|^{-1/2} \frac{\delta S_{\text{ren}}}{\delta \varphi} = \pi_{\varphi(\Delta)}. \tag{164b}$$

Note that these expressions are evaluated on the cut-off, i.e. they are covariant expressions of the induced metric and scalar field. Since these fields asymptotically behave as

$$\gamma_{ij} \sim e^{2r} g_{(0)ij}, \quad \varphi \sim e^{-(d-\Delta)r} \varphi_{(0)}, \tag{165}$$

and since $\mathcal{S}_{\text{ren}}$ has a finite limit as $r \to \infty$, it follows that we must multiply these one-point functions with a suitable factor of the radial coordinate to obtain finite values as $r \to \infty$. In particular, we define

$$\boxed{\begin{aligned} \langle \hat{\mathcal{T}}^{ij} \rangle_{\text{ren}} &:= \lim_{r\to\infty} e^{(d+2)r} \langle \mathcal{T}^{ij} \rangle_{\text{ren}} = -2|g_{(0)}|^{-1/2} \frac{\delta \hat{S}_{\text{ren}}}{\delta g_{(0)ij}} = -2\hat{\pi}_{(d)}{}^{ij}, \\ \langle \hat{\mathcal{O}} \rangle_{\text{ren}} &:= \lim_{r\to\infty} e^{\Delta r} \langle \mathcal{O} \rangle_{\text{ren}} = |g_{(0)}|^{-1/2} \frac{\delta \hat{S}_{\text{ren}}}{\delta \varphi_{(0)}} = \hat{\pi}_{\varphi(\Delta)}. \end{aligned}} \tag{166}$$

Using these expressions for the renormalized one-point functions we can now derive the holographic Ward identities. Inserting the expansions (140) into the momentum constraint (91) and matching terms of equal dilatation weight gives for the terms with weight d

$$-2D_j\pi_{(d)}{}^{ij} + \pi_{\varphi(\Delta)}\partial^i\varphi = 0. \tag{167}$$

Rescaling this with the appropriate radial factor and taking the limit $r \to \infty$ leads to the diffeomorphism Ward identity

$$\boxed{D_{(0)j}\langle\hat{\mathcal{T}}^{ij}\rangle_{\text{ren}} + \langle\hat{\mathcal{O}}\rangle_{\text{ren}}\partial^i\varphi_{(0)} = 0.} \tag{168}$$

Finally, in order to derive the trace Ward identity note that under an infinitesimal Weyl transformation the renormalized action transforms as

$$\delta_\sigma S_{\text{ren}} = \int_{\Sigma_r}\sqrt{\gamma}(-2\tilde{\mathcal{L}}_{(d)})\delta\sigma + \text{total derivative.} \tag{169}$$

This follows from the fact that such a transformation corresponds to the infinitesimal bulk diffeomorphism $r \to r + \delta\sigma(x)$. It follows that the conformal anomaly $\mathcal{A}$ is given by

$$\mathcal{A} := 2\tilde{\mathcal{L}}_{(d)}. \tag{170}$$

To see that this is compatible with the trace Ward identity, recall that we have shown in (148) that

$$-2\pi_{(d)} + (d-\Delta)\varphi\pi_{\varphi(\Delta)} = 2\tilde{\mathcal{L}}_{(d)}, \tag{171}$$

which, using the identifications (166), becomes

$$\boxed{\langle\hat{\mathcal{T}}_i^i\rangle_{\text{ren}} + (d-\Delta)\varphi_{(0)}\langle\hat{\mathcal{O}}\rangle_{\text{ren}} = \mathcal{A}.} \tag{172}$$

It should be emphasized that these Ward identities hold in the presence of arbitrary sources. This has important implications. Namely, even if the conformal anomaly vanishes numerically on a particular background where the sources are set to zero, the anomaly does contribute to some n-point function because the nth derivative of the anomaly with respect to the sources will not be zero even when evaluated at zero sources. The anomaly therefore is a genuine property of the quantum field theory and affects the dynamics even in flat space.

7 Fefferman–Graham Asymptotic Expansions

Having obtained the asymptotic solution of the HJ equation in the form of a covariant expansion in eigenfunctions of some suitable operator $\mathfrak{S}$, we can now use the first order flow equations (96) to construct the asymptotic Fefferman–Graham expansions for the induced fields γ_{ij} and φ. In order to integrate these expansions, however, we must pick a specific example and a specific solution of the HJ equation. We will therefore demonstrate how this works in the example we worked out above.

Inserting the expansions (140) in the flow equations (96) we get[13]

$$\dot{\gamma}_{ij} = 4\kappa^2 \left(\gamma_{ik}\gamma_{jl} - \frac{1}{3}\gamma_{kl}\gamma_{ij}\right)\left(\pi_{(0)}{}^{ij} + \pi_{(2)}{}^{ij} + \cdots + \tilde{\pi}_{(4)}{}^{ij}\log e^{-2r} + \pi_{(4)}{}^{ij} + \cdots\right),$$
$$\dot{\varphi} = \pi_{\varphi(1)} + \cdots + \tilde{\pi}_{\varphi(3)}\log e^{-2r} + \pi_{\varphi(3)} + \cdots . \tag{173}$$

From the expressions (162) above we obtain

$$\begin{aligned}
\pi_{(0)ij} - \frac{1}{3}\pi_{(0)}\gamma_{ij} &= \frac{1}{2\kappa^2}\gamma_{ij}, \\
\pi_{(2)ij} - \frac{1}{3}\pi_{(2)}\gamma_{ij} &= \frac{1}{4\kappa^2}\left(R_{ij} - \frac{1}{6}R\gamma_{ij}\right) + \frac{1}{12}\varphi^2\gamma_{ij}, \\
\pi_{(3)ij} - \frac{1}{3}\pi_{(3)}\gamma_{ij} &= -\frac{\kappa}{6}V_3\varphi^3\gamma_{ij}, \\
\tilde{\pi}_{(4)ij} - \frac{1}{3}\tilde{\pi}_{(4)}\gamma_{ij} &= \frac{1}{16\kappa^2}\Big[-2R^{kl}R_{kilj} + \frac{1}{3}D_iD_jR - \Box_\gamma R_{ij} + \frac{2}{3}RR_{ij} \\
&\qquad + \frac{1}{2}\gamma_{ij}\left(R^{kl}R_{kl} + \frac{1}{3}\Box_\gamma R - \frac{1}{3}R^2\right)\Big] \\
&\quad + \frac{1}{24}\left(R_{ij} - \frac{5}{6}R\gamma_{ij}\right)\varphi^2 - \frac{1}{24}\left(D_iD_j - \gamma_{ij}\Box_\gamma\right)\varphi^2 + \frac{1}{4}\partial_i\varphi\partial_j\varphi \\
&\quad - \frac{1}{8}\gamma_{ij}\partial^k\varphi\partial_k\varphi + \frac{\kappa^2}{12}\left(V_4 - \frac{9}{2}V_3^2 + \frac{1}{6}\right)\varphi^4\gamma_{ij} - \frac{1}{12}\varphi\Box_\gamma\varphi\gamma_{ij}, \\
\pi_{\varphi(1)} &= -\varphi, \\
\pi_{\varphi(2)} &= 3\kappa V_3\varphi^2, \\
\tilde{\pi}_{\varphi(3)} &= \frac{1}{12}R\varphi + \frac{1}{2}\Box_\gamma\varphi - 2\kappa^2\left(V_4 - \frac{9}{2}V_3^2 + \frac{1}{6}\right)\varphi^3.
\end{aligned} \tag{174}$$

Using these expressions we can integrate the flow equations (173) straightforwardly. There are two ways to solve these equations order by order asymptotically as $r \to \infty$. One way is to make an explicit Fefferman–Graham ansatz for the asymptotic expansions for γ_{ij} and φ and insert them in the flow equations. This will result in algebraic equations for the coefficients. A more general way that does not require prior knowledge of the form of the asymptotic expansion is expanding the induced fields formally as

$$\gamma_{ij} = \gamma_{ij}^{(0)} + \gamma_{ij}^{(1)} + \gamma_{ij}^{(2)} + \gamma_{ij}^{(3)} + \cdots, \quad \varphi = \varphi^{(0)} + \varphi^{(1)} + \varphi^{(2)} + \cdots, \tag{175}$$

[13] Note one needs to adjust these for the different normalization of the scalar.

where each order is assumed to be asymptotically subleading relative to the previous one, but without assuming a specific functional form. Inserting these expansions in the flow equations results in a sequence of differential equations that can be solved order by order. To leading order we get the homogeneous equations

$$\dot{\gamma}_{ij}^{(0)} = 2\gamma_{ij}^{(0)}, \quad \dot{\varphi}^{(0)} = -\varphi^{(0)}, \tag{176}$$

and hence

$$\gamma_{ij}^{(0)} = e^{2r} g_{(0)ij}, \quad \varphi^{(0)} = e^{-r}\varphi_{(0)}, \tag{177}$$

where $g_{(0)ij}(x)$ and $\varphi_{(0)}(x)$ are arbitrary integration sources. At the next order for γ_{ij} we still get the same homogeneous equation

$$\dot{\gamma}_{ij}^{(1)} = 2\gamma_{ij}^{(1)}. \tag{178}$$

However, we have already introduced an arbitrary source at order 0 and, since $\gamma_{ij}^{(1)}$ is asymptotically subleading relative to $\gamma_{ij}^{(0)}$ by the hypothesis, we must set $\gamma_{ij}^{(1)} = 0$. At the next order we obtain the inhomogeneous equations

$$\begin{aligned} \dot{\gamma}_{ij}^{(2)} &= 2\gamma_{ij}^{(2)} + R[g_{(0)}]_{ij} - \frac{1}{6}R[g_{(0)}]g_{(0)ij} + \frac{\kappa^2}{3}\varphi_{(0)}^2 g_{(0)ij}, \\ \dot{\varphi}^{(1)} &= -\varphi^{(1)} + 3\kappa V_3 \varphi_{(0)}^2 e^{-2r}. \end{aligned} \tag{179}$$

Discarding the homogeneous solutions again, the inhomogeneous solutions are

$$\begin{aligned} \gamma_{ij}^{(2)} &= -\frac{1}{2}\left(R_{ij}[g_{(0)}] - \frac{1}{6}R[g_{(0)}]g_{(0)ij} + \frac{\kappa^2}{3}\varphi_{(0)}^2 g_{(0)ij}\right), \\ \varphi^{(1)} &= -3\kappa V_3 e^{-2r}\varphi_{(0)}^2. \end{aligned} \tag{180}$$

At the next order for the metric we get

$$\gamma_{ij}^{(3)} = \frac{8}{9}\kappa^3 V_3 e^{-r}\varphi_{(0)}^3 g_{(0)ij}, \tag{181}$$

while, using the following expansions of the momenta

$$\begin{aligned} \pi_{(2)ij} - \frac{1}{3}\pi_{(2)}\gamma_{ij} = {} & \frac{1}{4\kappa^2}\left(R[g_{(0)}]_{ij} - \frac{1}{6}R[g_{(0)}]g_{(0)ij}\right) + \frac{1}{12}\varphi_{(0)}^2 g_{(0)ij} + \frac{1}{6}\varphi^{(0)}\varphi^{(1)}\gamma_{ij}^{(0)} \\ & + e^{-2r}\Bigg[\frac{1}{4\kappa^2}\left(R^k_{(i}[g_{(0)}]\gamma_{kj)}^{(2)} - R_i{}^k{}_j{}^l[g_{(0)}]\gamma_{kl}^{(2)} + D_{(0)(i}D^k_{(0)}\gamma_{kj)}^{(2)}\right. \\ & \left. -\frac{1}{2}\left(\Box_{(0)}\gamma_{ij}^{(2)} + g_{(0)}{}^{kl}D_{(0)i}D_{(0)j}\gamma_{kl}^{(2)}\right) - \frac{1}{6}R[g_{(0)}]\gamma_{ij}^{(2)}\right. \end{aligned}$$

$$
\begin{aligned}
&\qquad -\frac{1}{6}g_{(0)ij}\left(-R^{kl}[g_{(0)}]\gamma^{(2)}_{kl}+D^k_{(0)}D^l_{(0)}\gamma^{(2)}_{kl}-g^{kl}_{(0)}\Box_{(0)}\gamma^{(2)}_{kl}\right)\Big)\Big]\\
&\quad +\frac{1}{12}\left(\left[(\varphi^{(1)})^2+2\varphi^{(0)}\varphi^{(2)}\right]\gamma^{(0)}_{ij}+(\varphi^{(0)})^2\gamma^{(2)}_{ij}\right)+\mathcal{O}(e^{-3r}),
\end{aligned}
$$

$$
\begin{aligned}
\pi_{(3)ij}-\frac{1}{3}\pi_{(3)}\gamma_{ij} &= -\frac{\kappa}{6}V_3(\varphi^{(0)})^3\gamma^{(0)}_{ij}-\frac{\kappa}{2}V_3(\varphi^{(0)})^2\varphi^{(1)}\gamma^{(0)}_{ij}+\mathcal{O}(e^{-3r}),\\
\pi_{\varphi(2)} &= 3\kappa V_3(\varphi^{(0)})^2+6\kappa V_3\varphi^{(0)}\varphi^{(1)}+\mathcal{O}(e^{-4r}),
\end{aligned}
\tag{182}
$$

we obtain the next order equations

$$
\begin{aligned}
\dot\gamma^{(4)}_{ij} = 2\gamma^{(4)}_{ij} &+ (-2r)e^{-2r}\Bigg\{\frac{1}{16\kappa^2}\Big[-2R^{kl}[g_{(0)}]R_{kilj}[g_{(0)}]+\frac{1}{3}D_{(0)i}D_{(0)j}R[g_{(0)}]\\
&\qquad -\Box_{(0)}R_{ij}[g_{(0)}]+\frac{2}{3}R[g_{(0)}]R_{ij}[g_{(0)}]\\
&\qquad +\frac{1}{2}g_{(0)ij}\left(R^{kl}[g_{(0)}]R_{kl}[g_{(0)}]+\frac{1}{3}\Box_{(0)}R[g_{(0)}]-\frac{1}{3}R^2[g_{(0)}]\right)\Big]\\
&\quad +\frac{1}{24}\left(R_{ij}[g_{(0)}]-\frac{5}{6}R[g_{(0)}]g_{(0)ij}\right)\varphi^2_{(0)}-\frac{1}{24}\left(D_{(0)i}D_{(0)j}-g_{(0)ij}\Box_{(0)}\right)\varphi^2_{(0)}\\
&\quad +\frac{1}{4}\partial_i\varphi_{(0)}\partial_j\varphi_{(0)}-\frac{1}{8}g_{(0)ij}\partial^k\varphi_{(0)}\partial_k\varphi_{(0)}+\frac{\kappa^2}{12}\left(V_4-\frac{9}{2}V_3^2+\frac{1}{6}\right)\varphi^4_{(0)}g_{(0)ij}\\
&\quad +\frac{1}{6}\varphi_{(0)}\tilde\varphi_{(2)}g_{(0)ij}-\frac{1}{12}\varphi_{(0)}\Box_{(0)}\varphi_{(0)}g_{(0)ij}\Bigg\}\\
+e^{-2r}\Bigg\{&\hat\pi_{(4)ij}-\frac{1}{3}g_{(0)ij}\hat\pi_{(4)}+\frac{1}{4\kappa^2}\Big(R^k_{(i}[g_{(0)}]\gamma^{(2)}_{kj)}-R_i{}^k{}_j{}^l[g_{(0)}]\gamma^{(2)}_{kl}+D_{(0)(i}D^k_{(0)}\gamma^{(2)}_{kj)}\\
&-\frac{1}{2}\left(\Box_{(0)}\gamma^{(2)}_{ij}+g_{(0)}{}^{kl}D_{(0)i}D_{(0)j}\gamma^{(2)}_{kl}\right)-\frac{1}{6}R[g_{(0)}]\gamma^{(2)}_{ij}\\
&-\frac{1}{6}g_{(0)ij}\left(-R^{kl}[g_{(0)}]\gamma^{(2)}_{kl}+D^k_{(0)}D^l_{(0)}\gamma^{(2)}_{kl}-g^{kl}_{(0)}\Box_{(0)}\gamma^{(2)}_{kl}\right)\Big)\\
&+\frac{1}{12}\left(\left[9\kappa^2V_3^2\varphi^4_{(0)}+2\varphi_{(0)}\hat\varphi_{(2)}\right]g_{(0)ij}+\varphi^2_{(0)}\gamma^{(2)}_{ij}\right)+\frac{3\kappa^2}{2}V_3^2\varphi^4_{(0)}g_{(0)ij}\Bigg\},
\end{aligned}
\tag{183}
$$

and

$$
\begin{aligned}
\dot\varphi^{(2)} = -\varphi^{(2)}&\\
&+e^{-3r}(-2r)\left[\frac{1}{12}R[g_{(0)}]\varphi_{(0)}+\frac{1}{2}\Box_{(0)}\varphi_{(0)}-2\kappa^2\left(V_4-\frac{9}{2}V_3^2+\frac{1}{6}\right)\varphi^3_{(0)}\right]\\
&+e^{-3r}\left(\hat\pi_{\varphi(3)}-18\kappa^2V_3^2\varphi^3_{(0)}\right).
\end{aligned}
\tag{184}
$$

The inhomogeneous solutions of these equations take the form

$$
\gamma^{(4)}_{ij}=e^{-2r}\left(-2rh_{(4)ij}+g_{(4)ij}\right),\quad \varphi^{(2)}=e^{-3r}\left(-2r\widetilde\varphi_{(2)}+\hat\varphi_{(2)}\right),
\tag{185}
$$

where

$$\begin{aligned} h_{(4)ij} = -\kappa^2 \Bigg\{ \frac{1}{16\kappa^2} \Bigg[&-2R^{kl}[g_{(0)}]R_{kilj}[g_{(0)}] + \frac{1}{3}D_{(0)i}D_{(0)j}R[g_{(0)}] - \Box_{(0)}R_{ij}[g_{(0)}] \\ &+ \frac{2}{3}R[g_{(0)}]R_{ij}[g_{(0)}] + \frac{1}{2}g_{(0)ij}\Big(R^{kl}[g_{(0)}]R_{kl}[g_{(0)}] \\ &+ \frac{1}{3}\Box_{(0)}\ R[g_{(0)}] - \frac{1}{3}R^2[g_{(0)}]\Big)\Bigg] \\ &+ \frac{1}{24}\left(R_{ij}[g_{(0)}] - \frac{5}{6}R[g_{(0)}]g_{(0)ij}\right)\varphi_{(0)}^2 \\ &- \frac{1}{24}\left(D_{(0)i}D_{(0)j} - g_{(0)ij}\Box_{(0)}\right)\varphi_{(0)}^2 + \frac{1}{4}\partial_i\varphi_{(0)}\partial_j\varphi_{(0)} - \frac{1}{8}g_{(0)ij}\partial^k\varphi_{(0)}\partial_k\varphi_{(0)} \\ &+ \frac{\kappa^2}{12}\left(V_4 - \frac{9}{2}V_3^2 + \frac{1}{6}\right)\varphi_{(0)}^4 g_{(0)ij} + \frac{1}{6}\varphi_{(0)}\tilde{\varphi}_{(2)}g_{(0)ij} - \frac{1}{12}\varphi_{(0)}\Box_{(0)}\varphi_{(0)}g_{(0)ij}\Bigg\}, \end{aligned} \tag{186}$$

$$\begin{aligned} g_{(4)ij} = -\kappa^2 \Bigg\{ &\hat{\pi}_{(4)ij} - \frac{1}{3}g_{(0)ij}\hat{\pi}_{(4)} + \frac{1}{4\kappa^2}\Big(R^k_{(i}[g_{(0)}]\gamma^{(2)}_{kj)} - R_i{}^k{}_j{}^l[g_{(0)}]\gamma^{(2)}_{kl} + D_{(0)(i}D^k_{(0)}\gamma^{(2)}_{kj)} \\ &- \frac{1}{2}\left(\Box_{(0)}\gamma^{(2)}_{ij} + g_{(0)}{}^{kl}D_{(0)i}D_{(0)j}\gamma^{(2)}_{kl}\right) - \frac{1}{6}R[g_{(0)}]\gamma^{(2)}_{ij} \\ &- \frac{1}{6}g_{(0)ij}\left(-R^{kl}[g_{(0)}]\gamma^{(2)}_{kl} + D^k_{(0)}D^l_{(0)}\gamma^{(2)}_{kl} - g^{kl}_{(0)}\Box_{(0)}\gamma^{(2)}_{kl}\right)\Big) \\ &+ \frac{1}{12}\left(\left[9\kappa^2V_3^2\varphi^4_{(0)} + 2\varphi_{(0)}\hat{\varphi}_{(2)}\right]g_{(0)ij} + \varphi^2_{(0)}\gamma^{(2)}_{ij}\right) + \frac{3\kappa^2}{2}V_3^2\varphi^4_{(0)}g_{(0)ij}\Bigg\} - \frac{1}{2}h_{(4)ij}, \end{aligned} \tag{187}$$

$$\widetilde{\varphi}_{(2)} = -\frac{1}{3}\left[\frac{1}{12}R[g_{(0)}]\varphi_{(0)} + \frac{1}{2}\Box_{(0)}\varphi_{(0)} - 2\kappa^2\left(V_4 - \frac{9}{2}V_3^2 + \frac{1}{6}\right)\varphi^3_{(0)}\right], \tag{188}$$

and

$$\hat{\varphi}_{(2)} = -\frac{1}{3}\left(\hat{\pi}_{\varphi(3)} - 18\kappa^2V_3^2\varphi^3_{(0)} - \frac{2}{3}\widetilde{\varphi}_{(2)}\right). \tag{189}$$

This completes the computation since the coefficients $\hat{\pi}_{(4)ij}$ and $\hat{\pi}_{(3)}$ have been identified above with the renormalized one-point functions. In particular, taking the trace of the expression for $g_{(4)ij}$ relates the trace of $\hat{\pi}_{(4)ij}$ with the trace of $g_{(4)ij}$. Inserting this back in the expression for $g_{(4)ij}$ one obtains the renormalized stress tensor $\hat{\pi}_{(4)ij}$ in terms of $g_{(4)ij}$, its trace, and lower order terms that are explicitly expressed in terms of the sources.

Acknowledgments I would like to thank the International Institute of Physics, Natal, for the hospitality and financial support during the School on Theoretical Frontiers in Black Holes and Cosmology, June 8–19, 2015, where these lectures were delivered.

Appendix

ADM Identities

A few identities relating to the ADM decomposition (84) of the metric are collected in this appendix. In particular, in matrix form, the metric (84) and its inverse are

$$g = \begin{pmatrix} N^2 + N_k N^k & N_i \\ N_i & \gamma_{ij} \end{pmatrix}, \quad g^{-1} = \begin{pmatrix} 1/N^2 & -N^i/N^2 \\ -N^i/N^2 & \gamma^{ij} + N^i N^j/N^2 \end{pmatrix}, \tag{190}$$

where the indices $i = 1, \ldots, d$ are raised and lowered respectively with γ^{ij} and γ_{ij}. Moreover, the Christoffel symbols $\Gamma^{\rho}_{\mu\nu}[g]$ can be decomposed into the following components in terms of N, N_i and γ_{ij}:

$$\begin{aligned}
\Gamma^r_{rr} &= N^{-1}\left(\dot{N} + N^i \partial_i N - N^i N^j K_{ij}\right), \\
\Gamma^r_{ri} &= N^{-1}\left(\partial_i N - N^j K_{ij}\right), \\
\Gamma^r_{ij} &= -N^{-1} K_{ij}, \\
\Gamma^i_{rr} &= -N^{-1} N^i \dot{N} - N D^i N - N^{-1} N^i N^j \partial_j N + \dot{N}^i + N^j D_j N^i + 2 N N^j K^i_j \\
&\quad + N^{-1} N^i N^k N^l K_{kl}, \\
\Gamma^i_{rj} &= -N^{-1} N^i \partial_j N + D_j N^i + N^{-1} N^i N^k K_{kj} + N K^i_j, \\
\Gamma^k_{ij} &= \Gamma^k_{ij}[\gamma] + N^{-1} N^k K_{ij}.
\end{aligned} \tag{191}$$

Hamilton–Jacobi Primer

In this appendix we collect a few essential facts about HJ theory in classical mechanics. For an in-depth account of HJ theory we refer the interested reader to [24, 40]. A more abstract exposition can be found in [41].

Let $\mathcal{Q}$ be the configuration space of a point particle described by the action[14]

$$S = \int^t dt' L(q, \dot{q}; t), \tag{192}$$

where q^α are coordinates on $\mathcal{Q}$. In the Hamiltonian formalism the generalized coordinates q^α and the canonical momenta

$$p_\alpha = \frac{\partial L}{\partial \dot{q}^\alpha}, \tag{193}$$

[14] In this appendix a dot $\dot{}$ denotes a derivative with respect to time t.

are independent variables parameterizing the phase space, which is isomorphic to the cotangent bundle $T^*\mathcal{Q}$ of the configuration space $\mathcal{Q}$. The cotangent bundle is a symplectic manifold with a canonical closed 2-form (symplectic form)

$$\Omega = dp_\alpha \wedge dq^\alpha. \tag{194}$$

Since Ω is closed, it can be *locally* expressed as

$$\Omega = d\Theta, \tag{195}$$

where

$$\Theta = p_\alpha dq^\alpha, \tag{196}$$

is known as the canonical 1-form, or pre-symplectic form. The Hamiltonian, given by the Legendre transform of the Lagrangian,

$$H(p, q; t) = p_\alpha \dot{q}^\alpha - L, \tag{197}$$

is a map $H : T^*\mathcal{Q} \longrightarrow \mathbb{R}$ and governs the time evolution of the dynamical system through Hamilton's equations

$$\dot{q}^\alpha = \frac{\partial H}{\partial p_\alpha}, \quad \dot{p}_\alpha = -\frac{\partial H}{\partial q^\alpha}. \tag{198}$$

At this point it is instructive to distinguish two cases, depending on whether the Hamiltonian depends explicitly on time t or not.

- **Time-independent systems**

A section, s, of the cotangent bundle is a map $s : \mathcal{Q} \longrightarrow T^*\mathcal{Q}$, providing a 1-form over each point $q \in \mathcal{Q}$. A closed section of $T^*\mathcal{Q}$ is locally exact and so it can be written as $s = d\mathcal{W}$ for some function $\mathcal{W}(q)$ on $\mathcal{Q}$. Under the isomorphism between phase space and the cotangent bundle this means that locally

$$p_\alpha = \frac{\partial \mathcal{W}(q)}{\partial q^\alpha}. \tag{199}$$

Moreover,

$$\Theta \circ s = d\mathcal{W}, \quad \Omega \circ s = 0. \tag{200}$$

These results hold for any closed section s of $T^*\mathcal{Q}$. The HJ theorem relates certain closed sections, s, of the cotangent bundle to solutions of Hamilton's equations (198). In particular,

$$d\,(H\circ s)=\left(\frac{\partial H}{\partial q^\alpha}+\frac{\partial^2\mathcal{W}}{\partial q^\beta\partial q^\alpha}\frac{\partial H}{\partial p_\beta}\right)dq^\alpha$$
$$=\left(\frac{\partial H}{\partial q^\alpha}+\dot{p}^\alpha\right)dq^\alpha+\frac{\partial^2\mathcal{W}}{\partial q^\beta\partial q^\alpha}\left(\frac{\partial H}{\partial p_\beta}-\dot{q}^\beta\right)dq^\alpha, \tag{201}$$

which implies that the following two statements are equivalent (see Theorem 2.1 in [42]):

(i) If $\sigma:\mathbb{R}\to\mathcal{Q}$ satisfies the first of Hamilton's equations in (198), then $s\circ\sigma$ satisfies the second Hamilton equation.
(ii) $d\,(H\circ s)=0$.

Hence, a closed section $s=d\mathcal{W}$ of the cotangent bundle that satisfies the (time-independent) HJ equation

$$H\circ s=H\left(\frac{\partial\mathcal{W}}{\partial q^\alpha},q^\beta\right)=E, \tag{202}$$

where E is some constant, provides a solution of Hamilton's equations.

- **Time-dependent systems**

In order to accommodate systems with a Hamiltonian that explicitly depends on time we extend the configuration space by including time as a generalized coordinate so that $\mathcal{Q}_{\text{ext}}=\mathcal{Q}\times\mathbb{R}$ is now the extended configuration space. Phase space is accordingly extended by including $-H$ as the canonical momentum conjugate to t. This extended phase space is isomorphic to the cotangent bundle $T^*\mathcal{Q}_{\text{ext}}$, which carries the canonical symplectic form

$$\Omega_{\text{ext}}=d\Theta_{\text{ext}}=dp_\alpha\wedge dq^\alpha-dH\wedge dt. \tag{203}$$

Moreover, to Hamilton's equations we can now append the equation

$$\dot{H}=\frac{\partial H}{\partial t}. \tag{204}$$

A closed section of $T^*\mathcal{Q}_{\text{ext}}$ can be locally written as $s=d\mathcal{S}$ for some function on $\mathcal{Q}_{\text{ext}}$, and consequently

$$p_\alpha=\frac{\partial\mathcal{S}(q;t)}{\partial q^\alpha},\quad -H=\frac{\partial\mathcal{S}(q;t)}{\partial t}, \tag{205}$$

which imply that

$$\Theta_{\text{ext}}\circ s=d\mathcal{S},\quad \Omega_{\text{ext}}\circ s=0. \tag{206}$$

It follows that

$$
\begin{aligned}
0 &= d\left(H\circ s + \frac{\partial \mathcal{S}}{\partial t}\right)\\
&= \left[\frac{\partial H}{\partial q^\alpha} + \dot q^\beta \frac{\partial^2 \mathcal{S}}{\partial q^\beta \partial q^\alpha} + \frac{\partial^2 \mathcal{S}}{\partial t \partial q^\alpha} + \frac{\partial^2 \mathcal{S}}{\partial q^\beta \partial q^\alpha}\left(\frac{\partial H}{\partial p_\beta} - \dot q^\beta\right)\right] dq^\alpha\\
&\quad + \left[\frac{\partial H}{\partial t} + \dot q^\alpha \frac{\partial^2 \mathcal{S}}{\partial t \partial q^\alpha} + \frac{\partial^2 \mathcal{S}}{\partial t^2} + \frac{\partial^2 \mathcal{S}}{\partial q^\beta \partial t}\left(\frac{\partial H}{\partial p_\beta} - \dot q^\beta\right)\right] dt\\
&= \left[\frac{\partial H}{\partial q^\alpha} + \dot p^\alpha + \frac{\partial^2 \mathcal{S}}{\partial q^\beta \partial q^\alpha}\left(\frac{\partial H}{\partial p_\beta} - \dot q^\beta\right)\right] dq^\alpha\\
&\quad + \left[\frac{\partial H}{\partial t} - \dot H + \frac{\partial^2 \mathcal{S}}{\partial q^\beta \partial t}\left(\frac{\partial H}{\partial p_\beta} - \dot q^\beta\right)\right] dt,
\end{aligned}
\tag{207}
$$

which allows us to generalize the HJ theorem to time-dependent Hamiltonians. Namely, a closed section $s = d\mathcal{S}$ of $T^*\mathcal{Q}_{\text{ext}}$ that satisfies the HJ equation

$$
H\circ s + \frac{\partial \mathcal{S}}{\partial t} = H\left(\frac{\partial \mathcal{S}}{\partial q^\alpha}, q^\beta; t\right) + \frac{\partial \mathcal{S}}{\partial t} = 0, \tag{208}
$$

provides a solution to Hamilton's equations.

A few comments are in order at this point. Firstly, note that the HJ formalism for time-dependent Hamiltonians reduces to that for time-independent Hamiltonians upon setting

$$
\mathcal{S}(q;t) = \mathcal{W}(q) - Et. \tag{209}
$$

The function $\mathcal{S}(q;t)$ is known as Hamilton's principal function, while $\mathcal{W}(q)$ is called the characteristic function. Secondly, the expressions (94) for the canonical momenta and the Hamiltonian should be familiar from quantum mechanics. Indeed, Hamilton's principal function $\mathcal{S}(q;t)$ is related to the WKB wavefunction by

$$
\psi_{WKB}(q;t) \sim e^{i\mathcal{S}(q;t)/\hbar}, \tag{210}
$$

and so the expressions (94) are respectively the coordinate representation of the momentum operator and the identification of the Hamiltonian with the time evolution operator.

Finally, Hamilton's principal function $\mathcal{S}(q;t)$, defined as a solution of the HJ equation (208), is closely related to the on-shell action. To elucidate the relation, consider the action (192) on the semi-infinite line $(-\infty, t]$. A general variation of the action (192) gives

$$
\delta S = \int^t dt'\left(\frac{\partial L}{\partial q^\alpha}\delta q^\alpha + \frac{\partial L}{\partial \dot q^\alpha}\delta \dot q^\alpha\right) = \int^t dt'\left(\frac{\partial L}{\partial q^\alpha} - \frac{d}{dt'}\left(\frac{\partial L}{\partial \dot q^\alpha}\right)\right)\delta q + \left.\frac{\partial L}{\partial \dot q^\alpha}\delta q^\alpha\right|_t. \tag{211}
$$

To ensure that the variational principle implies the equations of motion we need to impose the boundary condition $\delta q^\alpha = 0$ at $t' = t$. The on-shell action therefore becomes a function of the fixed but arbitrary boundary condition $q^\alpha(t)$, namely $S_{\text{on-shell}}(q;t)$, while

$$p_\alpha|_t = \left.\frac{\partial L}{\partial \dot{q}^\alpha}\right|_t = \frac{\partial S_{\text{on-shell}}}{\partial q^\alpha}. \tag{212}$$

Moreover,

$$\dot{S}_{\text{on-shell}} = L = \frac{\partial S_{\text{on-shell}}}{\partial t} + \frac{\partial S_{\text{on-shell}}}{\partial q^\alpha}\dot{q}^\alpha, \tag{213}$$

and so $S_{\text{on-shell}}$ satisfies the HJ equation (208):

$$0 = p_\alpha \dot{q}^\alpha - L + \frac{\partial S_{\text{on-shell}}}{\partial t} = H\left(\frac{\partial S_{\text{on-shell}}}{\partial q^\alpha}, q^\beta; t\right) + \frac{\partial S_{\text{on-shell}}}{\partial t}. \tag{214}$$

We therefore conclude that the on-shell action as a function of the arbitrary but fixed boundary condition $q(t)$, $S_{\text{on-shell}}(q;t)$, can be identified with Hamilton's principal function $\mathcal{S}(q;t)$. The fact that the on-shell action is a solution of the HJ equation is the fundamental reason for the critical role that HJ theory has in holographic renormalization.

References

1. J.M. Maldacena, The large N limit of superconformal field theories and supergravity. Int. J. Theor. Phys. **38**, 1113–1133 (1999). arXiv:hep-th/9711200 [Adv. Theor. Math. Phys. **2**, 231(1998)]
2. E. Witten, Anti-de Sitter space and holography. Adv. Theor. Math. Phys. **2**, 253–291 (1998). arXiv:hep-th/9802150
3. S. Gubser, I.R. Klebanov, A.M. Polyakov, Gauge theory correlators from noncritical string theory. Phys. Lett. B **428**, 105–114 (1998). arXiv:hep-th/9802109
4. M. Henningson, K. Skenderis, The holographic Weyl anomaly. J. High Energy Phys. **07**, 023 (1998). arXiv:hep-th/9806087
5. V. Balasubramanian, P. Kraus, A stress tensor for Anti-de Sitter gravity. Commun. Math. Phys. **208**, 413–428 (1999). arXiv:hep-th/9902121
6. J. de Boer, E.P. Verlinde, H.L. Verlinde, On the holographic renormalization group. J. High Energy Phys. **08**, 003 (2000). arXiv:hep-th/9912012
7. P. Kraus, F. Larsen, R. Siebelink, The gravitational action in asymptotically AdS and flat space-times. Nucl. Phys. B **563**, 259–278 (1999). arXiv:hep-th/9906127
8. S. de Haro, S.N. Solodukhin, K. Skenderis, Holographic reconstruction of space-time and renormalization in the AdS/CFT correspondence. Commun. Math. Phys. **217**, 595–622 (2001). arXiv:hep-th/0002230
9. M. Bianchi, D.Z. Freedman, K. Skenderis, How to go with an RG flow. J. High Energy Phys. **08**, 041 (2001). arXiv:hep-th/0105276
10. M. Bianchi, D.Z. Freedman, K. Skenderis, Holographic renormalization. Nucl. Phys. B **631**, 159–194 (2002). arXiv:hep-th/0112119

11. D. Martelli, W. Mueck, Holographic renormalization and ward identities with the Hamilton-Jacobi method. Nucl. Phys. B **654**, 248–276 (2003). arXiv:hep-th/0205061
12. K. Skenderis, Lecture notes on holographic renormalization. Class. Quantum Gravity **19**, 5849–5876 (2002). arXiv:hep-th/0209067
13. I. Papadimitriou, K. Skenderis, AdS/CFT correspondence and geometry, *AdS/CFT correspondence: Einstein metrics and their conformal boundaries. Proceedings, 73rd Meeting of Theoretical Physicists and Mathematicians, Strasbourg, France, 11–13 September 2003* (2004), pp. 73–101. arXiv:hep-th/0404176
14. O. Aharony, N. Seiberg, and Y. Tachikawa, *Reading between the lines of four-dimensional gauge theories*, *JHEP***08** (2013) 115, 1305.0318
15. H. Osborn, Weyl consistency conditions and a local renormalization group equation for general renormalizable field theories. Nucl. Phys. B **363**, 486–526 (1991)
16. B.P. Dolan, Symplectic geometry and Hamiltonian flow of the renormalization group equation. Int. J. Mod. Phys. A **10**, 2703–2732 (1995). arXiv:hep-th/9406061
17. I. Papadimitriou, Holographic renormalization as a canonical transformation. J. High Energy Phys. **11**, 014 (2010). arXiv:1007.4592
18. D.Z. Freedman, K. Johnson, J.I. Latorre, Differential regularization and renormalization: a new method of calculation in quantum field theory. Nucl. Phys. B **371**, 353–414 (1992)
19. I. Papadimitriou, Holographic renormalization made simple: an example. Subnucl. Ser. **41**, 508–514 (2005)
20. G.W. Gibbons, S.W. Hawking, Action integrals and partition functions in quantum gravity. Phys. Rev. D **15**, 2752–2756 (1977)
21. I. Papadimitriou, K. Skenderis, Thermodynamics of asymptotically locally AdS spacetimes. J. High Energy Phys. **08**, 004 (2005). arXiv:hep-th/0505190
22. R. Arnowitt, S. Deser, C.W. Misner, Canonical variables for general relativity. Phys. Rev. **117**, 1595–1602 (1960)
23. I. Papadimitriou, Multi-trace deformations in AdS/CFT: exploring the vacuum structure of the deformed CFT. J. High Energy Phys. **05**, 075 (2007). arXiv:hep-th/0703152
24. M. Henneaux, C. Teitelboim, *Quantization of Gauge Systems* (University of Princeton, Princeton, 1992)
25. B. C. van Rees, Holographic renormalization for irrelevant operators and multi-trace counterterms. Physics **42** (2011). arXiv:1102.2239
26. B.C. van Rees, Irrelevant deformations and the holographic Callan-Symanzik equation. J. High Energy Phys. **10**, 067 (2011). arXiv:1105.5396
27. K.A. Intriligator, Maximally supersymmetric RG flows and AdS duality. Nucl. Phys. B **580**, 99–120 (2000). arXiv:hep-th/9909082
28. K. Skenderis, M. Taylor, Kaluza-Klein holography. J High Energy Phys. **05**, 057 (2006). arXiv:hep-th/0603016
29. K. Skenderis, M. Taylor, Holographic Coulomb branch vevs. J. High Energy Phys. **08**, 001 (2006). arXiv:hep-th/0604169
30. Y. Korovin, K. Skenderis, M. Taylor, Lifshitz as a deformation of Anti-de Sitter. J. High Energy Phys. **08**, 026 (2013). arXiv:1304.7776
31. W. Chemissany, I. Papadimitriou, Lifshitz holography: the whole shebang. J. High Energy Phys. **01**, 052 (2015). arXiv:1408.0795
32. A. O'Bannon, I. Papadimitriou, J. Probst, *A Holographic Two-Impurity Kondo Model*. arXiv:1510.08123
33. S.F. Ross, Holography for asymptotically locally Lifshitz spacetimes. Class. Quantum Gravity **28**, 215019 (2011). arXiv:1107.4451
34. T. Griffin, P. Horava, C.M. Melby-Thompson, Conformal Lifshitz gravity from holography. J. High Energy Phys. **05**, 010 (2012). arXiv:1112.5660
35. I. Papadimitriou, Holographic renormalization of general dilaton-axion gravity. J High Energy Phys. **1108**, 119 (2011). arXiv:1106.4826
36. W. Chemissany, I. Papadimitriou, Generalized dilatation operator method for non-relativistic holography. Phys. Lett. B **737**, 272–276 (2014). arXiv:1405.3965

37. I. Kanitscheider, K. Skenderis, M. Taylor, Precision holography for non-conformal branes. J. High Energy Phys. **09**, 094 (2008). arXiv:0807.3324
38. P. Breitenlohner, D.Z. Freedman, Positive energy in anti-De Sitter backgrounds and gauged extended supergravity. Phys. Lett. B **115**, 197 (1982)
39. I. Papadimitriou, K. Skenderis, Correlation functions in holographic RG flows. J. High Energy Phys. **0410**, 075 (2004). arXiv:hep-th/0407071
40. H.J. Rothe, K.D. Rother, *Classical and Quantum Dynamics of Constrained Hamiltonian Systems* (World Scientific, Singapore, 2010)
41. R. Abraham, J. Marsden, *Foundations of Mechanics*, 2nd edn. (Benjamin-Cumming, Reading, 1978)
42. M. de Leon, J.C. Marrero, D.M. de Diego, A geometric Hamilton-Jacobi theory for classical field theories (2008). arXiv:0801.1181

Nonsingular Black Holes in Palatini Extensions of General Relativity

Gonzalo J. Olmo

Abstract An introduction to extended theories of gravity formulated in metric-affine (or Palatini) spaces is presented. Focusing on spherically symmetric configurations with electric fields, we will see that in these theories the central singularity present in General Relativity is generically replaced by a wormhole structure. The resulting space-time becomes geodesically complete and, therefore, can be regarded as non-singular. We illustrate these properties considering two different models, namely, a quadratic $f(R)$ theory and a Born-Infeld like gravity theory.

1 Introduction

Shortly after the publication of Einstein's equations for the gravitational field, Karl Schwarzschild found an exact solution describing the vacuum region surrounding a spherical body of mass M. The line element characterizing this space-time takes the form

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \frac{1}{\left(1 - \frac{2M}{r}\right)} dr^2 + r^2 d\Omega^2 \tag{1}$$

where r is the radial coordinate and $d\Omega^2 \equiv d\theta^2 + \sin\theta^2 d\varphi^2$ represents the spherical sector. Given the smallness of the quantity $r_S \equiv 2M$, which for a star like the sun is about $r_S \sim 3\,\text{km}$, and the limited astrophysical knowledge about compact objects at that time, this line element was thought to be physically meaningful only

G.J. Olmo (✉)
Departamento de Física Teórica and IFIC,
Centro Mixto Universidad de Valencia - CSIC, Paterna, Spain
e-mail: gonzalo.olmo@uv.es

G.J. Olmo
Universidad de Valencia, Burjassot, 46100 Valencia, Spain

G.J. Olmo
Departamento de Física, Universidade Federal da Paraíba,
João Pessoa, Paraíba 58051-900, Brazil

R. Kallosh and E. Orazi (eds.), *Theoretical Frontiers in Black Holes and Cosmology*, Springer Proceedings in Physics 176,
DOI 10.1007/978-3-319-31352-8_5

in the exterior regions of stars. With the discovery of neutron stars, the physical existence of ultra compact objects was reconsidered and in the 1960s it was understood that geometries such as Schwarzschild's could be a physical reality. In fact, using powerful mathematical techniques it was concluded that under reasonable conditions, complete gravitational collapse is unavoidable for sufficiently massive objects [1–5]. Black holes, therefore, are an important prediction of Einstein's theory of General Relativity (GR).

The existence of black holes has a deep impact for the theoretical consistency of GR. In fact, given that the laws of Physics as we know them are defined on top of a dynamical geometry, the space-time, if the geometry becomes ill defined at some event then our ability to describe physical phenomena and make predictions will be seriously affected [6]. This is precisely what happens in the interior of black holes.

In the Schwarzschild case, for instance, any observer within the region $r < r_S$ is forced to travel towards decreasing values of r, being $r = 0$ reached in a finite proper time [7]. At that location, curvature scalars diverge and gravitational forces are so strong that any extended body is instantaneously crushed to zero volume. Thus, any observer reaching $r = 0$ is destroyed and disappears together with its ability to describe the physical processes taking place in that region. Under this circumstance, it is typically stated that the Schwarzschild black hole contains a singularity or that it describes a singular space-time.

The notion of singularity is a very elusive concept, though [8]. The Schwarzschild example suggests that curvature divergences can somehow be regarded as a signature of their existence. However, if one takes a space-time such as Minkowski and artificially removes a portion of it, any observer or signal that propagates through it and reaches the boundary of the removed portion simply vanishes there, as there is nowhere to go *beyond* that boundary. One can also find observer trajectories which intersect this boundary in their past, suggesting that they came into existence out of the blue. The potential creation and/or destruction of physical observers and/or light signals in a given space-time is thus fundamental to determine if an appropriate physical description is possible or not. For this reason, for the characterization of singular space-times one should not focus on the potential existence of infinities in the gravitational fields, which are absent in the amputated Minkowskian example, but rather one should be worried about the existence of physical observers at all times.

Following this line of reasoning, it is generally stated that a singular space-time is one in which there exist incomplete timelike and/or null geodesics, i.e., geodesics which cannot be extended to arbitrary values of their affine parameter in the past or in the future [9–11] (see also [12] for a more recent discussion of this point and references). Note, in this sense, that observers are identified with geodesic curves. The incompleteness of geodesics, therefore, hinges in the fact that in order to be able to provide a reliable description of phenomena on a given space-time, physical observers and/or signals should never be created or destroyed, i.e., their existence should be unrestricted along their worldline. The presence of curvature divergences is thus irrelevant for the determination of whether a space-time is singular or not: the potential *suffering* of observers due to intense tidal forces is not comparable to the importance of their very existence.

The fact that the Schwarzschild solution, as well as all other black hole solutions known to date, represent geodesically incomplete space-times is thus a serious conceptual limitation of GR. Improvements in the theory are thus necessary, which has motivated different approaches to the problem of singularities. Some of those are based on the idea of bounded curvature scalars [13–18] which, however, is logically unrelated to the notion of geodesic completeness.

In these lectures we will be dealing with certain (classical) extensions of GR in which simple non-rotating black hole solutions which are geodesically complete, and hence nonsingular, are possible. The approach presented here does not follow the intuitive and widespread idea that to get a nonsingular theory one should keep curvature scalars bounded. In our case, curvature divergences do arise in some regions but their presence is not an obstacle to have complete geodesic paths[1] [19]. Making a long story short, this is accomplished by the replacement of the black hole center by a wormhole [20, 21]. Unlike the case of GR, in our approach one does not need exotic matter sources to generate the wormhole. Rather, a simple free electric field will be able to do the job. Also, our geometries are not designed a priori but, rather, follow directly by integrating the field equations once the matter fields are specified. It is in this sense that these wormholes are more natural than those typically discussed in the context of GR, where one first defines the metric and then obtains the necessary stress-energy tensor by plugging it in Einstein's equations.

It is worth noting at this point that the use of nontrivial topologies (wormholes) in combination with self-gravitating free fields as a way to cure space-time singularities was suggested long ago by J.A. Wheeler [22]. We will see that our solutions represent an explicit example of geons in Wheeler's sense [23, 24] and, as such, avoid the well-known *problem of the sources* [25] that one finds in GR for the Schwarzschild and Reissner-Nordström black holes, for instance.

The content is organized as follows. In Sect. 2 our geometrical scenario is introduced, making emphasis on the importance of understanding gravitation as a geometric phenomenon and geometry as an issue of metrics and connections, i.e., as something else than a theory of just metrics. Once the fundamental notions of metric-affine geometry have been presented, in Sect. 3 we work out the field equations of GR à la Palatini, and in Sect. 4 we do the same for two models of interest, namely, a quadratic $f(R)$ theory and a Born-Infeld-like gravity theory. The first example appears naturally in that quadratic corrections in curvature are common to many different approaches to quantum and non-quantum extensions of GR. The simplicity of this model comes at the price of introducing a nonlinear theory of electrodynamics as matter source in order to obtain the desired effects in the equations. The Born-Infeld case, on the contrary, can be easily combined with a standard Maxwell electric field. In both cases, exact analytical black hole solutions can be found, which allows us to explore the behavior of geodesics in both geometries in detail. The equations

[1]This provides a counterexample to the correlation typically observed in GR between space-times with incomplete geodesics and which contain curvature divergences.

governing black hole structure are derived in a generic form in Sect. 5 and applied to the gravitational Born-Infeld model in Sect. 6 and to the $f(R)$ model in Sect. 7. The study of geodesics appears in Sect. 8. We conclude in Sect. 9 with a brief summary and discussion of the results.

2 Basic Framework: Metric-Affine Gravity

In elementary courses on gravitation [7] one learns that general covariance is accomplished by replacing flat Minkowskian derivatives ∂_μ by covariant derivatives ∇_μ, whose action on vector components (for instance) is of the form $\nabla_\mu A_\nu = \partial_\mu A_\nu - \Gamma^\lambda_{\mu\nu} A_\lambda$. Here $\Gamma^\lambda_{\mu\nu}$ is the so-called Levi-Civita connection, which is defined as

$$\Gamma^\lambda_{\mu\nu} = \frac{g^{\lambda\rho}}{2}\left[\partial_\mu g_{\rho\nu} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}\right], \tag{2}$$

with $g_{\mu\nu}$ representing the space-time metric. The connection has a non-tensorial transformation law which compensates the action of ∂_μ in such a way that $\nabla_\mu A_\nu$ transforms as a tensor under arbitrary changes of coordinates. With the connection one defines the Riemann curvature tensor as

$$R^\alpha{}_{\beta\mu\nu} = \partial_\mu \Gamma^\alpha_{\nu\beta} - \partial_\nu \Gamma^\alpha_{\mu\beta} + \Gamma^\kappa_{\nu\beta}\Gamma^\alpha_{\mu\kappa} - \Gamma^\kappa_{\mu\beta}\Gamma^\alpha_{\nu\kappa}, \tag{3}$$

and Einstein's equations take the form

$$R_{\beta\nu} - \frac{1}{2} g_{\beta\nu} R = \kappa^2 T_{\beta\nu}, \tag{4}$$

where $R_{\beta\nu} = R^\lambda{}_{\beta\lambda\nu}$ is the Ricci tensor, $R = g^{\mu\nu} R_{\mu\nu}$ the Ricci curvature scalar, $T_{\beta\nu}$ the stress-energy tensor of the matter, and $\kappa^2 = 8\pi G/c^4$. Written in this form, GR is a theory based on the metric tensor $g_{\mu\nu}$ as the field that describes gravitational interactions.

Interestingly, at the time Einstein formulated GR, the theory of affine connections had not been developed yet. Only Riemannian geometry, based on the metric tensor, was available to implement his idea of gravitation as a geometric phenomenon. Einstein's theory boosted the interest of mathematicians on differential geometry, giving rise to the study of non-Riemannian spaces [26]. It was then established that general covariance could be implemented without defining a metric structure. This is so because the non-tensorial transformation law of the connection is a property that does not depend on the particular form of the connection, i.e., it is independent of the definition (2). As a consequence, the Riemann curvature tensor (3) can be defined without referring it to a metric.

This point is very important because it opens a whole new range of possibilities to implement the idea of gravitation as a geometric phenomenon. Is the space-time geometry Riemannian? It is rather apparent that the Euclidean space of Newtonian mechanics is not appropriate to describe relativistic phenomena, but that does not lead uniquely to the Riemannian case (*the metric as the foundation of all*). Whether the space-time geometry is Riemannian or not is a fundamental question that must be answered by experiments, as Einstein himself stated [27]. We must, obviously, admit that the Riemannian description of GR is very successful at the length scales and energies accessible in laboratory and the Solar system (as well as in other systems whose orbital motions are well understood) [28]. However, there is still a broad range of energies and length scales that lie beyond direct experimental scrutiny. Demanding that the Riemannian condition (2), or $\nabla_\mu g_{\alpha\beta} = 0$, be satisfied at all scales might be an excessive assumption/constraint.

Aside from the purely theoretical interest in non-Riemannian geometries, there are other reasons to explore the effects that independent metric and affine degrees of freedom could have in gravitation. It turns out that in continuous systems with an ordered microstructure, such as in Bravais crystals or materials as popular as graphene, one needs a metric-affine geometry in order to correctly describe macroscopic properties like viscosity or plasticity [29, 30]. These properties are intimately related with the existence of defects in the microstructure. And these defects are responsible for the independence between metric and affine degrees of freedom. For instance, in a crystal without defects, one can introduce a notion of distance (metricity) by counting atoms along crystallographic directions (a special set of directions in the structure which minimize distances) [29, 31–34]. However, if there exist point defects such as missing atoms, the microscopic process of step counting breaks down and the idea of metricity cannot be translated to the continuum in any natural way.

The microscopic notion of distance can be extended to the continuum by defining an *auxiliary or idealized* structure without defects in which the step-counting procedure is naturally implemented. Physical distances can be defined once the density of defects is known, which allows to establish a correspondence between the idealized structure and the physical one. The idealized crystallographic directions need not coincide everywhere with the directions that minimize physical distances, which implies that the physical metric $g_{\alpha\beta}$ is not conserved along the idealized paths, i.e., $\nabla_\mu^{(\Gamma)} g_{\alpha\beta} \neq 0$, where Γ is the connection associated to the auxiliary metric. The quantity $Q_{\mu\alpha\beta} \equiv \nabla_\mu^{(\Gamma)} g_{\alpha\beta}$, known as non-metricity tensor, then plays a relevant role in the physical description of the continuized system.

Another interesting geometric structure arises when there exist dislocations (one-dimensional defects). It is well-known that dislocations are the discrete version of torsion [35, 36]. Crystals with a certain density of dislocations, therefore, lead to effective geometries with a metric and a non-symmetric connection, which is related to the Einstein–Cartan theory of gravity [30]. Given that point defects (vacancies and interstitial) can interact with dislocations (creating and/or destroying them), a complete theory should have into account the metric, the non-metricity tensor, and the torsion. If the space-time had a microstructure with defects, such as that suggested by the notion of space-time foam, the continuum that we perceive could require geometric structures beyond those typically considered in Einstein's theory of gravity [37–39].

It is for the above simple reasons that we are going to explore several examples of theories of gravity assuming that metric and connection are equally fundamental and a priori independent fields. Imposing a *principle of democracy*, we will derive the equations governing the metric and the connection from an action, without imposing any a priori constraint between them. The field equations should determine how metric and affine degrees of freedom interact between them and with the matter fields.

3 General Relativity à la Palatini

To begin with, it is useful to consider the metric-affine or Palatini version of GR [40]. The action functional for the Einstein–Palatini theory can be written as

$$S = \frac{1}{2\kappa^2}\int d^4x\sqrt{-g}g^{\mu\nu}R_{\mu\nu}(\Gamma) + S_m(g_{\mu\nu},\psi), \tag{5}$$

where $R_{\mu\nu}(\Gamma) = R^{\alpha}{}_{\mu\alpha\nu}$ is defined in terms of a connection which is a priori independent of the metric $g_{\mu\nu}$, S_m represents the matter action, and ψ denotes collectively the matter fields.[2]

Variation of the action with respect to the (inverse) metric and the connection leads to

$$\delta S = \frac{1}{2\kappa^2}\int d^4x\sqrt{-g}\left[\left(R_{\mu\nu}(\Gamma) - \frac{1}{2}g_{\mu\nu}g^{\alpha\beta}R_{\alpha\beta}(\Gamma) - \kappa^2 T_{\mu\nu}\right)\delta g^{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu}\right], \tag{6}$$

where

$$\delta R_{\mu\nu} = \nabla_\lambda\left(\delta\Gamma^\lambda_{\nu\mu}\right) - \nabla_\nu\left(\delta\Gamma^\lambda_{\lambda\mu}\right) + 2S^\rho_{\alpha\nu}\delta\Gamma^\alpha_{\rho\mu}, \tag{7}$$

and $S^\rho_{\alpha\nu} \equiv \frac{1}{2}\left(\Gamma^\rho_{\alpha\nu} - \Gamma^\rho_{\nu\alpha}\right)$ is the torsion tensor. For simplicity, in the following derivations we will skip all torsional terms.[3] After elementary manipulations, and knowing that $\nabla_\mu(\sqrt{-g}J^\mu) = \partial_\mu(\sqrt{-g}J^\mu) + 2S^\lambda_{\lambda\mu}(\sqrt{-g}J^\mu)$, (6) turns into

---

[2]For simplicity, in the matter action we have only assumed a dependence on the metric. This prescription is compatible with the experimental evidence on the Einstein equivalence principle [28]. However, dependence on the connection should also be allowed to explore its phenomenology in regimes not yet accessed experimentally. The coupling of fermions to gravity, whose spin may source the torsion tensor (antisymmetric part of the connection), is a particular case of interest which has been considered explicitly in supergravity theories and in the Einstein–Cartan theory [25], for example.

[3]We do this to focus our attention on the symmetric part of the connection but we do admit the possibility of having an antisymmetric part because fermions do exist in Nature. Note in this sense that, in general, assuming a symmetric connection before performing the variations or setting it to zero after the field equations have been obtained are inequivalent procedures. A detailed discussion with concrete examples can be found in [41].

$$\delta S = \frac{1}{2\kappa^2}\int d^4x\sqrt{-g}\left[\left(R_{\mu\nu}(\Gamma) - \frac{1}{2}g_{\mu\nu}g^{\alpha\beta}R_{\alpha\beta}(\Gamma) - \kappa^2 T_{\mu\nu}\right)\delta g^{\mu\nu} + \left(-\nabla_\lambda\left(\sqrt{-g}g^{\mu\nu}\right) + \delta^\mu_\lambda\nabla_\rho\left(\sqrt{-g}g^{\rho\nu}\right)\right)\delta\Gamma^\lambda_{\mu\nu}\right]. \quad (8)$$

The field equations are obtained by setting to zero the coefficients multiplying the independent variations $\delta g^{\mu\nu}$ and $\delta\Gamma^\lambda_{\mu\nu}$, which yields

$$R_{\mu\nu}(\Gamma) - \frac{1}{2}g_{\mu\nu}g^{\alpha\beta}R_{\alpha\beta}(\Gamma) = \kappa^2 T_{\mu\nu} \quad (9)$$

$$-\nabla_\lambda\left(\sqrt{-g}g^{\mu\nu}\right) + \delta^\mu_\lambda\nabla_\rho\left(\sqrt{-g}g^{\rho\nu}\right) = 0. \quad (10)$$

Contracting the indices μ and λ in (10) one finds that $\nabla_\rho\left(\sqrt{-g}g^{\rho\nu}\right) = 0$, which turns that equation into

$$\nabla_\lambda\left(\sqrt{-g}g^{\mu\nu}\right) = 0, \quad (11)$$

writting this equation explicitly, we get

$$g^{\mu\nu}\partial_\lambda\sqrt{-g} + \sqrt{-g}\partial_\lambda g^{\mu\nu} + \sqrt{-g}\left[-\Gamma^\alpha_{\alpha\lambda}g^{\mu\nu} + \Gamma^\mu_{\lambda\alpha}g^{\alpha\nu} + \Gamma^\nu_{\lambda\alpha}g^{\alpha\mu}\right] = 0, \quad (12)$$

and contracting with $g_{\mu\nu}$ we find that $\Gamma^\alpha_{\alpha\mu} = \partial_\mu \ln\sqrt{-g}$, where the relation $g_{\mu\nu}\partial_\lambda g^{\mu\nu} = -2\partial_\lambda \ln\sqrt{-g}$ has been used. Inserting this result in (12), one finds that (11) is equivalent to $\nabla_\lambda g^{\mu\nu} = 0$. Given that $g_{\mu\rho}g^{\rho\nu} = \delta^\nu_\mu$, one readily verifies that $\nabla_\lambda g^{\mu\nu} = 0$ also implies $\nabla_\lambda g_{\mu\nu} = 0$. This last relation can be used to obtain the form of $\Gamma^\alpha_{\mu\nu}$ as a function of the metric and its first derivatives by just using algebraic manipulations [42]. The result is simply that $\Gamma^\alpha_{\mu\nu}$ boils down to the Levi-Civita connection defined in (2). As a consequence, the Ricci tensor $R_{\mu\nu}(\Gamma)$ turns into the Ricci tensor of the metric $g_{\mu\nu}$ and (9) coincides with the Einstein equations (4).

In summary, the Einstein–Palatini action exactly recovers Einstein's equations (in the torsionless case) and implies that the geometry is Riemannian without the need of imposing the *compatibility condition* $\nabla_\lambda g_{\mu\nu} = 0$ as an input.

It is important to remark at this point that the constraint $\nabla_\lambda g_{\mu\nu} = 0$ between metric and connection is a property that belongs naturally to the Einstein–Palatini theory but which is not a priori guaranteed in other theories. Nonetheless, in most of the literature on extended theories of gravity it has been implicitly assumed as true, forcing the geometry to be Riemannian from the onset (see, however, [43] for a review on Palatini gravity). We will see in the following that relaxing this constraint and allowing the theory to determine the form of the connection from a variational principle, the compatibility between metric and connection is generically lost. The implications of this will be nontrivial, providing new phenomenology that will be relevant in the study of black hole interiors.

4 Beyond GR

Considering extensions of GR to address questions concerning high and very high energies one naturally finds the possibility of adding quadratic and/or higher order curvature corrections in the gravitational Lagrangian. Such corrections arise when one considers quantum fields propagating in curved space-times [44, 45], in the low-energy limits of string theories [25], and in effective field theory or phenomenological approaches [46, 47]. Theories such as $R + \lambda R^2 + \gamma R_{\mu\nu} R^{\mu\nu} + \beta R^{\alpha}{}_{\beta\mu\nu} R_{\alpha}{}^{\beta\mu\nu}$, for instance, have been typically considered in the literature on the early universe and in black hole scenarios [48–59]. The Riemann-squared dependence is typically removed because it can be combined with the other quadratic terms to give the so-called Gauss-Bonnet term, which does not contribute to the field equations and simply redefines the coefficients λ and γ.

The standard argument is that high-order curvature corrections could capture some relevant new physics beyond the range of applicability of GR but below the full quantum gravity regime. Given the higher-order character of the resulting field equations, analytical solutions are hard to find in general. Numerical solutions do exist and regular cases (in the sense of bounded curvature scalars [14]) have been found for static black hole configurations [60] coupled to nonlinear theories of electrodynamics using perturbative methods.

The extensive literature existing on the metric (or Riemannian) formulation of quadratic gravity contrasts with the little attention received by its metric-affine counterpart. Interestingly, through recent work carried out in the last years, it has been established that in the Palatini version of those theories one always finds analytical solutions [61–63]. In the following we will study the field equations of models similar to the quadratic theory mentioned above but formulated in the Palatini approach. We will then focus on spherically symmetric configurations in which new black hole solutions can be found.

4.1 $f(R)$ Theories

The derivation of the field equations for theories of the $f(\mathcal{R})$ type, where f represents a certain function of the Ricci scalar[4] $\mathcal{R} = g^{\mu\nu} R_{\mu\nu}(\Gamma)$, is straightforward and follows essentially the same steps as in the case of GR presented in Sect. 3. Variation of the action leads to the equations (see, for instance, [42, 43] for details)

[4]The typography $\mathcal{R}$ is used here to emphasize that this scalar is built by combining the metric $g_{\mu\nu}$ with the Ricci tensor of a connection $\Gamma^{\alpha}_{\mu\nu}$ whose relation with $g_{\mu\nu}$ is a priori unknown. Whenever $\Gamma^{\alpha}_{\mu\nu}$ be defined in terms of a metric $k_{\mu\nu}$, then we will use the notation $R(k) = k^{\mu\nu} R_{\mu\nu}(k)$.

$$f_{\mathcal{R}} R_{\mu\nu}(\Gamma) - \frac{1}{2} g_{\mu\nu} f(\mathcal{R}) = \kappa^2 T_{\mu\nu} \tag{13}$$

$$-\nabla_\lambda \left(\sqrt{-g} f_{\mathcal{R}} g^{\mu\nu}\right) + \delta^\mu_\lambda \nabla_\rho \left(\sqrt{-g} f_{\mathcal{R}} g^{\rho\nu}\right) = 0, \tag{14}$$

where we denote $f_{\mathcal{R}} \equiv df/d\mathcal{R}$. Manipulating the connection equation (14), one finds that it can be reduced to

$$\nabla_\lambda \left(\sqrt{-g} f_{\mathcal{R}} g^{\mu\nu}\right) = 0. \tag{15}$$

Before proceeding with further manipulations, it is important to interpret this equation in combination with (13). At first sight, one may think that (15) contains up to second order derivatives of the connection because $f_{\mathcal{R}}$ is being acted upon by a derivative operator and it already contains first-order derivatives of $\Gamma^\alpha_{\mu\nu}$ via its dependence on $\mathcal{R}$. However, taking the trace of (13) with $g^{\mu\nu}$, one finds the important relation

$$\mathcal{R} f_{\mathcal{R}} - 2f = \kappa^2 T, \tag{16}$$

which establishes an algebraic relation between $\mathcal{R}$ and T, generalizing in this way the case $\mathcal{R} = -\kappa^2 T$ to nonlinear Lagrangians. This allows us to reinterpret (15) as an equation in which the independent connection $\Gamma^\alpha_{\mu\nu}$ satisfies an algebraic linear equation which involves the matter fields through the function $f_{\mathcal{R}}$ and the metric.

A solution to this equation can be obtained [64] by considering the existence of a rank-two tensor $h_{\mu\nu}$ such that $\sqrt{-g} f_{\mathcal{R}} g^{\mu\nu}$ can be written as $\sqrt{-h} h^{\mu\nu}$. With this identification, (15) turns into $\nabla_\mu(\sqrt{-h} h^{\alpha\beta}) = 0$, with $h_{\mu\nu} = f_{\mathcal{R}} g_{\mu\nu}$, and the solution can be obtained in much the same way as in the GR case (see the manipulations following (11)). As a result, we find that $\Gamma^\alpha_{\mu\nu}$ can be written as the Levi-Civita connection of the *auxiliary metric* $h_{\mu\nu}$, i.e.,

$$\Gamma^\lambda_{\mu\nu} = \frac{h^{\lambda\rho}}{2} \left[\partial_\mu h_{\rho\nu} + \partial_\nu h_{\rho\mu} - \partial_\rho h_{\mu\nu}\right]. \tag{17}$$

This result is valid for any Palatini theory of the $f(\mathcal{R})$ type, including GR.

We now turn our attention to the metric field equations (13), which contains elements referred to the metric $g_{\mu\nu}$ and others, like $R_{\mu\nu}(\Gamma)$, that depend on $h_{\mu\nu}$. Given that $g_{\mu\nu} = (1/f_{\mathcal{R}}) h_{\mu\nu}$ are conformally related, one can express $R_{\mu\nu}(\Gamma)$ in terms of $R_{\mu\nu}(g)$ and derivatives of $f_{\mathcal{R}}$ using well-known formulas [65, 66] (see, for instance, Appendix D in Wald's book [11]). Another possibility is to express everything in terms of $h_{\mu\nu}$. This is the approach we will follow because it leads to a very compact expression of the form

$${R^\mu}_\nu(h) = \frac{\kappa^2}{f_{\mathcal{R}}^2} \left[\frac{f}{2\kappa^2} {\delta^\mu}_\nu + {T^\mu}_\nu\right], \tag{18}$$

where ${R^\mu}_\nu(h) = h^{\mu\lambda} R_{\lambda\nu}(h)$ and ${T^\mu}_\nu = g^{\mu\lambda} T_{\lambda\nu}$. Written in this form, it is apparent that the auxiliary metric $h_{\mu\nu}$ satisfies a set of second-order equations with a structure very similar to that found in GR. In fact, on the left-hand side we find a second-order differential operator acting on $h_{\mu\nu}$, whereas on the right-hand side we have the matter, represented by ${T_\mu}^\nu$ and by f and $f_\mathcal{R}$, which are both functions of the trace T of ${T_\mu}^\nu$.

With the equations written in this form, one may try to solve for $h_{\mu\nu}$ and then obtain $g_{\mu\nu}$ by just using the conformal relation $g_{\mu\nu} = (1/f_\mathcal{R})h_{\mu\nu}$. This strategy might not always be straightforward, but will be very useful in the cases we will be dealing with.

To conclude with the discussion of $f(R)$ theories, it is important to consider the vacuum solutions. Such solutions correspond to the case in which ${T_\mu}^\nu = 0$, which implies $T = 0$. As a result, the algebraic equation (16) implies $\mathcal{R} = \mathcal{R}_{vac}$, where $\mathcal{R}_{vac}$ is some constant which may depend on the parameters that characterize the specific $f(\mathcal{R})$ Lagrangian chosen (obviously, some models may yield more than one solution and the good ones should be selected on physically reasonable grounds). A constant $\mathcal{R}$ implies that any function of $\mathcal{R}$ is also a constant. A direct consequence of this is that the conformal factor relating $g_{\mu\nu}$ and $h_{\mu\nu}$ can be absorbed into an irrelevant redefinition of units, making the two metrics coincide. This means that in vacuum the connection (17) boils down to the Levi-Civita connection of $g_{\mu\nu}$. Also, the metric field equations (18) recover the equations of GR in vacuum, with an effective cosmological constant. All this implies that the vacuum solutions of the theory are exactly the same as those appearing in vacuum GR (although different boundary conditions may apply). Therefore, in order to explore new physics beyond GR, one must consider explicitly the presence of matter sources. In this sense, we note that though the Schwarzschild solution is a mathematically acceptable solution of all Palatini $f(R)$ theories in vacuum, one should carefully consider the boundary conditions necessary to match that solution with the solution in the region containing the sources. The intuitive view that a delta-like distribution at the center is valid is not guaranteed here, as some models exhibit upper bounds for the density and pressure [64, 67]. For this reason, vacuum solutions must be handled with care, and non-vacuum solutions should be explored to gain insight on the properties of these theories.

4.2 Born-Infeld Gravity

The Born-Infeld gravity model is defined by means of the following action

$$S = \frac{1}{\kappa^2 \varepsilon} \int d^4x \left[\sqrt{-|g_{\mu\nu} + \varepsilon R_{\mu\nu}|} - \lambda\sqrt{-|g_{\mu\nu}|}\right] + S_m[g_{\mu\nu}, \psi], \qquad (19)$$

where vertical bars inside the square-root denote the determinant of that quantity, and ε is a small parameter with dimensions of length squared. This model was first consider in metric formalism [68], where the model suffers from a ghost instability due to its nonlinear dependence on the Ricci tensor. In [69], the theory was studied within the Palatini formalism, finding that in that approach the ghost is avoided. The phenomenological consequences of this theory have since then been extensively explored in cosmology [70–80], astrophysics [81, 82], stellar structure [83–90], the problem of cosmic singularities [91, 92], black holes [93, 94], and wormhole physics [95–98], among many others. Extensions of the original formulation have also been considered [99–114].

In the limit $\varepsilon \to 0$, this action recovers the quadratic[5] gravity theory mentioned at the beginning of this section with specific coefficients in front of R^2 and $R_{\mu\nu}R^{\mu\nu}$ [94]. The parameter λ is related to the cosmological constant, which vanishes if $\lambda = 1$. From now on we will set $\lambda = 1$ for simplicity. Higher-order contractions of the Ricci tensor arise as higher-order corrections in ε are considered.

The derivation of the field equations is straightforward if one introduces the definition

$$h_{\mu\nu} = g_{\mu\nu} + \varepsilon R_{\mu\nu}, \tag{20}$$

which allows to express the action (19) in the more compact form

$$S = \frac{1}{\kappa^2 \varepsilon} \int d^4x \left[\sqrt{-h} - \sqrt{-g}\right] + S_m[g_{\mu\nu}, \psi]. \tag{21}$$

Variation of the action with respect to metric and connection [94, 99] leads to

$$\sqrt{-h}h^{\mu\nu} - \sqrt{-g}g^{\mu\nu} = -\varepsilon\sqrt{-g}\kappa^2 T^{\mu\nu} \tag{22}$$

$$\nabla_\mu(\sqrt{-h}h^{\alpha\beta}) = 0 \tag{23}$$

It is clear from (23) that one can formally solve for the connection as the Levi-Civita connection of the *auxiliary metric* $h_{\mu\nu}$. Accepting that possibility, then we find that on the right-hand side of our original definition (20) the Ricci tensor contains up to second-order derivatives of $h_{\mu\nu}$. This simply indicates that to obtain $h_{\mu\nu}$ we need to solve some differential equations which involve $g_{\mu\nu}$ and $R_{\mu\nu}(h)$. In order to be able to do it, we must first find the relation that exists between $h_{\mu\nu}$ and the pair $(g_{\mu\nu}, T_{\mu\nu})$. This relation is determined by (22). In fact, assuming that $h_{\mu\nu}$ and $g_{\mu\nu}$ are related by some *deformation* matrix in the form

$$h_{\mu\nu} = g_{\mu\alpha}\Omega^\alpha{}_\nu, \;\; h^{\mu\nu} = (\Omega^{-1})^\mu{}_\alpha g^{\alpha\nu}, \tag{24}$$

then we can write (22) as

$$\sqrt{|\Omega|}(\Omega^{-1})^\mu{}_\nu = \delta^\mu{}_\nu - \varepsilon\kappa^2 T^\mu{}_\nu. \tag{25}$$

[5]As mentioned before, in the quadratic theory the dependence on the Riemann squared term can be eliminated by a simple redefinition of the coefficients in front of R^2 and $R_{\mu\nu}R^{\mu\nu}$. It is this Ricci-dependent theory which is recovered from the Born-Infeld action. We also note that the Ricci tensor in the action is symmetric. Though this is not obvious *a priori*, it can be shown that it is indeed true when torsion is set to zero at the level of the field equations [41].

This equation tells us that the deformation that relates $h_{\mu\nu}$ with $g_{\mu\nu}$ is determined by the local distribution of energy-momentum. This is similar to what we already observed in the case of $f(R)$ theories, where the conformal factor relating the metrics was a function of the trace of $T_\mu{}^\nu$ (see (16)). Note also that for this model the explicit form of $\Omega^\alpha{}_\nu$ is

$$\Omega^\alpha{}_\nu = \delta^\alpha{}_\nu + \varepsilon g^{\alpha\beta} R_{\beta\nu}(h). \tag{26}$$

Equation (25) is thus telling us that the object $g^{\alpha\beta} R_{\beta\nu}(h)$, which is a hybrid tensor that mixes $g^{\alpha\beta}$ with $h_{\mu\nu}$, is an algebraic function of the stress-energy tensor $T^\mu{}_\nu$. This is analogous to the relation between the scalar quantities $\mathcal{R}$ and T in the $f(\mathcal{R})$ case.

Having established the explicit relation between $h_{\mu\nu}$ and $g_{\mu\nu}$, we can now go back to (20) and write an equation for $h_{\mu\nu}$ and the matter. With a bit of algebra, one finds that the corresponding equations can be written as

$$R^\mu{}_\nu(h) = \frac{\kappa^2}{\sqrt{|\Omega|}} \left[\frac{\sqrt{|\Omega|} - \lambda}{\kappa^2 \varepsilon} \delta^\mu{}_\nu + T^\mu{}_\nu \right]. \tag{27}$$

The structure of these equations is very similar to that found in the case of $f(\mathcal{R})$ theories, with the Ricci tensor of the metric $h_{\mu\nu}$ on the left-hand side and functions of the matter fields on the right. We will see that in some cases of interest it will be possible to solve for $h_{\mu\nu}$ and then use (24) to obtain $g_{\mu\nu}$.

We also note here that the vacuum solutions of this model recover the field equations of vacuum GR. This is clearly seen from (20), which in vacuum implies that the matrix $\Omega_\mu{}^\alpha$ is a constant times the identity (when $\lambda = 1$, this constant is just unity). As a result the two metrics are physically equivalent and one recovers the equations of vacuum GR. The exploration of new physics should thus be carried out considering explicitly the presence of matter sources.

4.3 Generic Field Equations

The field equations obtained in the previous subsections for two different types of gravity models suggests that there exists a basic structure for the field equations in Palatini theories. This similarity is even more transparent when one realizes that the gravity Lagrangian in the case of $f(\mathcal{R})$ theories is $\mathcal{L}_G = f(\mathcal{R})/2\kappa^2$ and in the Born-Infeld case, $\mathcal{L}_G = \frac{\sqrt{|\Omega|}-\lambda}{\kappa^2\varepsilon}$. Moreover, in the $f(\mathcal{R})$ theories, the conformal relation between the metrics can be seen as a particular case in which $\Omega_\mu{}^\nu = f_\mathcal{R} \delta_\mu{}^\nu$. This allows us to express the field equations in the generic form

$$R^\mu{}_\nu(h) = \frac{\kappa^2}{\sqrt{|\Omega|}} \left[\mathcal{L}_G \delta^\mu{}_\nu + T^\mu{}_\nu \right], \tag{28}$$

with $\Omega_\mu{}^\nu$ representing the relations (24), and the explicit dependence of $\Omega^\mu{}_\nu$ with the matter fields determined by the field equations of the specific theory. With formal manipulations, it is possible to show that this representation of the field equations in terms of the auxiliary metric $h_{\mu\nu}$ is indeed correct for large families of theories of gravity in which $\mathcal{L}_G$ is just a functional of the inverse metric $g^{\mu\nu}$ and the Ricci tensor of an independent connection [113, 115] (when torsion is set to zero at the end of the variation). In vacuum configurations, the field equations recover GR plus an effective cosmological constant.

For convenience, we will use the generic equations (28) to obtain formal expressions for the solutions of static, spherically symmetric configurations in which the stress-energy tensor possesses certain algebraic properties. These formal expressions will then be particularized to specific gravity plus matter models.

5 Static, Spherically Symmetric Solutions

In this section we will be concerned with stress-energy tensors with a specific algebraic structure, namely

$$T^\mu{}_\nu = \begin{pmatrix} T_+ \hat{I}_{2\times 2} & \hat{O} \\ \hat{O} & T_- \hat{I}_{2\times 2} \end{pmatrix}, \tag{29}$$

where $T_\pm$ are some functions of the space-time coordinates, $\hat{I}_{2\times 2}$ is the 2×2 identity matrix, and $\hat{O}$ is the 2×2 zero matrix. Examples of stress-energy tensors with this structure arise in the case of electric fields and also for certain anisotropic fluids. The extension to higher-dimensions is straightforward using similar notation (see for instance [115, 116]). Given that the deformation matrix $\Omega^\mu{}_\nu$ will be determined by the stress-energy tensor, we may assume that it also has a similar algebraic structure, i.e., we can take

$$\Omega^\mu{}_\nu = \begin{pmatrix} \Omega_+ \hat{I}_{2\times 2} & \hat{O} \\ \hat{O} & \Omega_- \hat{I}_{2\times 2} \end{pmatrix}, \tag{30}$$

where $\Omega_\pm$ are given functions that should be provided by the field equations of the specific model considered. This point has been verified in several models explicitly and, therefore, appears as a reasonable assumption to proceed in a formal manner.

With the above assumptions, we find that the field equations (28) become

$$R^\mu{}_\nu(h) = \frac{\kappa^2}{\sqrt{|\Omega|}} \begin{pmatrix} (\mathcal{L}_G + T_+)\hat{I}_{2\times 2} & \hat{O} \\ \hat{O} & (\mathcal{L}_G + T_-)\hat{I}_{2\times 2} \end{pmatrix}. \tag{31}$$

Now we need to focus on the form of the left-hand side to proceed further. For static, spherically symmetric configurations, we can take the line element of the space-time metric $g_{\mu\nu}$ as

$$ds^2 = g_{ab}(x^0, x^1)dx^a dx^b + r^2(x^0, x^1)(d\theta^2 + \sin^2\theta d\varphi^2), \tag{32}$$

where (x^0, x^1) represent the coordinates of the 2×2 sector orthogonal to the 2-spheres. Analogously, one can define a line element for the auxiliary metric $h_{\mu\nu}$ of the form

$$d\tilde{s}^2 = h_{ab}(x^0, x^1)dx^a dx^b + \tilde{r}^2(x^0, x^1)(d\theta^2 + \sin^2\theta d\varphi^2). \tag{33}$$

Using the generic relations (24) between $h_{\mu\nu}$ and $g_{\mu\nu}$ together with (30), one finds that

$$h_{ab} = \Omega_+ g_{ab} \tag{34}$$
$$\tilde{r}^2 = \Omega_- r^2. \tag{35}$$

For static configurations, we further specify the form of $h_{\mu\nu}$ as follows:

$$d\tilde{s}^2 = -A(x)e^{2\Phi(x)}dt^2 + \frac{1}{A(x)}dx^2 + \tilde{r}^2(x)(d\theta^2 + \sin^2\theta d\varphi^2). \tag{36}$$

Computing the Ricci tensor associated to this line element, one finds the following relations:

$$R_t{}^t(h) = R_x{}^x(h) + \frac{4}{\tilde{r}}(\tilde{r}_{xx} - \Phi_x \tilde{r}_x) \tag{37}$$
$$R_\theta{}^\theta(h) = \frac{1}{\tilde{r}^2}\left[1 - A\tilde{r}_x^2 - \tilde{r}A\left(\tilde{r}_{xx} + \tilde{r}_x\left\{\frac{A_x}{A} + \Phi_x\right\}\right)\right]. \tag{38}$$

Given that the right-hand side of (31) implies that $R_t{}^t = R_x{}^x$, it follows that $(\tilde{r}_{xx} - \Phi_x \tilde{r}_x) = 0$. This equation allows us to take $\Phi(x) \to 0$ and $\tilde{r} \to x$, without loss of generality, and write the line element (36) in the form

$$d\tilde{s}^2 = -A(x)dt^2 + \frac{1}{A(x)}dx^2 + x^2(d\theta^2 + \sin^2\theta d\varphi^2). \tag{39}$$

As a result, $R_\theta{}^\theta$ gets simplified as

$$R_\theta{}^\theta(h) = \frac{1}{x^2}(1 - A - xA_x). \tag{40}$$

It is now useful to insert the Ansatz

$$A(x) = 1 - \frac{2M(x)}{x}, \tag{41}$$

which in combination with the right-hand side of (31) leads to the general expression

$$\frac{2M_x}{x^2} = \frac{\kappa^2}{\sqrt{|\Omega|}}(\mathcal{L}_G + T_-). \tag{42}$$

Given that we are dealing with a static, spherically symmetric space-time, the functions appearing in the right-hand side of this equation are just functions of x (or of $r(x)$). Therefore, by integrating this first-order equation, the geometry will be completely determined. In practice, however, one still needs to find the explicit relation between the area functions $r^2(x)$ and x^2, which is specified by (35). Recall, in this sense, that $\tilde{r}(x) \equiv x$ implies that $x^2 = \Omega_- r^2$ and that, in general, Ω_- will be a function of r. This point will become clear when we consider explicit examples.

In the examples that we will consider below, the functions $\Omega_\pm$ depend on x via $r(x)$. For this reason, it is convenient to express (42) in terms of the derivative with respect to r. This is immediate by just noting that $x^2 = \Omega_- r^2$ implies

$$\frac{dr}{dx} = \frac{1}{\Omega_-^{1/2}\left[1 + \frac{1}{2}\frac{\Omega_{-,r}}{\Omega_-}\right]}. \tag{43}$$

The resulting expression for M_r is thus

$$M_r = \frac{\kappa^2 \Omega_-^{1/2}}{2\Omega_+}(\mathcal{L}_G + T_-)r^2\left[1 + \frac{r}{2}\frac{\Omega_{-,r}}{\Omega_-}\right]. \tag{44}$$

By integrating this equation, the space-time line element (defined by the metric $g_{\mu\nu}$) becomes

$$ds^2 = -\frac{A(x)}{\Omega_+}dt^2 + \frac{1}{A(x)\Omega_+}dx^2 + r^2(x)(d\theta^2 + \sin^2\theta d\varphi^2). \tag{45}$$

In the next two sections we consider explicit examples that give concrete form to the above formulas.

6 Solutions in Born-Infeld Gravity

Let us consider the coupling of the Born-Infeld gravity model to a spherically symmetric, static electric field defined by the action $S_M = -\frac{1}{16\pi}\int d^4x\sqrt{-g}F_{\mu\nu}F^{\mu\nu}$, being $F_{\mu\nu}$ the electromagnetic field strength tensor. For this matter source, the stress energy tensor can be written as

$$T^\mu{}_\nu = \frac{q^2}{8\pi r^4}\begin{pmatrix} -\hat{I}_{2\times 2} & \hat{O} \\ \hat{O} & +\hat{I}_{2\times 2}\end{pmatrix}, \tag{46}$$

where q represents the electric charge. Inserting this expression in (20), one finds that the components of $\Omega^{\mu}{}_{\nu}$ are just

$$\Omega_{\pm} = 1 \mp \frac{\varepsilon\kappa^2 q^2}{8\pi r^4}. \tag{47}$$

Now we make a specific choice for the parameter ε. Given that it has dimensions of squared length, we take $\varepsilon = -2l_{\varepsilon}^2$, where l_{ε} represents some characteristic length scale. The sign of ε and the factor 2 have been chosen in such a way that the resulting solutions are identical to those found in the quadratic theory[6]

$$S = \frac{1}{2\kappa^2}\int d^4x\sqrt{-g}\left[R + l_{\varepsilon}^2(aR^2 + R_{\mu\nu}R^{\mu\nu})\right] - \frac{1}{16\pi}\int d^4x\sqrt{-g}F_{\mu\nu}F^{\mu\nu}. \tag{48}$$

This is a curious property of the Born-Infeld and quadratic gravity theories that occurs in four space-time dimensions with stress-energy tensors of the form (29). With this choice, we can introduce a dimensionless variable $z = r/r_c$ such that $r_c^4 \equiv l_{\varepsilon}^2 r_q^2$, with $r_q^2 \equiv \kappa^2 q^2/4\pi$, which turns (47) into

$$\Omega_{\pm} = 1 \pm \frac{1}{z^4}. \tag{49}$$

We can now use (35), recalling that $\tilde{r} = x$, to find that

$$r^2 = \frac{x^2 + \sqrt{x^4 + 4r_c^4}}{2}. \tag{50}$$

This relation puts forward that the area of the 2−spheres has a minimum of magnitude $A_c = 4\pi r_c^2$ at $x = 0$. In other words, the sector $r < r_c$ is excluded from the range of values of the area function $A = 4\pi r^2(x)$ (Fig. 1).

The mass function determined by (44) has a constant contribution and a term that comes from integrating over the electric field. The constant piece is identified with the Schwarzschild mass and will be denoted as M_0. To simplify the analysis, it is convenient to parametrize the mass function as follows:

$$M(r) = M_0(1 + \delta_1 G(z)), \tag{51}$$

[6]From an algebraic point of view, it is much easier to deal with the Born-Infeld model [94] than with the above quadratic theory [61], though from an effective field theory approach it is easier to motivate the latter. For this reason we analyzed the field equations of the Born-Infeld model but restrict the discussion of solutions to those with more interest in the quadratic theory. We note that the sign in front of l_{ε}^2 in (48) has been chosen in such a way that cosmological models with perfect fluids yield regular, bouncing solutions in both isotropic and anisotropic scenarios [117].

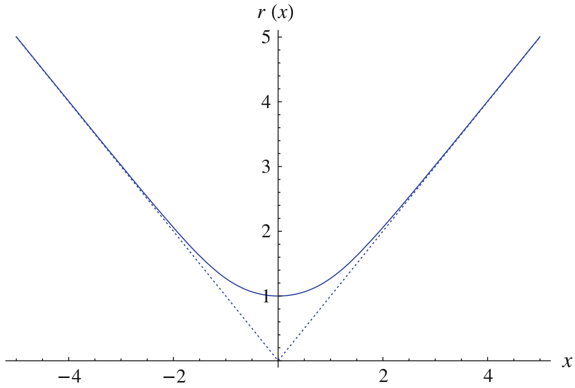


Fig. 1 Representation of $r(x)$ (*solid curve*), defined in (50), as a function of the radial coordinate x in units of the scale r_c. The *dotted lines* represent the function $|x|$

where δ_1 is a dimensionless constant and $G(z)$ encodes the contribution of the electric field. Inserting this form of $M(r)$ in (44), one finds

$$G_z = \frac{1}{z^4}\frac{(1+z^4)}{\sqrt{z^4-1}}, \tag{52}$$

and

$$\delta_1 = \frac{r_c^3}{2r_S l_\varepsilon^2} = \frac{1}{2r_S}\sqrt{\frac{r_q^3}{l_\varepsilon}}, \tag{53}$$

where $r_S \equiv 2M_0$ denotes the Schwarzschild radius. The integration of G_z is immediate and yields an infinite power series expansion of the form [61]

$$G(z) = -\frac{1}{\delta_c} + \frac{1}{2}\sqrt{z^4-1}\left[f_{3/4}(z) + f_{7/4}(z)\right], \tag{54}$$

where $f_\lambda(z) = {}_2F_1[\frac{1}{2}, \lambda, \frac{3}{2}, 1-z^4]$ is a hypergeometric function, and $\delta_c \approx 0.572069$ is a constant. Having obtained explicit solutions for $r^2(x)$ and $G(z)$, the space-time metric is completely specified.

6.1 Properties and Interpretation of the Solutions

One can verify from (52) that for $z \gg 1$, $G(z) \approx -1/z$ yields the expected Reissner-Nordström solution of GR, with $\Omega_\pm \approx 1$, $r^2(x) \approx x^2$, and

$$A(x) \approx 1 - \frac{r_S}{r} + \frac{r_q^2}{2r^2} + O\left(\frac{r_c^4}{r^4}\right). \tag{55}$$

From this expression one readily verifies that the typical configurations in terms of horizons found for Reissner-Nordström black holes also arise here, at least when the location of the horizon is much bigger than the scale r_c [61]. This occurs, in particular, when the charge-to-mass ratio δ_1 is greater than δ_c. We will refer to these configurations as RN-like. When $\delta_1 < \delta_c$, the solutions only have one horizon, like the Scharzschild black hole (Schwarzschild-like from now on). In some sense, the case $\delta_1 < \delta_c$ describes the limit in which the charge is much smaller than the mass. When $\delta_1 = \delta_c$, one finds a richer structure: depending on the number of charges, one can have one horizon, like in Schwarzschild, or have no horizons. More details on this will be given later.

It is apparent from (52) and (54) that the variable $z \equiv r/r_c$ can not become smaller than unity. This is consistent with (50) and tells us that something relevant occurs at $r = r_c$ (or $z = 1$ or $x = 0$). Some information in this direction can already be extracted from the action that defines the theory. The fact that we are considering the combination of gravity with an electric field without sources means that our theory does not know about the existence of *sources* for the electric field. In GR, the Reissner-Nordström solution is derived under similar assumptions, and one considers that the solution is only valid outside of the sources, which are supposed to be somehow concentrated at the origin. This picture, however, is not completely satisfactory, and a precise description of the sources is still an open question (see Chap. 8 of [25] for details). In our case, the combination of a minimum area for the two-spheres of the spherical sector together with the existence of an electric flux without sources points towards the notions of *geon* [22] and *wormhole* [118] suggested by J.A. Wheeler and C.W. Misner in the decade of 1950.

It is well-known that an electric field flowing through a hole in the topology (wormhole) can generate a charge which, from all perspectives, acts exactly in the same way as point charges. Wormholes are characterized by having a minimum area, which defines their throat [119]. The Born-Infeld theory combined with a free Maxwell field considered here, therefore, is yielding self-gravitating wormhole solutions for which there is no need to consider additional sources [120].

One should now note that in the derivation of the field equations we used a radial variable x which was different from $r(x)$. The reason for this is that r can only be used as a coordinate in those intervals in which it is a monotonic function of x [121], and $r(x)$ has a minimum at the wormhole throat ($x = 0$). Consistency of our *model of gravity plus electric field without sources* together with this behavior in the radial function implies the existence of a wormhole, in such a way that the range of x is the whole real line (from $-\infty$ to $+\infty$). The theory is thus describing a spherically symmetric electric field which flows from one universe into another through a wormhole located at $x = 0$ [120]. On one of the sides, the electric field lines point in the direction of increasing area thus defining a positive charge. On the other side, the electric field points into the direction of decreasing area, defining in this way a negative charge. This type of configuration is similar to that envisioned by Einstein and Rosen [122] when they used the Schwarzschild geometry to build a geometric model of elementary particles. A clear advantage of our model is that the wormhole structure arises naturally from the field equations and, therefore, one needs not follow a cut-and-paste strategy gluing together two exterior Schwarzschild geometries through the horizon to build the bridge that represents the particle in the Einstein–Rosen model. Moreover, a simple electric field has been able to generate a wormhole. This contrasts with the typical situation in GR, where wormholes supported by electric fields (linear like Maxwell's or nonlinear) are not possible [123], being necessary exotic energy sources that violate the energy conditions [21, 119].

Having established the wormhole nature of our solutions, one should re-think the meaning of the classification given above regarding event horizons. What we called Schwarzschild-like actually represents a wormhole with one horizon located somewhere on the $x > 0$ side of the $x-$axis and another horizon symmetrical with this one but on the $x < 0$ side. The RN-like configurations may have up to two horizons on each side of the $x-$axis. In the case with $\delta_1 = \delta_c$, depending on the amount of electric charge (which is a measure of the intensity of the electric flux), we can have Schwarzschild-like configurations (one horizon on each side of the axis), a case in which the two horizons converge at $x = 0$, and a horizonless family of (traversable) wormholes. This classification follows from a numerical study of the solutions of the equation $g_{tt} = -A/\Omega_+ = 0$ (see [61] for details).

An analytical discussion of the behavior near the wormhole throat is possible and useful. In fact, defining the number of charges as $N_q = q/e$, where e is the proton charge, we have

$$\lim_{r\to r_c} g_{tt} \approx \frac{l_P}{2l_\varepsilon}\frac{N_q}{N_c}\left[-\frac{(\delta_1-\delta_c)}{2\delta_1\delta_c}\sqrt{\frac{r_c}{r-r_c}} + \left(1-\frac{l_\varepsilon}{l_P}\frac{N_c}{N_q}\right) + O\left(\sqrt{r-r_c}\right)\right], \quad (56)$$

where, for convenience, we have introduced the Planck length $l_P = \sqrt{\hbar G/c^3}$ and $N_c \equiv \sqrt{2/\alpha_{em}} \approx 16.55$, with α_{em} representing the electromagnetic fine structure constant. This expression puts forward that the metric is finite at $r = r_c$ only for $\delta_1 = \delta_c$, diverging otherwise. By direct computation one can verify that curvature scalars generically diverge at $r = r_c$ except for those solutions with $\delta_1 = \delta_c$, where constant scalars are obtained. For this regular case, (56) also shows that the wormhole is hidden behind an event horizon if the sign of $\left(1-\frac{l_\varepsilon}{l_P}\frac{N_c}{N_q}\right)$ is positive, because then $g_{tt} > 0$ near the throat.

If we take $l_\varepsilon = l_P$, i.e., if the characteristic length scale of the gravity sector coincides with the Planck scale, then the event horizon for the regular solutions exists if $N_q > N_c$. For smaller values of the charge, $N_q \leq 16.55$, the horizon disappears and we are left with a regular horizonless object which could be interpreted as a black hole remnant. The existence of this type of solutions is interesting for theoretical as well as for astrophysical reasons. Theoretically, the existence of regular remnants could have important implications for the quantum information loss in the process of black hole evaporation [124]. From an astrophysical perspective, the existence of remnants could justify the lack of observational evidence for black hole explosions. Moreover, solutions of this type could contribute to the so-called dark matter in the form of very massive neutral atoms [120]. In fact, from the charge-to-mass constraint

$\delta_1 = \delta_c$, one finds that the mass of these solutions is completely determined by their electric charge according to the formula

$$M_0 = n_{BI} m_P \left(\frac{N_q}{N_c}\right)^{3/2} \left(\frac{l_P}{l_\varepsilon}\right)^{1/2}, \tag{57}$$

where $n_{BI} = \pi^{3/2}/(3\Gamma[3/4]^2) \approx 1.23605$ is a number that also arises in the determination of the total electrostatic energy of a point charge in the Born-Infeld theory of electrodynamics[7] (formulated in flat Minkowski space-time). With the mass formula (57), one can verify that Hawking's original predictions regarding the mass and charge spectrum of primordial black holes [125] formed in the early universe are in consonance with our results. He found that collapsed objects of order the Planck mass and above and with up to ± 30 electron charges could have been formed by large density fluctuations. It is typically argued that the existence of a quantum instability due to the horizon would make the lightest primordial black holes decay and evaporate. With the above explicit results, it is apparent that new mechanisms could lead to the formation of stable remnants which could survive until our times.

As a curiosity, from (57) one also finds that a solar mass black hole (with $\sim 10^{57}$ protons) of this type would require only $N_q \sim 3 \times 10^{26}$ charges (or $\sim$484 moles) to make the metric and all curvature scalars regular at the origin. Moreover, the external horizon of such an object would almost coincide with the Schwarzschild radius predicted by GR, making these objects astrophysically identical to those found in GR. This amount of charge certainly allows us to get rid of a number of important problems at a very low price. However, one should recall that (57) is only strictly valid for the $\delta_1 = \delta_c$ configuration, which suggests that only fine tuned configurations would be satisfactory. This raises a natural question: given that for $\delta_1 = \delta_c$ the geometry is completely regular and that infinitesimal deviations from this relation imply the development of curvature divergences and infinities in the metric, what happens to geodesics? In the $\delta_1 = \delta_c$ case we expect geodesics to be complete, as there is no reason to expect any pathological behavior that limits their extendibility at or near the wormhole throat. What happens to them when $\delta_1 \neq \delta_c$? Answering this question will provide us with useful information on the relation between curvature divergences and the existence of observers. In other words, this model offers us a good opportunity to better understand the correlation existing in GR between curvature divergences and geodesic incompleteness. We will resume this discussion later on, when we consider the geodesic equation in Sect. 8.

---

[7] In fact, using a notation similar to ours, in the Born-Infeld electromagnetic theory, whose Lagrangian is $\mathcal{L}_{BI} = \beta^2 \left(\sqrt{-|\eta_{\mu\nu} + \beta^{-1} F_{\mu\nu}|} - \sqrt{-|\eta_{\mu\nu}|}\right)$, one finds that the total electrostatic energy of a point particle is $\mathcal{E}_{BI} = \sqrt{2} n_{BI} m_P c^2 \left(\frac{N_q}{N_c}\right)^{3/2} \left(\frac{l_P}{l_\beta}\right)^{1/2}$, where $l_\beta^2 \equiv (4\pi/\kappa^2 c \beta^2)$ is a length scale associated to the β parameter of the theory.

7 Solutions in $f(\mathcal{R}) = \mathcal{R} - \lambda\mathcal{R}^2$

In Sect. 4 we discussed the field equations of the Palatini version of $f(R)$ theories. Now we would like to find nontrivial black hole solutions and study their properties to see how their geodesic structure compares with that provided by GR. A natural procedure would be to consider the coupling of an electric field as we did in the previous section in the case of Born-Infeld gravity. However, given that the stress-energy tensor of Maxwell's electrodynamics is traceless and that the modified dynamics of Palatini $f(R)$ theories depends crucially on nonlinear functions of this trace, we find that electrovacuum solutions in these theories are identical to those found in GR with a cosmological constant. Thus, in order to explore new physics, we need to consider matter sources whose stress-energy tensor has a non-zero trace.

To proceed, we consider a generic anisotropic fluid with stress-energy tensor of the form [95, 98]

$$T_\mu{}^\nu = \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & P_r & 0 & 0 \\ 0 & 0 & P_\theta & 0 \\ 0 & 0 & 0 & P_\varphi \end{pmatrix} \tag{58}$$

and set $P_r = -\rho$ and $P_\theta = P_\varphi = K(\rho)$, where $K(\rho)$ is some function of the fluid density, such that our fluid has the same structure as the generic stress-energy tensor considered in Sect. 5

$$T_\mu{}^\nu = diag[-\rho, -\rho, K(\rho), K(\rho)]. \tag{59}$$

It is worth noting that this structure of the stress-energy tensor allows us to see it as corresponding to a non-linear theory of electrodynamics [126]. In fact, for a theory where the electromagnetic Lagrangian goes from $X = -\frac{1}{2}F_{\mu\nu}F^{\mu\nu}$ to $\varphi(X)$, the stress-energy tensor becomes

$$T_\mu{}^\nu = \frac{1}{8\pi} diag[\varphi - 2X\varphi_X, \varphi - 2X\varphi_X, \varphi, \varphi]. \tag{60}$$

We can thus establish the correspondences $-8\pi\rho = (\varphi - 2X\varphi_X)$ and $K(\rho) = \varphi(X)$, which allow to solve for $\varphi(X)$ once a function $K(\rho)$ is specified.

Considering the fluid representation, the conservation equation $\nabla_\mu T^\mu{}_\nu = 0$ for a line element of the form $ds^2 = -C(x)dt^2 + B^{-1}(x)dx^2 + r^2(x)(d\theta^2 + \sin^2\theta d\varphi^2)$ leads to the relation $\rho_x + 2[\rho + K(\rho)]r_x/r = 0$. This expression can be readily integrated to obtain a formal relation between $\rho(x)$ and $r(x)$ given by

$$r^2(x) = r_0^2 \exp\left[-\int^\rho \frac{d\tilde{\rho}}{\tilde{\rho} + K(\tilde{\rho})}\right], \tag{61}$$

where r_0 is an integration constant with dimensions of length. In order to simplify our discussion, we shall restrict ourselves to the case $K(\rho) = \alpha\rho + \beta\rho^2$, where α

is a dimensionless constant and β has dimensions of inverse density. This example yields analytical solutions and covers a number of interesting cases. In particular, one finds that the relation between $\rho(x)$ and $r(x)$ turns into

$$\rho(r) = \frac{(1+\alpha)\rho_0}{\left(\frac{r}{r_0}\right)^{2(1+\alpha)} - \beta\rho_0}. \tag{62}$$

One readily verifies that when $\alpha = 1$ and $\beta = 0$, this fluid has the same stress-energy tensor as the Maxwell electric field (46), with $\rho_0 r_0^4 = q^2/8\pi$. The inclusion of the parameters α and β allows to generate a non-zero trace in the stress energy tensor. The case with $\beta = 0$ and $0 < \alpha < 1$ was studied in detail in [126]. Here we shall take $\alpha = 1$ and focus on the case $\beta < 0$ (a more exhaustive discussion will be presented elsewhere [127]). This family of models rapidly recovers the usual RN solution away from the center but regularizes the energy density, which is everywhere finite and bounded above by $\rho_m = \frac{(1+\alpha)}{|\beta|}$. We note that the effect of the parameter $\beta > 0$ is to shift the location of the divergence in the density from $r = 0$ to $(|\beta|\rho_0)^{1/(2+2\alpha)} r_0$. With our choice of negative β, we regularize the divergence of the matter sector.

To proceed, we set $\alpha = 1$, $\beta = -\tilde{\beta}/\rho_0$, and introduce a dimensionless variable $z^4 = r^4/\tilde{\beta} r_0^4$, in such a way that the density is now given by

$$\rho = \frac{\rho_m}{1+z^4}. \tag{63}$$

Using the trace equation (16) and the quadratic model $f = \mathcal{R} - \lambda\mathcal{R}^2$, one readily finds that $\mathcal{R} = -\kappa^2 T$, which is the same linear relation as in GR (this is just an accident of the quadratic model in four dimensions). We thus find that the function $f_{\mathcal{R}}$ takes the simple form

$$f_{\mathcal{R}} = 1 - \frac{\gamma}{(1+z^4)^2}, \tag{64}$$

where $\gamma \equiv \rho_m/\rho_\lambda$ and $\rho_\lambda \equiv 1/8\kappa^2\lambda$.

Following the same approach as in the Born-Infeld gravity theory studied above, we find that parametrizing the mass function as $M(r) = M_0(1 + \delta_1 G(z))$ leads to

$$G_z = \frac{z^2}{(1+z^4) f_{\mathcal{R}}^{3/2}} \left(1 - \frac{\gamma}{(1+z^4)^3}\right)\left(1 - \frac{\gamma(1-3z^4)}{(1+z^4)^3}\right) \tag{65}$$

$$\delta_1 \equiv \frac{\kappa^2 \rho_m (r_0 \tilde{\beta}^{\frac{1}{4}})^3}{r_S} \tag{66}$$

The function $G(z)$ can be obtained easily in terms of power series expansions and the solutions are classified in two types, depending on the value of the parameter γ. If $\gamma > 1$ then z is bounded from below, $z \geq z_c$, with $z_c^4 = \gamma^{1/2} - 1$ representing the location where $f_{\mathcal{R}} = 0$. At that point, the function G_z diverges, as can be easily understood from the expression (65), which has a term $f_{\mathcal{R}}^{3/2}$ in the denominator.

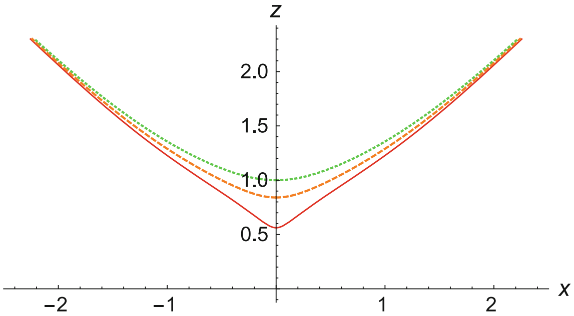


Fig. 2 Representation of $z(x)$ (*solid curve*) as a function of the radial coordinate x (in units of the scale $r_c = |\tilde{\beta}|^{1/4} r_0$) for different values of the parameter γ. The *solid* (*red*) *curve* corresponds to $\gamma = 1.1$, the *dashed* (*orange*) *curve* is $\gamma = 1.5$, and $\gamma = 2$ is the *dotted* (*green*) one

The lower bound on z signals the presence of a wormhole, in much the same way as we already observed in the case of Born-Infeld gravity. This is confirmed by the relation between the radial functions x and z given by $x^2 = f_{\mathcal{R}} z^2$, which is plotted in Fig. 2. Having this wormhole structure in mind, one finds that near z_c we have $f_{\mathcal{R}} \approx \frac{8z_c^3}{1+z_c^4}(z - z_c)$ and $G_z \approx C/(z - z_c)^{3/2}$, with $C > 0$ a constant (whose explicit form can be computed but is not necessary). This leads to $\lim_{z\to z_c} G(z) \approx -2C/\sqrt{z - z_c}$.

It is obvious that for $0 < \gamma < 1$ there are no real solutions for z_c. One finds that for that case, and also for $\gamma = 1$, the range of z is comprised between 0 and ∞, which implies that there is no wormhole, G_z is finite everywhere, and $G(z)$ tends to a constant as $z \to 0$. In fact, near $z = 0$ we can approximate $G(z) \approx -\frac{1}{\delta_c^{(\gamma)}} + (1-\gamma)^{1/2} z^3/3 + \frac{(7\gamma-1)}{\sqrt{1-\gamma}} z^7/7 + O(z^{11})$, where $\delta_c^{(\gamma)}$ is a constant. The case $\gamma = 1$ admits an analytical solution in terms of special functions and its series expansion must be considered separately, yielding $G(z) \approx -1/\delta_c^{(1)} + \frac{9z^5}{5\sqrt{2}} - \frac{13z^9}{4\sqrt{2}} + O(z^{13})$. One can easily verify that for $z \gg 1$ (65) rapidly converges to the GR prediction $G_z \approx 1/z^2$ regardless of the value of γ.

Let us now discuss the geometry near the center in the two cases distinguished above in terms of γ. Consider first the wormhole case, $\gamma > 1$, for which $\lim_{z\to z_c} f_{\mathcal{R}} \approx \frac{8z_c^3}{1+z_c^4}(z - z_c)$ and $\lim_{z\to z_c} G(z) \approx -2C/\sqrt{z - z_c}$. The area of the two spheres is determined by solving the relation $x^2 = f_{\mathcal{R}} r^2$. Denoting $r = z r_c$, $x = \tilde{x} r_c$, and $r_c = r_0 \tilde{\beta}^{1/4}$, one finds

$$\tilde{x} \approx \sqrt{\frac{8z_c^5}{1+z_c^4}}(z - z_c)^{1/2}, \tag{67}$$

which leads to

$$r^2(x) \approx r_c^2 z_c^2 + \frac{(1+z_c^4)}{4z_c^4} x^2. \tag{68}$$

This relation puts forward that the physical 2−spheres have a minimum area at $x = 0$, thus signaling the presence of a wormhole, as already advanced above. The g_{tt} component of the metric can be written as

$$g_{tt} = -\frac{1}{f_{\mathcal{R}}}\left(1 - \frac{r_S(1+\delta_1 G(z))}{x}\right) \approx -\frac{\tilde{C}}{(z-z_c)^2}, \tag{69}$$

where $\tilde{C}$ is a positive constant whose explicit form is not relevant. It is clear that for this type of solutions the metric diverges at $z = z_c$. One can also verify that curvature scalars generically diverge on that surface. We note that the properties of the solutions with $\gamma > 1$ are shared by all those models in which $f_{\mathcal{R}}$ has a simple pole at $z = z_c$. One can easily verify that if $f_R = b_0(z - z_c)$, then the two spheres satisfy a relation like (68) and the metric has a quadratic divergence at z_c.

When $0 < \gamma \leq 1$, the properties of the solutions largely depart from those observed in the case of having a pole in $f_{\mathcal{R}}$. Given that the function $f_{\mathcal{R}}$ does not vanish in this case, we find that near the center $\tilde{x} \approx \sqrt{1-\gamma}\, z$. The g_{tt} component of the metric then becomes

$$g_{tt} \approx -\frac{1}{(1-\gamma)}\left(1 - \frac{r_S(\delta_c^{(\gamma)} - \delta_1)}{r_c\delta_c^{(\gamma)}\sqrt{1-\gamma}\, z} - \frac{r_S\delta_1}{2r_c}z^2 \ldots\right). \tag{70}$$

This indicates that for the choice $\delta_1 = \delta_1^{(\gamma)}$, the metric is regular everywhere. Curvature scalars, however, do have divergences. For $\gamma = 1$, the above expression must be replaced by

$$g_{tt} \approx \frac{r_S}{2r_c\sqrt{2}z^7} - \frac{1}{2z^4} + O(z^{-3}). \tag{71}$$

We note that the case $\gamma \to 0$ yields the limit in which this anistropic fluid is coupled to GR. One can verify that the behavior of the solutions with $0 < \gamma \leq 1$ near the origin is similar to that of models of nonlinear electrodynamics coupled to GR [128–142].

8 Geodesics

The modified gravitational dynamics generated by the models considered in the previous sections has an impact on the space-time metric $g_{\mu\nu}$ and, consequently, on its associated geodesics. Since we are interested in determining whether the space-times derived above are geodesically complete or not, in this section we solve the geodesic equation and explore their behavior in those regions where GR typically yields incomplete paths.

The geodesics of a given connection $\Gamma^{\mu}_{\alpha\beta}$ are determined by the equation

$$\frac{d^2x^\mu}{d\lambda^2}+\Gamma^{\mu}_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}=0. \tag{72}$$

Here we will focus on the geodesics of the metric $g_{\mu\nu}$, which are the ones that matter fields can see according to the Einstein equivalence principle. We thus take $\Gamma^{\mu}_{\alpha\beta}$ as defined in (2). In order to solve these equations, we introduce a Hamiltonian approach that simplifies the analysis. To proceed, we first note that (72) can be derived from an action of the form [143]

$$S=\frac{1}{2}\int d\lambda g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}, \tag{73}$$

which for a line element like $ds^2=-C(x)dt^2+B^{-1}(x)dx^2+r^2(x)d\Omega^2$ becomes

$$S=\frac{1}{2}\int d\lambda\left[-C(x)\dot{t}^2+\frac{1}{B(x)}\dot{x}^2+r^2(x)\dot{\theta}^2+r^2(x)\sin^2\theta\dot{\varphi}^2\right]. \tag{74}$$

From this representation, one easily verifies that the momenta associated to the variables (t, x, θ, φ) are

$$P_t=-\frac{\partial L}{\partial \dot{t}}=\dot{t}C(x) \tag{75}$$

$$P_x=\frac{\partial L}{\partial \dot{x}}=\dot{x}/B(x) \tag{76}$$

$$P_\theta=\frac{\partial L}{\partial \dot{\theta}}=r^2(x)\dot{\theta} \tag{77}$$

$$P_\varphi=\frac{\partial L}{\partial \dot{\varphi}}=r^2(x)\sin^2\theta\dot{\varphi}. \tag{78}$$

With these momenta one finds that the Hamiltonian $H=-P_t\dot{t}+P_x\dot{x}+P_\theta\dot{\theta}+P_\varphi\dot{\varphi}-L$ coincides with the Lagrangian (due to the absence of potential terms) and can be written as

$$H=\frac{1}{2}g^{\mu\nu}(x)P_\mu P_\nu. \tag{79}$$

The geodesic equations can thus be written as

$$\dot{x}^\mu=\frac{\partial H}{\partial P_\mu}=g^{\mu\nu}P_\nu \tag{80}$$

$$\dot{P}_\mu=-\frac{\partial H}{\partial x^\mu}=-\frac{1}{2}(\partial_\mu g^{\alpha\beta})P_\alpha P_\beta \tag{81}$$

From these equations one readily sees that P_t and P_φ are constants of the motion, as $\dot{P}_t = 0 = \dot{P}_\varphi$. These equations also imply that $dH/d\lambda = 0$, showing that H is another conserved quantity. We thus have

$$P_t = \left(\frac{dt}{d\lambda}\right) C(x) = E \tag{82}$$

$$P_\varphi = \left(\frac{d\varphi}{d\lambda}\right) r^2(x) \sin^2\theta = L \tag{83}$$

$$2H = -\frac{P_t^2}{C(x)} + B(x)P_x^2 + \frac{P_\theta^2}{r^2(x)} + \frac{P_\varphi^2}{r^2(x)\sin^2\theta} = -\frac{E^2}{C(x)} + \frac{\dot{x}^2}{B(x)} + \frac{L^2}{r^2}, \tag{84}$$

where in the last equality we have set $\theta = \pi/2$ without loss of generality (because the motion takes place on a plane). When $H \neq 0$, a constant rescaling of the affine parameter $\lambda \to \lambda/\sqrt{|2H|}$ makes it clear that only the sign of H is physically relevant. This sign allows to classify the geodesics in three families: those with $H > 0$ (space-like), those with $H < 0$ (time-like), and those with $H = 0$ (null), which clarifies the meaning of this conserved quantity. Denoting $k \equiv 2H$ (with $k = 1, 0, -1$ corresponding to spatial, null, and time-like geodesics, respectively), (84) can be recast as

$$\frac{C(x)}{B(x)}\left(\frac{dx}{d\lambda}\right)^2 = E^2 - C(x)\left(\frac{L^2}{r^2(x)} - k\right), \tag{85}$$

which will be used to study the range of λ in different scenarios.

8.1 Geodesics in GR

Let us consider the Schwarzschild and Reissner-Nordström solutions of GR, whose line element takes the form

$$ds^2 = -C(r)^2 dt^2 + \frac{1}{C(r)}dr^2 + r^2 d\Omega^2, \tag{86}$$

with $C(r) = 1 - \frac{r_S}{r} + \frac{r_q^2}{2r^2}$, $r_S = 2GM_0/c^2$, $r_q^2 = \kappa^2 q^2/4\pi$ (for Schwarzschild, $r_q^2 = 0$), and $\kappa^2 = 8\pi G/c^4$. Given that here $C(r) = B(r)$, we find that (85) turns into

$$\left(\frac{dr}{d\lambda}\right)^2 = E^2 - C(r)\left(\frac{L^2}{r^2} - k\right). \tag{87}$$

This equation has the same structure as that of a particle with energy $\mathcal{E} = E^2$ in an effective one-dimensional potential of the form $V_{eff}(r) = C(r)\left(\frac{L^2}{r^2} - k\right)$, which facilitates its interpretation.

Let us consider first the uncharged (Schwarzschild) case. In this scenario, the function $C(r)$ becomes negative inside the horizon. As a result, the effective potential becomes an infinitely attractive well of the form $V_{eff} \approx -\frac{r_S}{r}\left(\frac{L^2}{r^2} - k\right)$, and the causal structure is such that all observers and light rays are forced to move in the direction of decreasing r as time goes by. This can be seen straightforwardly by just writing the line element (86) in ingoing Eddington-Finkelstein coordinates

$$ds^2 = -C(r)^2 dv^2 + 2dvdr + r^2 d\Omega^2, \tag{88}$$

where $dv = dt + dr/C(r)$ now plays the role of time coordinate. Inside the event horizon, where $A(r) < 0$, we see that

$$-2dvdr = -C(r)^2 dv^2 - ds^2 + r^2 d\Omega^2 \tag{89}$$

implies that as time goes by ($dv > 0$) we must have $dr < 0$ for time-like and null trajectories ($ds^2 \leq 0$). Thus, regardless of their point of origin, all physical observers and light rays will sooner or later end up at $r = 0$. The precise evolution of the affine parameter near the center is determined by $dr/d\lambda \approx -\sqrt{r_S/r}$ for radial timelike geodesics ($L = 0$) and by $dr/d\lambda \approx -\sqrt{r_S L^2/r^3}$ for timelike and null geodesics with $L \neq 0$. By integrating these expressions, we find $\lambda(r) = \lambda_0 - \frac{2}{3}\sqrt{r^3/r_S}$ and $\lambda(r) = \lambda_0 - \frac{2}{5}\sqrt{r^5/r_S L^2}$, respectively, where λ_0 represents the value of the affine parameter at $r = 0$. Given that the affine parameter cannot be extended *beyond* the center, these geodesics are incomplete in the future. A similar analysis can be carried out in the *white hole* region of the Schwarzschild geometry, where all geodesics are outgoing ($dr > 0$ with growing time). In that case, geodesics are incomplete in the past, i.e., they cannot be extended into $\lambda \to -\infty$. This space-time, therefore, can be regarded as singular.

In the Reissner-Nordström case, the situation is quite different from Schwazschild. As one approaches the center, the charge term dominates and $C(r) \sim \frac{r_q^2}{2r^2} > 0$ implies that for time-like observers ($k = -1$) $dr/d\lambda$ in (87) must vanish at some point before reaching $r = 0$ regardless of the value of L. These observers, therefore, bounce before reaching the center due to the presence of an infinite potential barrier and continue their trip in the direction of growing r, having the possibility of getting into new asymptotically flat regions if horizons are present. Something similar happens also to light rays ($k = 0$) with nonzero angular momentum L. However, for radial null geodesics ($k = 0$ and $L = 0$), we find $r(\lambda) = \pm E(\lambda - \lambda_0)$, where the minus sign represents ingoing rays and the plus sign outgoing rays. Ingoing rays cannot be extended beyond $\lambda = \lambda_0$, whereas outgoing rays are *created* at some finite λ. Thus, the Reissner-Nordström geometry is incomplete as far as radial null geodesics are concerned.

8.2 Geodesics in Born-Infeld Gravity

From our discussion of the spherically symmetric charged solutions found in Sect. 6 for the Born-Infeld theory, it is clear that geodesics in that space-time are essentially the same as in GR as soon as one moves a few r_c units away from the central wormhole [19]. In fact, in Fig. 1 one can readily see that $r(x) \approx x$ as soon as one reaches $|x| \approx 2r_c$. The g_{tt} component of the metric also converges quickly to the GR prediction, as shown in (55), with corrections that decay rapidly as $\sim(r_c/r)^4$. We thus only need to focus on the behavior of geodesics near the wormhole to explore the impact of curvature divergences on their completeness. Recall, in this sense, that the different metric solutions could be classified according to whether the charge-to-mass ratio δ_1, defined in (53), was smaller, equal, or larger than the characteristic value $\delta_c \approx 0.572069$ that arises in the electric field contribution to the mass function of (54). The case $\delta_1 = \delta_c$ was completely regular (no metric or curvature divergences [61]), whereas $\delta_1 < \delta_c$ (Schwarzschild-like) and $\delta_1 > \delta_c$ (RN-like) had divergences at the wormhole throat, $x = 0$ (or $r = r_c$ or $z = 1$).

Using the identifications $C(x) = A(x)/\Omega_+$ and $B(x) = A(x)\Omega_+$ together with the expression for $r^2(x)$ found in (50), (85) turns into

$$\frac{1}{\Omega_+^2}\left(\frac{dx}{d\lambda}\right)^2 = E^2 - \frac{A(x)}{\Omega_+}\left(\frac{L^2}{r^2(x)} - k\right). \tag{90}$$

For radial null geodesics ($L = 0,\ k = 0$), which are incomplete in both the Schwarzschild and RN solutions of GR, the above equation becomes independent of the function $A(x)$ and an exact solution can be found analytically. Using (50), one finds that $dx/dr = \pm\Omega_+/\Omega_-^{1/2}$, with the minus sign corresponding to $x \leq 0$. This turns (91) into

$$\frac{1}{\Omega_-}\left(\frac{dr}{d\lambda}\right)^2 = E^2, \tag{91}$$

which can be integrated to obtain

$$\pm E \cdot \lambda(x) = \begin{cases} {}_2F_1[-\frac{1}{4}, \frac{1}{2}, \frac{3}{4}; \frac{r_c^4}{r^4}]r & x \geq 0 \\ 2x_0 - {}_2F_1[-\frac{1}{4}, \frac{1}{2}, \frac{3}{4}; \frac{r_c^4}{r^4}]r & x \leq 0 \end{cases}, \tag{92}$$

where ${}_2F_1[a, b, c; y]$ is a hypergeometric function, $x_0 = {}_2F_1[-\frac{1}{4}, \frac{1}{2}, \frac{3}{4}; 1] = \frac{\sqrt{\pi}\Gamma[3/4]}{\Gamma[1/4]} \approx 0.59907$, and the $\pm$ sign corresponds to outgoing/ingoing null rays in the $x > 0$ region. It should be noted that given that $dr/d\lambda$ is a continuous function, the solution (92) is unique. One can easily verify that as $x \to \infty$ the series expansion of (92) yields $\pm E\lambda(x) \approx r + O(r^{-3}) \approx x$ and naturally recovers the GR behavior

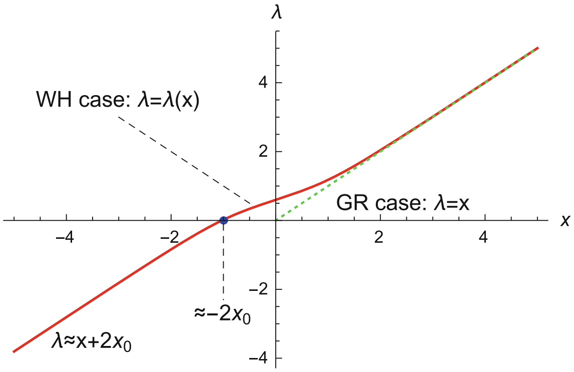


Fig. 3 Affine parameter $\lambda(x)$ as a function of the radial coordinate x for radial null geodesics (outgoing in $x > 0$). In the GR case (*green dashed curve* in the upper right quadrant), $\lambda = x$ is only defined for $x \geq 0$. For radial null geodesics in our wormhole spacetime (*solid red curve*), $\lambda(x)$ interpolates between the GR prediction and a shifted *straight line* $\lambda(x) \approx x + 2x_0$, with $x_0 \approx 0.59907$. In this plot $E = 1$ and the horizontal axis is measured in units of r_c

for large radii (see Fig. 3). As $x \to -\infty$, we get $\pm E\lambda(x) \approx x + 2x_0$, which also recovers the linear behavior of GR but shifted by a (negligible) constant factor.

Given that the radial coordinate x can naturally take negative values due to the wormhole structure, it follows that the affine parameter for radial null geodesics can be extended over the whole real line. As a result, these geodesics are complete. This was expected for the regular case with $\delta_1 = \delta_c$, for which the metric and all curvature scalars are finite everywhere, but was not obvious a priori for the other cases. Remarkably, the fact that this result is independent of the details of the function $A(x)$, which contains the information about δ_1, confirms that radial null geodesics are complete for all our solutions. This puts forward that a space-time can be geodesically complete even when there exist divergences in the metric and/or in curvature scalars. The wormhole has thus crucially contributed to allow the extendibility of the most critical geodesics of GR.

For nonradial and/or time-like geodesics, the discussion must take into account whether the geometry is Schwarzschild-like or RN-like. Considering the limit $x \to 0$, (91) turns into

$$\frac{1}{4}\left(\frac{dx}{d\lambda}\right)^2 = E^2 - V_{eff}(x) \tag{93}$$

$$V_{eff}(x) \approx -\frac{a}{|x|} - b, \tag{94}$$

with $a = \left(\kappa + \frac{L^2}{r_c^2}\right)\frac{(\delta_c - \delta_1)}{2\delta_c\delta_2}$, $b = \left(\kappa + \frac{L^2}{r_c^2}\right)\frac{(\delta_1 - \delta_2)}{2\delta_2}$, and $\delta_2 \equiv \delta_1 \frac{N_c}{N_q}\frac{l_\varepsilon}{l_P}$. From the above expression it is easy to see that in the RN-like configuration the coefficient a is negative, thus implying that the right-hand side of (93) must vanish at some point before reaching the wormhole. The situation is thus analogous to that already observed in the case of GR, with $L \neq 0$ geodesics bouncing before reaching the center (or the wormhole in our case). In the Schwarzschild-like configurations, the effective potential represents an infinite attractive well with the possibility of having a maximum before reaching the throat. As a consequence, all geodesics with energy above that maximum hit the wormhole (see [19] for more details). Using (93) and (94), one finds that the affine parameter behaves as

$$\lambda(x) \approx \lambda_0 \pm \frac{x}{3}\left|\frac{x}{a}\right|^{\frac{1}{2}}\left(1 - \frac{3(b+E^2)}{10}\left|\frac{x}{a}\right|\right). \tag{95}$$

This solution (which is unique) guarantees the extendibility of the affine parameter accross $x = 0$. Therefore, all time-like and null geodesics in these space-times are complete regardless of the existence of curvature divergences at the wormhole throat.

8.3 Geodesics in $f(R)$ Gravity

In the $f(\mathcal{R})$ case, our general approach for the description of geodesics leads to the following equation

$$\frac{1}{f_{\mathcal{R}}^2}\left(\frac{dx}{d\lambda}\right)^2 = E^2 - \frac{A(x)}{f_{\mathcal{R}}}\left(\frac{L^2}{r^2(x)} - k\right). \tag{96}$$

Let us consider first the case with $0 < \gamma < 1$, for which there is no wormhole structure. In these cases, as $x \to 0$ we find $f_{\mathcal{R}} \approx (1-\gamma)$, $r(x) \approx x/\sqrt{1-\gamma}$, and

$$A(z) \approx 1 - \frac{r_S(\delta_c^{(\gamma)} - \delta_1)}{r_c\delta_c^{(\gamma)}\sqrt{1-\gamma}\, z} - \frac{r_S\delta_1}{2r_c}z^2 + \cdots \tag{97}$$

With this, near the center (96) can be written as

$$\left(\frac{dr}{d\lambda}\right)^2 = \tilde{E}^2 - A(r)\left(\frac{L^2}{r^2} - k\right), \tag{98}$$

with $\tilde{E}^2 = (1-\gamma)E^2$. The discussion now proceeds in much the same way as in models of non-linear electrodynamics coupled to GR. One can find configurations for which the metric is regular at the origin, $\delta_c^{(\gamma)} = \delta_1$, and others with divergences, $\delta_c^{(\gamma)} \neq \delta_1$. A detailed discussion of geodesics in such configurations will be provided elsewhere [127]. The key point to note here is that the absence of a wormhole implies that radial null geodesics, $\left(\frac{dr}{d\lambda}\right)^2 = \tilde{E}^2$, always reach $r = 0$ in a finite proper time with no possibility of extension beyond that point. Thus, similarly as in the Reissner-Nordström case of GR, such solutions can be regarded as singular.

Let us now consider the case with $\gamma > 1$, for which there is a wormhole. From previous results, we know that as the wormhole is approached, we have $A(x) \approx \tilde{C}/(z - z_c)$ and $f_{\mathcal{R}} \approx \frac{8z_c^3}{1+z_c^4}(z - z_c)$, which implies that the right-hand side of (96) must vanish at some $z > z_c$ if $L \neq 0$ or $k = -1$ (time-like observers). This means that such geodesics never reach the wormhole throat, which is similar to what we already observed in the case of Reissner-Nordström in GR, where time-like observers and $L \neq 0$ geodesics never reach the center. If we consider radial null geodesics, (96) turns into

$$\frac{1}{f_{\mathcal{R}}^2}\left(\frac{dx}{d\lambda}\right)^2 = E^2. \tag{99}$$

Far from the wormhole $f_{\mathcal{R}} \to 1$ and this recovers the standard behavior $r \approx x \approx \pm E(\lambda - \lambda_0)$, with the $+/-$ sign corresponding to outgoing/ingoing rays. Now, near the wormhole, we can use the relation $r^2 f_{\mathcal{R}} = x^2$ and the fact that $r \to r_c$ as $x \to 0$ to write (99) as

$$\frac{r_c^4}{x^4}\left(\frac{dx}{d\lambda}\right)^2 = E^2, \tag{100}$$

which leads to

$$-\frac{1}{x} = \pm\frac{E}{r_c^2}(\lambda - \lambda_0). \tag{101}$$

From this it follows that as $x \to 0$, $\lambda \to -\infty$ for outgoing rays, while for ingoing rays $\lambda \to +\infty$. Stated in words, ingoing rays which started their trip from $x \to +\infty$ and $\lambda \to -\infty$ approach the wormhole at $x \to 0$ as $\lambda \to +\infty$, whereas outgoing rays which started their trip near the wormhole at $\lambda \to -\infty$ propagate to infinity as $\lambda \to +\infty$. Thus, all time-like and null geodesics in these configurations ($\gamma > 1$) are complete. Curvature divergences, which arise at the wormhole throat, cannot be reached in a finite affine parameter and, therefore, do not belong to the physically accessible region. These solutions are nonsingular even though one can never go through the wormhole. If one considers the region $x < 0$, identical conclusions are obtained.

9 Summary and Conclusions

In these Lectures we have studied the classical problem of black hole singularities from a four dimensional geometric perspective. Motivated by the fact that GR predicts the existence of singularities in simple static, spherically symmetric configurations, we have considered extensions of the theory to test the robustness of this disturbing result. In our study we have not followed the traditional approach of implicitly assuming that the space-time geometry is Riemannian. Rather, we have emphasized that the type of geometry associated with the gravitational interaction is an empirical question that must be settled by experiments, not imposed by convention or tradition. Whether the geometry is Riemannian or not is as fundamental a question as the number of space-time dimensions or the existence of supersymmetry, which are aspects that have received much attention in the last years.

We have thus considered a metric-affine geometrical framework for the formulation of our extensions of GR, with the additional simplification of setting torsion to zero (Palatini approach [40, 43]). This choice is justified on simplicity grounds, as a first step in the exploration of new gravitational physics. The inclusion of fermionic matter, whose spin sources the torsion, would require a detailed treatment beyond the Palatini approach.

An unusual property of the gravity theories considered here, as compared to the more standard metric or Riemannian approach, is that their modified dynamics arises as a result of nonlinearities generated by the matter fields rather than by the emergence of new dynamical degrees of freedom. In fact, the field equations of $f(\mathcal{R})$ theories, the Born-Infeld model, or any Lagrangian which is just a function of the inverse metric and the Ricci tensor à la Palatini admit a generic representation that exactly recovers the equations of GR (with an effective cosmological constant) in vacuum when the matter fields are absent [42, 113, 115, 144]. This means that generically these theories neither exhibit ghosts nor massive gravitons. These properties together with the second-order character of the field equations should be regarded as general characteristics of the metric-affine formulation.

In our opinion, the most remarkable aspect of the theories presented here is that they do what they were expected to do in a simple and clean manner. They were conceived as extensions of GR which could bring new relevant physics at high energies, and they yield solutions which are in agreement with GR almost everywhere, except in regions of very high energy density. The modifications that they introduce are such that black hole centers acquire a nontrivial structure that allows to preserve the completeness of geodesics. In the Born-Infeld type model, geodesics can go through the central wormhole, whereas in the $f(\mathcal{R})$ case, the wormhole (when it exists) lies beyond the reach of the geodesics.

Following the standard definition of space-time singularities given in the specialized literature and main text books on gravitation [9–12], we have concluded that the solutions containing wormholes are nonsingular because they are geodesically complete. And this is so despite the appearance of curvature divergences at the wormhole throat. One should note, however, that there exists a widespread tendency in the literature to simplify the complex notion of space-time singularity and associate the divergence of certain quantities (such as curvature scalars or tensor components) with its definition. This tendency can be partly justified by the *strong correlation* existing between the occurrence of divergences and the incompleteness of some geodesics. Somehow, one tends intuitively to associate divergences with geodesic incompleteness as if the former were the cause/reason for the latter [8]. We have shown here with several explicit examples that black hole space-times can be geodesically complete and at the same time have curvature divergences, thus breaking the correlation typically found in GR.

Divergences in curvature tensors/scalars are obviously associated with strong tidal forces. The effects of such forces have been investigated in the literature by means of geodesic congruences in an attempt to classify the strength of singularities [145–151]. In that context, extended physical objects are represented as congruences of geodesics, and the evolution of their relative distance as curvature divergences are approached provides information about their fate. Those methods have been applied in the general charged solutions of the Born-Infeld model studied here finding that the different parts of a body that goes through the wormhole never lose causal contact among them despite the existence of infinite accelerations at the throat [152]. This offers a new view on the problem which should be further investigated to better understand if curvature divergences possess any *destructive* power. We would like to emphasize that though in the Born-Infeld model physical observers do interact with the curvature divergence as the wormhole is crossed, in the $f(\mathcal{R})$ case, the divergence is never reached in a finite affine distance. Therefore, the $f(\mathcal{R})$ model is free from the potential drawbacks of directly interacting with a curvature divergence, as it lies beyond the physically accessible space-time.

Though much research is still needed to better understand gravitational and non-gravitational physics in metric-affine spaces, the point is that two analytically tractable *toy models* with nontrivial results about black holes are already available.

Before concluding, we must note that our approach has assumed that particles and observers can be viewed as structureless entities (geodesics), whereas physical measurements are carried out by means of probes with wave-like properties because matter fields are of a quantum nature. One should thus study the propagation of waves in these space-times to see how they behave and interact with regions of intense gravitational fields such as wormhole throats, where curvature scalars typically diverge. A first analysis in this direction was carried out in [93], where the scattering of scalar waves in horizonless (naked) configurations was considered. Despite the infinite potential barrier that curvature divergences generate, one verifies that the propagation through the wormhole is smooth and that transmission and reflection coefficients can be computed numerically and contrasted with analytical estimates, yielding good agreement. These results, therefore, give further support to the absence of singularities in these geometries.

Acknowledgments The author is supported by a Ramon y Cajal contract, the Spanish grants FIS2014-57387-C3-1-P and FIS2011-29813-C02-02 from MINECO, the grants i-LINK0780 and i-COOPB20105 of the Spanish Research Council (CSIC), the Consolider Program CPANPHY-1205388. This work has also been supported by the Severo Ochoa Grant SEV-2014-0398 (Spain) and the CNPq project No. 301137/2014-5 (Brazilian agency).

References

1. R. Penrose, Phys. Rev. Lett. **14**, 57 (1965)
2. R. Penrose, Riv. Nuovo Cim. Numero Speciale **1**, 252 (1969)

3. R. Penrose, Gen. Relativ. Gravit. **34**, 1141 (2002)
4. S.W. Hawking, Phys. Rev. Lett. **17**, 444 (1966)
5. B. Carter, Phys. Rev. Lett. **26**, 331 (1971)
6. S.W. Hawking, Phys. Rev. D **14**, 2460 (1976)
7. C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation and Cosmology* (W.H. Freeman, London, 1973)
8. E. Curiel, P. Bokulich, Singularities and black holes, in *The Stanford Encyclopedia of Philosophy* ed. by E.N. Zalta. (http://plato.stanford.edu/archives/fall2012/entries/spacetime-singularities/-2012)
9. R.P. Geroch, Ann. Phys. **48**, 526 (1968)
10. S.W. Hawking, G.F.R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge University Press, Cambridge, 1973)
11. R.M. Wald, *General Relativity* (University Press, Chicago, 1984)
12. J.M.M. Senovilla, D. Garfinkle, Class. Quantum Gravit. **32**(12), 124008 (2015). arXiv:1410.5226 [gr-qc]
13. V.F. Mukhanov, R.H. Brandenberger, Phys. Rev. Lett. **68**, 1969 (1992)
14. S. Ansoldi, arXiv:0802.0330 [gr-qc]
15. J.P.S. Lemos, V.T. Zanchin, Phys. Rev. D **83**, 124005 (2011)
16. E. Spallucci, A. Smailagic, P. Nicolini, Phys. Lett. B **670**, 449 (2009)
17. K.A. Bronnikov, J.C. Fabris, Phys. Rev. Lett. **96**, 251101 (2006)
18. S.A. Hayward, Phys. Rev. Lett. **96**, 031103 (2006)
19. G.J. Olmo, D. Rubiera-Garcia, A. Sanchez-Puente, Phys. Rev. D **92**(4), 044047 (2015)
20. M. Visser, *Lorentzian Wormholes: From Einstein to Hawking* (American Institute of Physics, New York, 1995)
21. F.S.N. Lobo, Classical and Quantum Gravity Research (Nova Science Publishers, New York, 2008), pp. 1–78. arXiv:0710.4474 [gr-qc]
22. J.A. Wheeler, Phys. Rev. **97**, 511 (1955)
23. F.S.N. Lobo, G.J. Olmo, D. Rubiera-Garcia, JCAP 1307, 011 (2013); Eur. Phys. J. C **74**, 2924 (2014)
24. F.S.N. Lobo, J. Martinez-Asencio, G.J. Olmo, D. Rubiera-Garcia, Phys. Lett. B **731**, 163 (2014)
25. T. Ortin, *Gravity and Strings (Cambridge Monographs on Mathematical Physics)*, 2nd edn. (2015)
26. L.P. Eisenhart, Non-Riemannian Geometry. Am. Math. Soc. **184** (1927)
27. H. Feigl, M. Brodbeck (eds.), *Readings in the Phylosophy of Science* (Appleton-Century-Crofts, New York, 1953). (See page 193)
28. C.M. Will, Living Rev. Relativ. **17**, 4 (2014). arXiv:1403.7377 [gr-qc]
29. E. Kröner, Z. angew. Math. Mech. **66**, 5 (1986)
30. H. Kleinert, *Gauged Fields in Condensed Matter*, vol. II (World Scientific, Singapore, 1989)
31. E. Kröner, Int. J. Theor. Phys. **29**, 1219 (1990)
32. R. de Witt, Int. J. Eng. Sci. **19**, 1475 (1981)
33. E. Kröner, Trends in applications of pure mathematics to mechanics. Lect. Notes Phys. **249**, 281 (1986)
34. E. Kröner, Int. J. Solids Struct. **29**, 1849 (1992)
35. K. Kondo, Proceedings of the 2nd Japan Kat. Congr. of Appl. at Max-Planek-lnstitut fur Mctallforschung, Postfach 800665, D-7000 Stuttgart 80, BRD. Mechanics, p. 41 (1952)
36. B.A. Bilby, R. Bullough, E. Smith, Proc. Roy. Soc. Lond. Ser. A **231**, 263 (1955)
37. F.S.N. Lobo, G.J. Olmo, D. Rubiera-Garcia, Phys. Rev. D **91**(12), 124001 (2015). arXiv:1412.4499 [hep-th]
38. G.J. Olmo, D. Rubiera-Garcia, Int. J. Mod. Phys. D **24**(09), 1542013 (2015). arXiv:1507.07777 [hep-th]
39. G.J. Olmo, D. Rubiera-Garcia, J. Phys. Conf. Ser. **600**(1), 012041 (2015). arXiv:1506.02139 [gr-qc]
40. M. Ferraris, M. Francaviglia, C. Reina, Gen. Relativ. Gravit. **14**(3), 244 (1982)

41. G.J. Olmo, D. Rubiera-Garcia, Phys. Rev. D **88**, 084030 (2013). arXiv:1306.4210 [hep-th]
42. G.J. Olmo, Chapter of the book 'Open Questions in Cosmology', ed. by G.J. Olmo (InTech Publishing, Rijeka, Croatia, 2012) ISBN 978-953-51-0880-1. (Downloadable from http://www.intechopen.com/books/open-questions-in-cosmology). arXiv:1212.6393
43. G.J. Olmo, Int. J. Mod. Phys. D **20**, 413 (2011). arXiv:1101.3864 [gr-qc]
44. L. Parker, D.J. Toms, *Quantum Field Theory in Curved Spacetime: Quantized Fields and Gravity* (Cambridge University Press, Cambridge, 2009)
45. N.D. Birrel, P.C.W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, 1982)
46. J.A.R. Cembranos, Phys. Rev. Lett. **102**, 141301 (2009). arXiv:0809.1653 [hep-ph]
47. R. Schimming, R.J. Schmidt, NTM Schriftenr. Gesch. Naturw. Tech. Med. **27**, 41 (1990). arXiv:gr-qc/0412038
48. R. Utiyama, B.S. DeWitt, J. Math. Phys. **3**, 608 (1962)
49. K.S. Stelle, Phys. Rev. D **16**(3), 953 (1977)
50. K.S. Stelle, Gen. Relativ. Gravit. **9**, 353 (1978)
51. H. Weyl, Ann. der Phys. **59**, 101 (1919)
52. H. Weyl, Phys. Zeits. **22**, 473 (1921)
53. W. Pauli, Phys. Zeits. **20**, 457 (1919)
54. A.A. Starobinsky, Phys. Lett. **91B**, 99 (1980)
55. P. Anderson, Phys. Rev. D **28**, 271 (1983)
56. A. Economou, C.O. Lousto, Phys. Rev. D **49**(1), 5278 (1994). arXiv:gr-qc/9310021
57. H. Lu, C. Pope, Phys. Rev. Lett. **106**, 181302 (2011). arXiv:1101.1971 [hep-th]
58. I. Gullu, M. Gurses, T.C. Sisman, B. Tekin, Phys. Rev. D **83**, 084015 (2011). arXiv:1102.1921 [hep-th]
59. H. Liu, H. Lu, M. Luo, Int. J. Mod. Phys. D **21**, 1250020 (2012). arXiv:1104.2623 [hep-th]
60. W. Berej, J. Matyjasek, D. Tryniecki, M. Woronowicz, Gen. Relativ. Gravit. **38**, 885 (2006). arXiv:hep-th/0606185
61. G.J. Olmo, D. Rubiera-Garcia, Phys. Rev. D **86**, 044014 (2012)
62. G.J. Olmo, D. Rubiera-Garcia, Int. J. Mod. Phys. D **21**, 1250067 (2012)
63. G.J. Olmo, D. Rubiera-Garcia, Eur. Phys. J. C **72**, 2098 (2012)
64. G.J. Olmo, H. Sanchis-Alepuz, S. Tripathi, Phys. Rev. D **80**, 024013 (2009). arXiv:0907.2787 [gr-qc]
65. G.J. Olmo, Phys. Rev. D **72**, 083505 (2005). arXiv:gr-qc/0505135
66. G.J. Olmo, Phys. Rev. Lett. **95**, 261102 (2005). arXiv:gr-qc/0505101
67. J. Martinez-Asencio, G.J. Olmo, D. Rubiera-Garcia, Phys. Rev. D **86**, 104010 (2012). arXiv:1209.3371 [gr-qc]
68. S. Deser, G.W. Gibbons, Class. Quantum Gravit. **15**, L35 (1998). arXiv:hep-th/9803049
69. D.N. Vollick, Phys. Rev. D **69**, 064030 (2004). arXiv:gr-qc/0309101
70. X.-L. Du, K. Yang, X.-H. Meng, Y.X. Liu, arXiv:1403.0083 [gr-qc]
71. H.C. Kim, arXiv:1312.0703 [gr-qc]
72. S.I. Kruglov, Phys. Rev. D **89**, 064004 (2014). arXiv:1310.6915 [gr-qc]
73. K. Yang, X.-L. Du, Y.-X. Liu, Phys. Rev. D **88**, 124037 (2013). arXiv:1307.2969 [gr-qc]
74. P.P. Avelino, R.Z. Ferreira, Phys. Rev. D **86**, 041501 (2012). arXiv:1205.6676 [astro-ph.CO]
75. A. De Felice, B. Gumjudpai, S. Jhingan, Phys. Rev. D **86**, 043525 (2012). arXiv:1205.1168 [gr-qc]
76. C. Escamilla-Rivera, M. Bañados, P.G. Ferreira, Phys. Rev. D **85**, 087302 (2012). arXiv:1204.1691 [gr-qc]
77. I. Cho, H.-C. Kim, T. Moon, Phys. Rev. D **86**, 084018 (2012). arXiv:1208.2146 [gr-qc]
78. J.H.C. Scargill, M. Bañados, P.G. Ferreira, Phys. Rev. D **86**, 103533 (2012). arXiv:1210.1521 [astro-ph.CO]
79. C. Escamilla-Rivera, M. Bañados, P.G. Ferreira, arXiv:1301.5264 [gr-qc]
80. M. Bañados, P.G. Ferreira, Phys. Rev. Lett. **105**, 011101 (2010). arXiv:1006.1769 [astro-ph.CO]

81. T. Harko, F.S.N. Lobo, M.K. Mak, S.V. Sushkov, Mod. Phys. Lett. A **29**, 1450049 (2014). arXiv:1305.0820 [gr-qc]
82. P.P. Avelino, Phys. Rev. D **85**, 104053 (2012). arXiv:1201.2544 [astro-ph.CO]
83. P.P. Avelino, JCAP **1211**, 022 (2012). arXiv:1207.4730 [astro-ph.CO]
84. Y.-H. Sham, L.-M. Lin, P.T. Leung, Astrophys. J. **781**, 66 (2014). arXiv:1312.1011 [gr-qc]
85. H.-C. Kim, arXiv:1312.0705 [gr-qc]
86. T. Harko, F.S.N. Lobo, M.K. Mak, S.V. Sushkov, Phys. Rev. D **88**, 044032 (2013). arXiv:1305.6770 [gr-qc]
87. Y.H. Sham, P.T. Leung, L.M. Lin, Phys. Rev. D **87**, 061503 (2013). arXiv:1304.0550 [gr-qc]
88. Y.-H. Sham, L.-M. Lin, P.T. Leung, Phys. Rev. D **86**, 064015 (2012). arXiv:1208.1314 [gr-qc]
89. P. Pani, T.P. Sotiriou, Phys. Rev. Lett. **109**, 251102 (2012). arXiv:1209.2972 [gr-qc]
90. P. Pani, V. Cardoso, T. Delsate, Phys. Rev. Lett. **107**, 031101 (2011). arXiv:1106.3569 [gr-qc]
91. M. Bouhmadi-Lopez, C.Y. Chen, P. Chen, arXiv:1302.5013 [gr-qc]
92. R. Ferraro, F. Fiorini, J. Phys. Conf. Ser. **314**, 012114 (2011). arXiv:1011.5196 [gr-qc]
93. G.J. Olmo, D. Rubiera-Garcia, A. Sanchez-Puente, Eur. Phys. J. C. **76**(3), 143 (2016)
94. G.J. Olmo, D. Rubiera-Garcia, H. Sanchis-Alepuz, Eur. Phys. J. C **74**, 2804 (2014). arXiv:1311.0815 [hep-th]
95. R. Shaikh, arXiv:1505.01314 [gr-qc]
96. C. Bambi, G.J. Olmo, D. Rubiera-Garcia, Phys. Rev. D **91**(10), 104010 (2015). arXiv:1504.01827 [gr-qc]
97. F.S.N. Lobo, G.J. Olmo, D. Rubiera-Garcia, Eur. Phys. J. C. **74**(6), 2924 (2014)
98. T. Harko, F.S.N. Lobo, M.K. Mak, S.V. Sushkov, arXiv:1307.1883 [gr-qc]
99. S.D. Odintsov, G.J. Olmo, D. Rubiera-Garcia, Phys. Rev. D **90**(4), 044003 (2014). arXiv:1406.1205 [hep-th]
100. A.N. Makarenko, S. Odintsov, G.J. Olmo, arXiv:1403.7409 [hep-th]
101. A.N. Makarenko, S.D. Odintsov, G.J. Olmo, Phys. Lett. B **734**, 36 (2014). arXiv:1404.2850 [gr-qc]
102. D. Comelli, Phys. Rev. D **72**, 064018 (2005)
103. D. Comelli, A. Dolgov, JHEP **0411**(062), 062 (2004)
104. J.A. Feigenbaum, Phys. Rev. D **58**, 124023 (1998)
105. J.A. Feigenbaum, P.G.O. Freund, M. Pigli, Phys. Rev. D **57**, 4738 (1998)
106. R. Ferraro, F. Fiorini, Phys. Lett. B **692**, 206 (2010)
107. F. Fiorini, Phys. Rev. Lett. **111**, 041104 (2013)
108. I. Gullu, T.C. Sisman, B. Tekin, Class. Quantum Gravit. **27**, 162001 (2010)
109. I. Gullu, T.C. Sisman, B. Tekin, Phys. Rev. D **81**, 104018 (2010)
110. J.A. Nieto, Phys. Rev. D **70**, 044042 (2004)
111. M.N.R. Wohlfarth, Class. Quantum Gravit. **21**, 1927 (2004); Corrigendum, Class. Quantum Gravit. **21**, 5297 (2004)
112. J.B. Jiménez, L. Heisenberg, G.J. Olmo, JCAP **1411**, 004 (2014). arXiv:1409.0233 [hep-th]
113. J.B. Jiménez, L. Heisenberg, G.J. Olmo, JCAP **1506**, 026 (2015). arXiv:1504.00295 [gr-qc]
114. J.B. Jiménez, L. Heisenberg, G.J. Olmo, C. Ringeval, arXiv:1509.01188 [gr-qc]
115. D. Bazeia, L. Losano, R. Menezes, G.J. Olmo, D. Rubiera-Garcia, Class. Quantum Gravit. **32**(21), 215011 (2015). arXiv:1509.04895 [hep-th]
116. D. Bazeia, L. Losano, G.J. Olmo, D. Rubiera-Garcia, A. Sanchez-Puente, Phys. Rev. D **92**(4), 044018 (2015). arXiv:1507.07763 [hep-th]
117. C. Barragan, G.J. Olmo, Phys. Rev. D **82**, 084015 (2010). arXiv:1005.4136 [gr-qc]
118. C.W. Misner, J.A. Wheeler, Ann. Phys. **2**, 525 (1957)
119. M. Visser, *Lorentzian Wormholes: From Einstein to Hawking* (AIP, Woodbury, 1995)
120. F.S.N. Lobo, G.J. Olmo, D. Rubiera-Garcia, JCAP **1307**, 011 (2013). arXiv:1306.2504 [hep-th]
121. H. Stephani, E. Herlt, M. MacCullum, C. Hoenselaers, D. Kramer, *Exact Solutions of Einstein's Field Equations* (Cambridge University Press, Cambridge, 2003), p. 726
122. A. Einstein, N. Rosen, Phys. Rev. **48**, 73 (1935)
123. A.V.B. Arellano, F.S.N. Lobo, Class. Quantum Gravit. **23**, 7229 (2006). arXiv:gr-qc/0604095

124. A. Fabbri, J. Navarro-Salas, *Modeling Black Hole Evaporation* (Imperial College Press, London, 2005)
125. S. Hawking, Mon. Not. Roy. Astron. Soc. **152**, 75 (1971)
126. G.J. Olmo, D. Rubiera-Garcia, Universe 2015, **1**(2), 173–185 (2015). arXiv:1509.02430 [hep-th]
127. C. Bejarano, G.J. Olmo, and D. Rubiera-Garcia, Work in Progress
128. G.J. Olmo, D. Rubiera-Garcia, Phys. Rev. D **84**, 124059 (2011). arXiv:1110.0850 [gr-qc]
129. A. García, H. Salazar, J.F. Plebański, Nuovo Cimento Soc. Ital. Fis. **84B**, 65 (1984)
130. M. Demianski, Found. Phys. **16**, 187 (1986)
131. H.P. de Oliveira, Class. Quantum Gravit. **11**, 1469 (1994)
132. G.W. Gibbons, D.A. Rasheed, Nucl. Phys. B **454**, 185 (1995)
133. E. Ayón-Beato, A. García, Phys. Rev. Lett. **80**, 5056–5059 (1998)
134. E. Ayón-Beato, A. García, Gen. Relativ. Gravit. **31**, 629–633 (1999)
135. E. Ayón-Beato, A. García, Phys. Lett. B **464**, 25 (1999)
136. N. Breton, Phys. Rev. D **67**, 124004 (2003)
137. M. Hassaine, C. Martinez, Phys. Rev. D **75**, 027502 (2007)
138. M. Hassaine, C. Martinez, Class. Quantum Gravit. **25**, 195023 (2008)
139. J. Diaz-Alonso, D. Rubiera-Garcia, Phys. Rev. D **81**, 064021 (2010)
140. J. Diaz-Alonso, D. Rubiera-Garcia, Phys. Rev. D **82**, 085024 (2010)
141. J. Diaz-Alonso, D. Rubiera-Garcia, Gen. Relativ. Gravit. **45**, 1901 (2013)
142. J.A.R. Cembranos, A. de la Cruz-Dombriz, J. Jarillo, JCAP **1502**(02), 042 (2015)
143. S. Chandrasekhar, *The Mathematical Theory of Black Holes* (Oxford University Press, New York, 1992)
144. G.J. Olmo, AIP Conf. Proc. 1458, 222–237 (2011). arXiv:1207.4815v1 [gr-qc]
145. G.F.R. Ellis, B.G. Schmidt, Gen. Relativ. Gravit. **8**, 915 (1977)
146. F.J. Tipler, Phys. Rev. D **15**, 942 (1977)
147. C.J.S. Clarke, A. Krolak, J. Geom. Phys. **2**, 127 (1985)
148. B.C. Nolan, Phys. Rev. D **60**, 024014 (1999)
149. A. Ori, Phys. Rev. D **61**, 064016 (2000)
150. F.J. Tipler, Phys. Lett. A **64**, 8 (1977)
151. B.C. Nolan, Phys. Rev. D **62**, 044015 (2000)
152. A. Sanchez-Puente, G.J. Olmo, D. Rubiera-Garcia, Impact of Curvature Divergences on Physical Observers in a Wormhole Space-time with Horizons. Class. Quant. Grav., in press (2016). arXiv:1602.01798 [hep-th]

Inflation: Observations and Attractors

Diederik Roest and Marco Scalisi

Abstract In these lecture notes, we present the latest status of CMB observations and outline a particular set of inflationary models to explain these data. As an introduction, we provide the necessary background to understand the Planck results on the temperature fluctuations of the CMB. We then explain how these results can be interpreted in terms of the number of e-folds during inflation. Finally, we discuss theoretical models that underpin this interpretation and yield robust predictions for future CMB observables.

1 Introduction and Outline

These notes are an extended write-up of a set of lectures given by the first author in the school "Theoretical Frontiers in Black Holes and Cosmology" in Natal, Brazil, from June 8–12, 2015. They do not aim to give an exhaustive overview of cosmological inflation; instead we will highlight a number of recent developments, both at the observational as well as theoretical front, with an interesting interplay between them. We hope they serve as an interesting stand-alone introduction to these particular aspects of inflation. When possible we will avoid technical details, deferring these to the original literature, and take a more pedestrian approach.

We will first introduce the standard cosmological viewpoint. This leads one to conjecture a period of inflation in the very early Universe. In order to understand the consequences of this phase, we study a consistent quantum formulation of the paradigm where initial quantum fluctuations represent the natural seeds for the formation of the cosmological structures. This allows us to present the most recent observations on the cosmic microwave background, and provide a theoretical interpretation

D. Roest (✉) · M. Scalisi
Van Swinderen Institute for Particle Physics and Gravity, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
e-mail: d.roest@rug.nl

M. Scalisi
e-mail: m.scalisi@rug.nl

R. Kallosh and E. Orazi (eds.), *Theoretical Frontiers in Black Holes and Cosmology*, Springer Proceedings in Physics 176,
DOI 10.1007/978-3-319-31352-8_6

of them. Finally, we discuss progress in inflationary model building, focusing on the notion of cosmological attractors. Throughout these notes we will refer to some of the relevant papers. Complementary material can be found in more extensive reviews, see e.g. [1–5]

2 Standard Cosmology in a Nutshell

In 1929 the astronomer Edwin Hubble made a discovery [6] which has revolutionized the understanding of our Universe as a whole, and has given rise to the subsequent establishment of cosmology as a science. He observed the mutual recession of galaxies, which was almost immediately interpreted as first evidence that we live in an expanding Universe. This simple idea led to the development of the standard model of Big Bang cosmology, whose predictions are in excellent agreement with observations. Despite the name, the model says nothing about the "Big Bang" which remains a mathematical singularity as well as an unsolved physical question. On the other hand, it furnishes a clear and precise picture of the cosmic evolution from a few seconds after this mysterious start: the temperature decreases as the expansion of the Universe proceeds, light elements form during a process called Big Bang Nucleosynthesis (BBN), recombination of nuclei and electrons takes place followed by the last scattering of photons which freely reach us today as cosmic microwave background (CMB) radiation, observed in the sky at the temperature $T = 2.73\,\mathrm{K}$.

Although the model has had many successful experimental confirmations, it contains some serious theoretical shortcomings which can be better understood once we know the geometric properties of the Universe we live in.

2.1 FRW Geometry and Dynamics

A dynamical Universe is what comes naturally from Einstein theory of general relativity which relates the geometry of spacetime to its matter-energy content, through the field equations (throughout these notes we have fixed Newton's constant by setting the reduced Planck mass to unity: $M_{Pl} = 1$)

$$G_{\mu\nu} = T_{\mu\nu}. \tag{1}$$

Prior to Hubble's discovery, Einstein had already noticed such a genuine prediction of a non-static Universe. However, puzzled by its cosmological implications, he augmented his equations with a specific cosmological constant in order to avoid such a phenomenon. Hubble's discovery however confirmed that we do live in a non-static Universe.

The simple observation that our Universe is homogeneous and isotropic at large scales ($>100\,\mathrm{Mpc}$) imposes stringent constraints on the form of both sides of (1).

Originally an assumption, this so-called *cosmological principle* has been beautifully confirmed by the observations of the distribution of galaxies at large scales REF and the homogeneity and isotropy of the CMB radiation REF. Assuming these symmetries leads to the Friedmann–Robertson–Walker (FRW) metric which, written in terms of polar spherical coordinates (r, θ, σ), reads

$$ds^2 = -dt^2 + a(t)^2 \left[\frac{dr^2}{1 - r^2} + r^2 (d\theta^2 + \sin^2\theta \, d\sigma^2) \right]. \tag{2}$$

The scale factor $a(t)$ sets the physical distances among objects and can vary with respect to the cosmic time t (the proper time as measured by a comoving observer at constant spatial coordinates) allowing, then, for an expanding Universe. The coordinates (r, θ, σ) reflect the symmetries assumed and are called "comoving coordinates" as they are decoupled from the effect of the expansion. An FRW Universe can be thought as an expanding grid where objects can be fixed on it (i.e. at constant comoving coordinates) and still recede from each other as an effect of a growing scale factor. Typical scales, e.g. the wavelength λ of a photon, will increase as $\lambda \propto a$ as the expansion proceeds. However, the comoving wavelength λ/a will remain constant in time, if no other external process occurs (see Fig. 1).

Homogeneity and isotropy still allow for a constant curvature of the 3-dimensional spatial slices which can correspond to an open, flat or closed Universe and is parametrized by $\kappa = -1, 0, 1$, respectively. Moreover, the stress-energy tensor $T_{\mu\nu}$, compatible with such symmetries, is the one of a perfect fluid, that is

$$T^{\mu}{}_{\nu} = \mathrm{diag}(-\rho, p, p, p), \tag{3}$$

where ρ is the energy density and p the pressure as measured in the rest frame of the fluid.

Due to the symmetries assumed, the independent equations (1) turn out to be two which are known as *Friedmann equations* and read

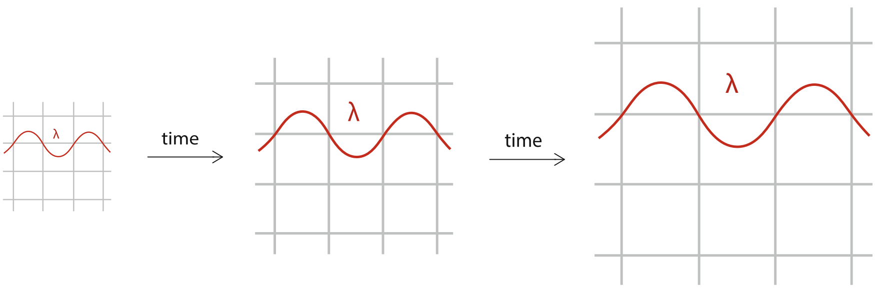


Fig. 1 The expanding Universe with a typical scale λ. The grid schematically represents comoving coordinates which do not change with time. Physical distances increase proportionally with the scale factor $a(t)$

$$H^2 = \frac{\rho}{3} - \frac{\kappa}{a^2}, \quad \frac{\ddot{a}}{a} = -\frac{1}{6}(\rho + 3p), \tag{4}$$

where dots denote derivatives with respect to the time t and we have defined the Hubble parameter as

$$H \equiv \frac{\dot{a}}{a}. \tag{5}$$

In order to extract the evolution of the scale factor $a(t)$, one must specify the type of matter and solve (4). In fact, these two equations can be combined into the continuity equation

$$\dot{\rho} + 3H(\rho + p) = 0, \tag{6}$$

which, alternatively, can be also derived from the condition of energy conservation $\nabla_\mu T^{\mu\nu} = 0$. Depending on the relation between energy density and pressure, dictated by the equation of state parameter

$$p = w\rho, \tag{7}$$

we obtain the following scaling for the energy density

$$\rho \propto a^{-3(1+w)}, \tag{8}$$

which, plugged back into (4), yields

$$a(t) \propto \begin{cases} t^{\frac{2}{3(1+w)}}, & w \neq -1 \\ e^{Ht}, & w = -1 \end{cases} \tag{9}$$

in the case of flat curvature ($\kappa = 0$). The parameter w can be assumed to be constant and depends on the specific species filling the Universe at any epoch:

- *Radiation*, or any species with dominating kinetic energy (e.g. photons or neutrinos), is characterized by $w = 1/3$. The energy density scales as $\rho \propto a^{-4}$ which implies that a Universe dominated by such type of matter expands as $a \propto t^{1/2}$.
- *Matter*, or any pressure-less species where kinetic energy is negligible with respect to the mass (e.g. baryons or dark matter), is characterized by $w = 0$. One has $\rho \propto a^{-3}$ and a Universe dominated by matter will have a scaling $a \propto t^{2/3}$.
- *Dark energy*, the mysterious component dominating the Universe nowadays, is characterized by $w = -1$ (when described by a cosmological constant) with negative pressure and constant energy density. A Universe dominated by that will expand exponentially as given by (9).

In standard cosmology, therefore, the history of the Universe is characterized by early times dominated by radiation, a moment of matter-radiation equality and subsequent domination of matter. Just recently we have entered an era in which dark

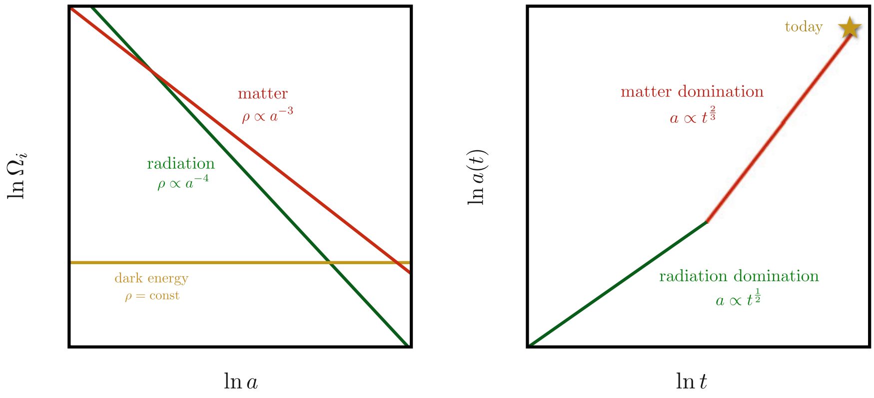


Fig. 2 Standard evolution of the energy densities (*left panel*) and the scale factor (*right panel*). According to the standard cosmological model, going back in time, the Universe becomes radiation dominated and the scale factor shrinks up to a singular point $a = 0$, commonly called "Big Bang"

energy constitutes most of the total energy in the Universe, at present 68.3% of the entire content. This evolution is shown in Fig. 2.

Finally, one may write the Friedmann equation in a form which is better for the discussion of the shortcomings affecting the standard cosmological model. By looking at (4), one may define, at any time, a critical energy density

$$\rho_c \equiv 3H^2 \tag{10}$$

corresponding to a perfect flat sectional curvature $\kappa = 0$. After normalizing all energy densities as

$$\Omega_i \equiv \frac{\rho_i}{\rho_c}, \tag{11}$$

one can rewrite (4) as

$$\boxed{\Omega \equiv \sum_i \Omega_i = 1 + \frac{\kappa}{(aH)^2}}. \tag{12}$$

2.2 Flatness Problem

In standard cosmology, an expanding Universe is naturally driven away from flatness. This can be well understood by differentiating (12), that is

$$\dot{\Omega} = H\Omega\,(\Omega - 1)\,(1 + 3w), \tag{13}$$

which can be rewritten as

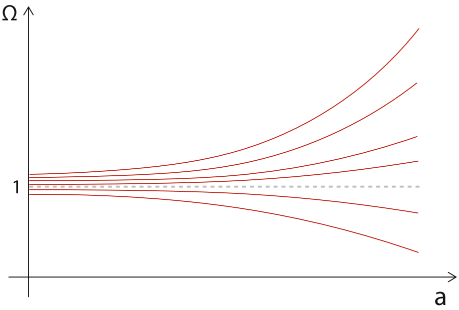


Fig. 3 Evolution of the total energy density in standard cosmology. The point $\Omega = 1$, corresponding to flat curvature, is a repeller

$$\frac{d|\Omega - 1|}{d \ln a} = \Omega |\Omega - 1|(1 + 3w). \tag{14}$$

A Universe with a growing scale factor $a(t)$ that is dominated by ordinary matter (subject to the strong energy condition $1 + 3w \geq 0$) therefore has $\Omega = 1$ as an unstable fixed point as displayed in Fig. 3.

This is exactly what happens in the standard cosmological picture where the Universe has been dominated by such type of energy from the beginning until the present time, as shown in Fig. 2. A Universe starting with generic initial curvature is driven away from flatness during its evolution. The same conclusion can be reached by looking at (12) and noticing that, in a Universe filled with radiation or matter, the sum of the energy densities Ω_i diverges from unity as the quantity $(aH)^{-1}$ increases with time.

The surprise comes with cosmological observations that suggest that the Universe today must be flat with an accuracy of 10^{-2}. This implies that, going back in time, the curvature of the Universe should have been even closer to perfect flatness: at the BBN epoch $|\Omega - 1| \lesssim 10^{-16}$, at the Planck scale $|\Omega - 1| \lesssim 10^{-64}$. Generally, such an incredible amount of fine-tuning for the initial conditions of the Universe makes physicists uncomfortable. A dynamical explanation of what we observe today would be certainly more desirable.

2.3 Horizon Problem

Given a space–time, the scale of causal physics is set by null geodesics, being the paths of photons. In an FRW Universe, with flat curvature, radial null geodesics (i.e. at constant θ and ϕ) are defined as

$$ds^2 = -dt^2 + a(t)^2 dr^2 = 0 \quad \Rightarrow \quad dr = \pm \frac{dt}{a(t)} \equiv \pm d\tau \tag{15}$$

where, in the last step, we have introduced the *conformal time* τ which simplifies the description of the causal structure of the FRW metric: the propagation of light is the same as in Minkowski space and take place diagonally (at 45°) in the (r, τ) plane.

If we assume the standard picture given by Fig. 2, the Universe was dominated by ordinary matter with state parameter $w > -1/3$ for most of its evolution and, going back in time, the scale factor $a(t)$ decreases up to the singular point $a(0) = 0$. In this case there is a maximum distance to which an observer, at time t_0, can see a light-signal sent at $t = 0$. In comoving coordinates, this is given by the so-called *comoving particle horizon*, that is

$$r_{ph} = \int_0^{t_0} \frac{dt}{a(t)} = \int_0^{a_0} (aH)^{-1} d \ln a. \tag{16}$$

If the comoving distance between two particles is greater than r_{ph}, they could have never talked to each other. Assuming (9) and integrating (16), we get

$$r_{ph} \sim a_0^{\frac{1}{2}(1+3w)} \sim (a_0 H_0)^{-1}. \tag{17}$$

Then, in an expanding Universe filled with ordinary matter, the horizon grows with time which means that comoving scales entering the horizon today have been never in causal contact before, as shown in Fig. 2.

The quantity $(aH)^{-1}$ is called *comoving Hubble radius* and determines the distance over which one cannot communicate at a given time. It basically fixes the causal structure of the space–time and its time-evolution is crucial for the particle horizon in (16).

3 Inflation

The shortcomings of standard cosmology concern the initial conditions of our Universe that require serious fine-tuning in order to reproduce what we observe today. The flatness problem can be solved by assuming that the initial value of the curvature was precisely flat. Similarly, in order to solve the horizon problem, one should imagine at least 10^6 causally disconnected spatial patches to have started their evolution exactly in the same physical conditions, in particular at the same temperature and same magnitude of perturbations. Postulating all this is possible but hardly attractive to a physicist that aims to understand the very early Universe.

In order to do better, inflation was proposed in the 1980s [7–9] to solve these problems all at once. The fundamental idea is that the primordial Universe underwent a finite phase of quasi-exponential expansion (similar to the one we are experiencing nowadays with dark energy) which changed the causal structure and how information propagates. As a bonus, one gets a physical mechanism to explain the presence of very small inhomogeneities as quantum fluctuations in the very early Universe; ultimately, these represent the seeds for the large scale structures we observe in the sky.

3.1 Basic Idea

Standard cosmology assumes that the early Universe was dominated by some form of energy satisfying the strong energy condition $\rho + 3p \geq 0$ which implies a decelerating phase of the scale factor, $\ddot{a} < 0$, as dictated by (4). This is at the core of both the flatness and horizon problems.

Inflation is nothing but inverting such a behavior and postulating a phase of accelerated expansion such as

$$\ddot{a} > 0, \tag{18}$$

which implies that the Universe was filled with some kind of matter with negative pressure, satisfying

$$\rho + 3p < 0. \tag{19}$$

The idea that, at very early times, neither matter nor radiation represented the dominant components of energy is not in contrast with any well-tested physical theory. In fact, the standard model of particles physics (SM) cannot be assumed to work up to the first moments after the Big Bang, when energies were several orders of magnitude higher than the domain of validity of the SM (which extends up to around one TeV). Inflation lives off the idea that something non-trivial might have happened due to high-energy physics.

3.2 Decreasing Hubble Radius

Interestingly, the condition (18) turns out to be equivalent to a decreasing comoving Hubble radius

$$\frac{d}{dt}(aH)^{-1} < 0, \tag{20}$$

which gives a deeper insight into the causal structure of a Universe undergoing a phase of inflationary expansion. Typical scales, being initially inside the horizon, leaves the radius of causal contact as inflation proceeds and the Hubble radius $(aH)^{-1}$ decreases. They start reentering the horizon when inflation ends, the standard cosmological evolution progresses and $(aH)^{-1}$ increases. This situation is illustrated in Fig. 4.

The horizon problem is solved if one allows for enough inflation such that also the largest scales we observe in the sky today (CMB and LSS scales) were inside the horizon at early times. Then, the CMB photons had enough time to exchange information and thermalize. Quantitatively, this means that the comoving scales of the observable Universe today $(a_0H_0)^{-1}$ must fit inside the comoving Hubble radius at the beginning of inflation $(a_iH_i)^{-1}$, that is

$$(a_iH_i)^{-1} > (a_0H_0)^{-1}. \tag{21}$$

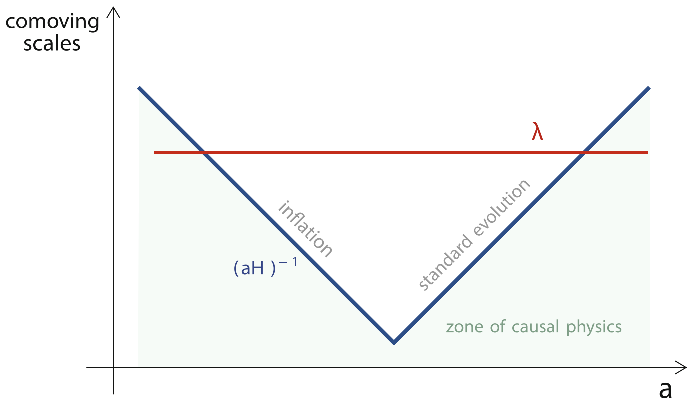


Fig. 4 The Hubbleand a typical comoving scale as a function of the scale factor. Due to the anomalous scaling of the comoving Hubble radius, which does not remain constant in time as it happens for all typical scales, the zone of causal physics change with time

The amount of inflation needed to allow for this resolution is quantified by the number of e-folds N:

$$e^N = \frac{a_{\rm end}}{a_{\rm i}}, \tag{22}$$

determined by the increase of the scale factor during inflation. A number $N \gtrsim 50$–60 suffices to explain the thermalization of the largest observational scales at present.

The flatness problem is overcome by means of the same mechanism. A decreasing comoving Hubble radius $(aH)^{-1}$ drives the value of the total energy density Ω to unity, providing a physical explanation for this apparently fine-tuned configuration. After inflation, the curvature will start diverging from $\Omega \approx 1$, as it happens in a Universe filled with ordinary matter. Interestingly, the same amount of inflation needed to solve the horizon problem is enough to explain the flatness we observe today. In fact, during inflation we have

$$\Omega - 1 = \frac{\kappa^2}{(aH)^2} \propto e^{-2N} \to 0. \tag{23}$$

The same number of e-folds quoted before would give the accuracy required for the value observed today.

3.3 *Scalar Field Dynamics and Slow-Roll Inflation*

The Einstein equations tell us that inflation should be supported by some form of matter with a negative pressure, as given by (19). However, we are still left with the issue of identifying the origin of such an incredible energy which led the scale factor to increase by an order of 10^{28}.

The simplest example is to imagine that (a small portion of) the primordial Universe is filled with a scalar field, often called *inflation* field, minimally coupled to gravity with Lagrangian

$$\mathcal{L} = \sqrt{-g}\left[\tfrac{1}{2}R - \tfrac{1}{2}g^{\mu\nu}\partial_\mu\phi\,\partial_\nu\phi - V(\phi)\right], \tag{24}$$

leading to the energy-momentum tensor

$$T_{\mu\nu} = \partial_\mu\phi\,\partial_\nu\phi - g_{\mu\nu}\left[\tfrac{1}{2}\partial^\sigma\phi\,\partial_\sigma\phi + V(\phi)\right]. \tag{25}$$

In the case of a homogeneous scalar field $\phi(t)$ filling a patch of the Universe with flat FRW metric (2), the energy density and pressure turn out to be simply

$$\rho \equiv T_{00} = \tfrac{1}{2}\dot{\phi}^2 + V(\phi), \quad p \equiv T_{ii} = \tfrac{1}{2}\dot{\phi}^2 - V(\phi). \tag{26}$$

The dynamics and interaction of the spacetime metric and scalar field is described by the two equations

$$H^2 = \frac{1}{3}\left[\frac{\dot{\phi}^2}{2} + V(\phi)\right], \quad \ddot{\phi} + 3H\dot{\phi} + V' = 0, \tag{27}$$

where primes denote derivatives with respect to ϕ. The first is simply the Friedmann equation (4), with $\kappa = 0$. The second is the equation of motion for the scalar field which is derived by varying its action. It describes a particle rolling down along its potential and subject to a friction due to the expansion term $3H\dot{\phi}$.

This region of the Universe will inflate if the state parameter $w = p/\rho < -1/3$, which is easily realizable if the potential energy dominates over the kinetic energy, that is

$$V(\phi) \gg \dot{\phi}^2. \tag{28}$$

The regime described by (28) is said *slow-roll inflation* as the field will evolve really slowly with respect to the quasi-exponential growth of the scale factor. Further, in order to have an inflationary period lasting long enough, one must ensure a small acceleration of the field and therefore impose

$$|\ddot{\phi}| \ll |3H\dot{\phi}|. \tag{29}$$

Intuitively, such a scenario is possible any time that the shape of the potential is sufficiently flat (in some measure) as it is shown in the cartoon of Fig. 5.

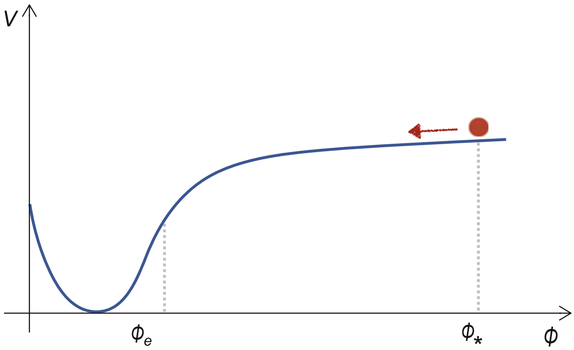


Fig. 5 Cartoon picture of a typical inflationary potential. The scalar field slowly rolls down along the shape driving the quasi-exponential expansion. Inflation ends at ϕ_e and starts at ϕ_*, at least around 60 e-foldings before the end

Within the slow-roll regime, the dynamical equations (27) become

$$H^2 \approx \frac{V(\phi)}{3} \approx \text{constant}, \quad \dot{\phi} \approx -\frac{V'}{3H}. \tag{30}$$

Given a scalar field with its potential $V(\phi)$, one can verify whether such scenario is suitable for inflation or not by calculating the so-called *slow-roll parameters*, defined as

$$\varepsilon \equiv \frac{1}{2}\left(\frac{V'}{V}\right)^2, \qquad \eta \equiv \frac{V''}{V}, \tag{31}$$

and check that

$$\{\varepsilon, |\eta|\} \ll 1, \tag{32}$$

which is equivalent to (28) and (29).

Eventually, inflation must end and give way to the standard cosmological evolution (with an increasing Hubble radius and ordinary matter domination). This happens when the conditions (32) are violated: the trajectory becomes first too steep and the inflaton eventually falls into a local minimum. The oscillations around the vacuum convert the inflationary energy into ordinary particles, within a process called *reheating*.

4 Quantum to Classical Perturbations

4.1 The Inhomogeneous Universe

The inflationary paradigm elegantly solves the standard cosmological puzzles, providing a natural explanation for the homogeneity and isotropy at large distances. However, at scales smaller than 100 Mpc, we do observe structures in form of galaxies, stars and so on. The standard cosmological theory allows us to accurately trace the evolution of such structures back in time. We are able to identify their origin in the gravitational instability of small density perturbations of a primordial plasma made up of photons and baryons, which have evolved into the large-scale structures of the present Universe.

This idea of structure formation is confirmed by the oldest snapshot we have of our Universe: the cosmic microwave background (CMB). It was produced at the time when electrons and nuclei have just recombined, around 300,000 years after the Big Bang, leaving the CMB photons to freely stream. The tiny temperature fluctuations of order $\delta T/T \sim 10^{-5}$, indicated in Fig. 6, reflect the presence of regions with slightly different densities; the wavelength of the photons is red-shifted or blue-shifted depending on the value of the local density. Indeed the properties of the CMB can be time-evolved into a forecast for the Universe that has an excellent match with our observed one.

Despite the stunning success of the theory of structure formation, we are left with some puzzling questions: *what set those initial density perturbations? Which is their fundamental origin? Why are they of the same magnitude at any scale? Why were they there at all?*

Surprisingly, inflation suggests a possible answer that is in excellent agreement with observations, thus definitively establishing itself as the leading paradigm for the understanding of the early Universe physics. This answer stems from adding quantum mechanics to the fundamental inflationary dynamics. The scalar field implementation provides once more a very useful stage in order to discuss such a physics. In fact, quantum fluctuations $\delta\phi$ are unavoidable in the homogeneous background represented by $\phi(t)$. These source metric perturbations via the Einstein equations and vice versa according to the following scheme

$$\phi(t, \mathbf{x}) = \phi(t) + \delta\phi(t, \mathbf{x}) \quad \Leftrightarrow \quad g_{\mu\nu}(t, \mathbf{x}) = g_{\mu\nu}(t) + \delta g_{\mu\nu}(t, \mathbf{x}), \tag{33}$$

where $g_{\mu\nu}(t)$ is simply the unperturbed FRW metric, as given by (2). Due to the symmetries and gauge invariance of the coupled system, the resulting physical perturbations reduce to a scalar and a tensor one (vector perturbations decay during the quasi-exponential expansion). Intuitively, quantum fluctuations excite all the light

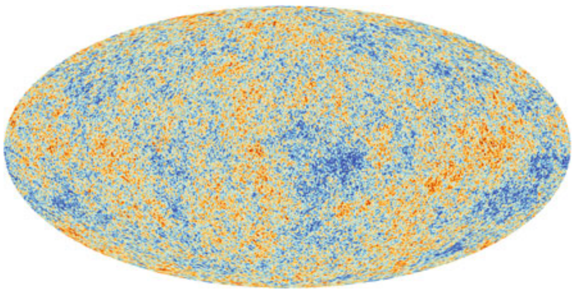


Fig. 6 The fluctuations of 1 part in 10^5 around the average temperature of $T = 2.73$ of the CMB. *Image* ESA

particles, in the minimal scenario being the inflaton and the graviton. The scalar perturbations couple to the energy density and eventually lead to the inhomogeneities and anisotropies observed in the CMB. The tensor perturbations are often referred to primordial gravitational waves. They do not couple to the density but induce polarization in the CMB spectrum [10–15]. This is considered to be a unique signature of inflation and many current and proposed experiments are searching for it in the sky.

A detailed treatment of the cosmological perturbations theory goes beyond the aim of the present lecture notes. The interested reader might consult the references [2, 3, 5]. In the following, we would like just to sketch the main consequences of a consistent quantum formulation of the inflationary paradigm. In order to simplify the discussion, we will firstly discuss the pure de Sitter and massless case. In the Sect. 5, we will focus on the proper inflationary analysis, regarded as a small deviation from the case studied here, and eventually extrapolate the significant observational parameters.

4.2 Quantum Scalar Fluctuations During Inflation

Scalar fluctuations can be fully attributed to the quantum nature of the inflaton field living in an unperturbed FRW background. This corresponds to a specific gauge (usually called *spatially flat slicing*) where metric perturbations are set equal to zero. It is a perfectly consistent choice in order to discuss the relevant physics and show how scalar fluctuations behave in an inflationary background metric. The decreasing Hubble radius $(aH)^{-1}$ will play again a crucial role, as we will see.

Let us consider the inflaton field $\phi(t, \mathbf{x})$ with a small spatial dependence as given by (33). The corresponding equation of motion is

$$\ddot{\phi} + 3H\dot{\phi} - \frac{\nabla^2}{a^2}\phi + V' = 0, \tag{34}$$

which differs from the homogeneous equation (27) of the background field $\phi(t)$ for the third extra term. We can Fourier expand the fluctuations such as

$$\delta\phi(t, \mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} \delta\phi_k(t) e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{35}$$

with $\mathbf{x}$ and $\mathbf{k}$ being respectively the comoving coordinates and momenta. Note that the Fourier modes $\delta\phi_k$ depend just on the modulo $k = |\mathbf{k}|$ because of the isotropy of the background metric. Then, we can perturb at first order (34), plug the decomposition (35) in and get

$$\delta\ddot{\phi}_k + 3H\delta\dot{\phi}_k + \frac{k^2}{a^2}\delta\phi_k = 0, \tag{36}$$

where we have neglected the additional term $V''\delta\phi_k$ due to the slow-roll conditions (32) during inflation. Equation (36) can be rewritten in a simpler form, without the Hubble friction term, once we introduce the variable

$$v_k \equiv a\delta\phi_k, \tag{37}$$

and switch to conformal time τ. This was defined by (15) and it is naturally related to the comoving Hubble radius as

$$\tau = -\frac{1}{aH}, \tag{38}$$

during a perfect exponential expansion with H constant. Then, the dynamics of the scalar perturbations can be described simply by the equation of a collection of independent harmonic oscillators

$$\boxed{\frac{d^2}{d\tau^2}v_k + \omega_k^2(\tau)v_k = 0}, \tag{39}$$

with time-dependent frequencies

$$\omega_k^2(\tau) = k^2 - \frac{2}{\tau^2} = k^2 - 2(aH)^2. \tag{40}$$

The quantization of the physical system now becomes very easy and one proceeds as in the case of the simple harmonic oscillator, following the canonical procedure. In particular, the modes v_k become nothing but the coefficients of the decomposition of the quantum operator

$$\hat{v}(\tau, \mathbf{k}) = v_k(\tau)\hat{a}_{\mathbf{k}} + v_k^*(\tau)\hat{a}_{\mathbf{k}}^\dagger, \tag{41}$$

where the creation and annihilation operators satisfy the canonical commutation relation

$$\left[\hat{a}_{\mathbf{k}}, \hat{a}_{\mathbf{k}'}^\dagger\right] = \delta^3\left(\mathbf{k} - \mathbf{k}'\right). \tag{42}$$

The quantum zero-point fluctuations are given by

$$\left\langle 0 \,\middle|\, \hat{v}^\dagger(\tau, \mathbf{k})\hat{v}(\tau, \mathbf{k}') \,\middle|\, 0 \right\rangle = |v_k(\tau)|^2\delta^3\left(\mathbf{k} - \mathbf{k}'\right) \tag{43}$$

where the vacuum is defined by $\hat{a}_{\mathbf{k}}\,|0\rangle = 0$ for any $\mathbf{k}$. Therefore, computing the quantum perturbations of the inflaton field reduces to solving the classical equation (39) and, then, extracting the time dependence of the Fourier modes $v_k(\tau)$.

The physics of the mode functions v_k, during inflation, is non-trivial and crucially depends on the fact that the comoving Hubble radius shrinks with time. In fact, fluctuations are produced on every scale λ and therefore with any momentum k.

While initially being inside the horizon, they leave the zone of causal physics at one point of the accelerated expansion, as schematically shown in Fig. 4.

One can prove that an exact solution of (39) is

$$v_k(\tau) = \alpha \frac{e^{-ik\tau}}{\sqrt{2k}} \left(1 - \frac{i}{k\tau}\right) + \beta \frac{e^{ik\tau}}{\sqrt{2k}} \left(1 + \frac{i}{k\tau}\right), \tag{44}$$

where α and β are some free parameters to be set by means of the initial conditions. These are defined at very early times, when the relevant scales were still inside the horizon. In the *sub-horizon limit* ($k \ll aH$), that is when $k|\tau| \to \infty$, the frequencies (40) become time-independent and (39) reduces to

$$\frac{d^2}{d\tau^2} v_k + k^2 v_k = 0, \tag{45}$$

basically the one of a simple harmonic oscillator. We can exploit this fact in order to get the correct normalized solution

$$\lim_{k|\tau| \to \infty} v_k = \frac{e^{-ik\tau}}{\sqrt{2k}}, \tag{46}$$

which comes from the requirement of a unique vacuum (so-called *Bunch–Davies* vacuum) being the ground state of energy. This sets $\alpha = 1$ and $\beta = 0$ in (44), thus yielding the definitive expression for the Fourier modes

$$\boxed{v_k(\tau) = \frac{e^{-ik\tau}}{\sqrt{2k}} \left(1 - \frac{i}{k\tau}\right)}. \tag{47}$$

Once we have the complete solution (47), we are particularly interested in studying when the modes leave the horizon. We would like indeed to understand how they behave after inflation and affect late time physics. How can quantum fluctuations produced during inflation source density perturbation at CMB decoupling? These events are separated by a huge amount of time where physics is very uncertain. Fortunately, something special happens as we explain below.

The *super-horizon limit* ($k \gg aH$), that is when $k|\tau| \to 0$, corresponds to the solution

$$\lim_{k|\tau| \to 0} v_k = -\frac{i}{\sqrt{2} k^{3/2} \tau}. \tag{48}$$

Since the conformal time is related to the scale factor by (15), the latter represents a growing mode $v_k \propto a$, in de Sitter background. Switching to the physical scalar perturbations by means of (37), one obtains that the amplitude $\delta\phi_k$ remains constant as long as the Hubble radius is smaller than their typical length. Modes freeze outside the horizon and this is a crucial result in order to connect the physics of the early Universe to the time when the density perturbations are created. It is a great bonus we get from inflation as we do not need to worry about the time evolution of such fluctuations for a very substantial part of the cosmic evolution.

Now we can return to (43) and properly evaluate the dimensionless *power spectrum* Δ_v^2 of the quantum fluctuations v_k, defined as

$$\left\langle 0 \left| \hat{v}^\dagger(\tau, \mathbf{k}) \hat{v}(\tau, \mathbf{k}') \right| 0 \right\rangle \equiv \frac{2\pi^2}{k^3} \Delta_v^2(k)\, \delta^3\left(\mathbf{k} - \mathbf{k}'\right). \tag{49}$$

Then, the power spectrum of the fluctuations after horizon crossing is

$$\lim_{k|\tau|\to 0} \Delta_v^2(k) = \frac{k^3}{2\pi^2}|v_k|^2 = \left(\frac{aH}{2\pi}\right)^2, \tag{50}$$

where we have used (43) in the first step while (48) and (38) in the last. Therefore, the power spectrum of the physical fluctuations of the inflaton field on super-horizon scales is

$$\boxed{\Delta_{\delta\phi}^2(k) = \left(\frac{H}{2\pi}\right)^2}, \tag{51}$$

which is scale-invariant as no k-dependence enters the expression above. Note that this result was first derived in [16], in a perfect de Sitter approximation, before inflation was proposed. A proper inflationary analysis would bring corrections of order $\mathcal{O}(\varepsilon, \eta)$.

4.3 Classical Curvature and Density Perturbations

In the previous section, we have learned that quantum fluctuations, produced during inflation, stop oscillating once they are stretched to super-horizon scales. Their amplitude freezes at some nonzero value, with scale invariant power spectrum given by (51). This situation lasts for a very long period until the point when the modes re-enter the horizon, during the standard cosmological evolution, as schematically shown in Fig. 4. At horizon re-entry, the amplitude of the modes starts oscillating again inducing the density perturbations. However, the energy density directly interacts with the gravitational potential. Therefore, *how do quantum fluctuations of the inflaton affect the metric curvature and ultimately become density perturbations?* Here, we present a very simple and heuristic derivation, mainly based on the *time-delay formalism* developed in [17].

The presence of quantum fluctuations $\delta\phi(t, \mathbf{x})$ over the smooth background $\phi(t)$ translates into local differences δN of the duration of the inflationary expansion, directly related to curvature perturbations ζ. In fact, not every point in space will end inflation at the same time thus leading to local variations of the scale factor a. Then, fluctuations $\delta\phi$ induce curvature perturbations equal to

$$\zeta = \delta N = H\frac{\delta\phi}{\dot{\phi}} = \frac{\delta a}{a}. \tag{52}$$

The corresponding dimensionless power spectrum is

$$\Delta_\zeta^2(k) = \frac{H^2}{\dot{\phi}^2}\Delta_{\delta\phi}^2(k) = \frac{H^2}{4\pi^2\dot{\phi}^2}, \tag{53}$$

which, during slow-roll, reads

$$\boxed{\Delta_\zeta^2 = \frac{1}{12\pi^2}\frac{V^3}{V'^2} = \frac{1}{24\pi^2}\frac{V}{\varepsilon}}, \tag{54}$$

where we have used (30) in the first equality and (31) in the second one.

Once inflation ends and the standard cosmological history begins, the energy density will evolve as $\rho = 3H^2$ and, then, decrease as given by (8) (the evolution is shown in Fig. 2). Local delays of the expansion lead to local differences in the density, schematically being $\delta N \sim \delta\rho/\rho$. The amplitude of the density fluctuations will be directly related to the amplitude of the curvature perturbations with power spectrum (54).

4.4 Primordial Gravitational Waves

Primordial quantum fluctuations excite also the graviton, corresponding to tensor perturbations δh of the metric. These have two independent and gauge-invariant degrees of freedom, associated to the polarization components of gravitational waves (usually denoted by h_+ and $h_\times$). One can prove that the Fourier modes of these functions satisfy an equation analogous to (36). Therefore, one may proceed identically to what done in Sect. 4.2. The dimensionless power spectrum turns out to be

$$\Delta_h^2(k) = 2 \times 4 \times \left(\frac{H}{2\pi}\right)^2, \tag{55}$$

where the factor 2 is due to the two polarizations and the factor 4 is related to different normalization.

5 Observations and Extrapolation

The last 50 years have seen extraordinary success in the development of observational techniques and in the experimental confirmation of our cosmological theories. The discovery of the CMB in 1965 [18] gave the start to a new scientific era where speculative ideas about the very early Universe have found empirical verification. Analysing this primordial light has become our fundamental tool for the investigation of the very early Universe physics.

Via CMB measurements, we are able to probe the inflationary era and set stringent constraints on the fundamental dynamical mechanism. In the language of the scalar field implementation, we can use observational inputs to impose restrictions on the form of the scalar potential $V(\phi)$. The reason why we are able to have access to such a primordial era is closely connected to the mechanism outlined in the previous section: fluctuations produced during inflation freeze outside the horizon thus providing a link between two very separated moments in time. This situation is depicted in Fig. 7.

In the following, we sketch the basic strategy to extract the inflationary parameters from the CMB data. However, as we will explain, the observational window we have access to is quite small (red region in Fig. 7) and corresponds to a short period around 50–60 e-folds before the end of inflation (this number was derived in Sect. 3 in order to account for the homogeneity and isotropy of the CMB at its largest scale). This implies that different scenarios, with very diverse potentials, may lead to the same observational consequences, as long as they agree in that CMB window. Extrapolating generic predictions, beyond the specific details of the model, and identifying related universality properties will be our primary interest. A description of inflation in terms of the number of e-folds N will turn out to be very useful.

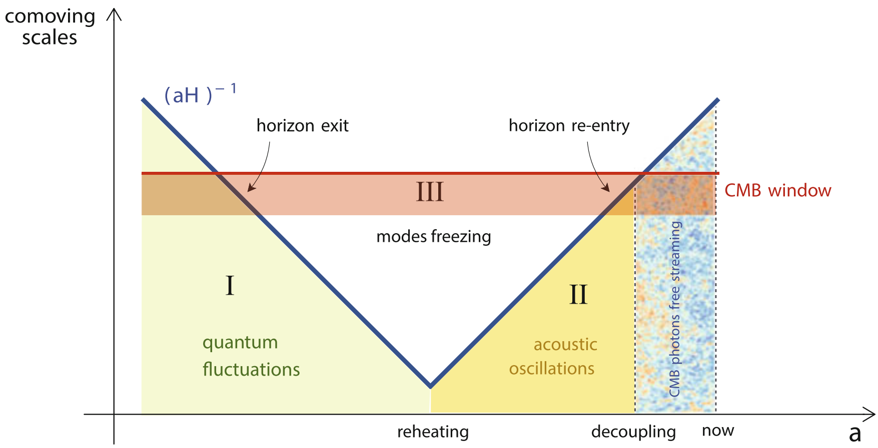


Fig. 7 Quantum fluctuations produced during inflation (*green area*) freeze at the horizon exit. They reenter the horizon after reheating thus sourcing acoustic oscillations of the plasma (*yellow part*). At decoupling time, the CMB photons freely stream towards us who measure their power spectrum just in the small *red* window

5.1 CMB and Inflationary Observables

The CMB is essentially the farthest point we can push our observations to. It is nothing but an almost isotropic 2D surface surrounding us and beyond which nothing can directly reach our telescopes. One can draw an analogy to the surface of the Sun: the inner dense plasma does not allow any light to freely stream outwards and the analysis of the last scattering photons (around 8 min old) becomes essential in order to probe the internal structure. In fact, the homogeneity and isotropy of the CMB together with its tiny and characteristic temperature anisotropy (see Fig. 6) naturally led us to study inflation in Sects. 3 and 4 and consider it as our best probe of what lies beyond that last scattering surface, around 13.4 billions years old.

The power spectrum of the temperature fluctuations in the CMB contains valuable information on the dynamics of inflation. The characteristic shape is simply dictated by the two-point correlation function of the inflaton fluctuations calculated in Sect. 4. A proper investigation of the CMB physics is required in order to understand the functional form, which goes beyond the scope of the present work (see e.g. [2, 19] for a detailed treatment). In practice, it is the so-called *transfer function* which relates the two power spectra: it contains all the information regarding the evolution of the initial fluctuations from the moment when they re-enter the horizon to the time of photon-decoupling (yellow part in Fig. 7) and, subsequently, their projection in the sky as we observe them today. The final result is the solid line of Fig. 8 with the peculiar Doppler peaks originated from the acoustic oscillations of the baryon-photon plasma. The first peak corresponds to a mode that had just time to compress once before decoupling. The other peaks underwent more oscillations and, on small scales, are damped. The high suppression of the power spectrum, at small angular scales, reflects why we are able to probe just a small window of the inflationary era.

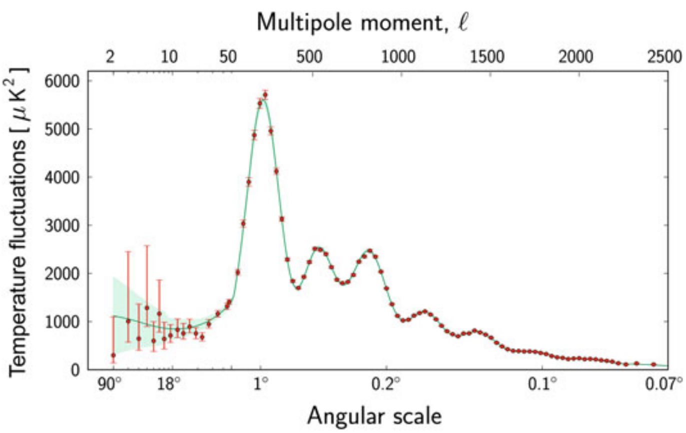


Fig. 8 Power spectrum of the CMB temperature anisotropy as measured by Planck 2015. *Image* ESA

In terms of the number of e-folds this corresponds to about $\Delta N \approx 7$. On the contrary, scales to the left of the first peak show no oscillations as they were superhorizon at the time of decoupling, and hence have not experienced any oscillations.

In Sect. 4, we have derived the power spectrum of perturbations in a perfect de Sitter ($H \approx$ const) and massless ($V'' \approx 0$) approximation. However, an appropriate inflationary analysis would bring some corrections (order slow-roll) and hence a small k-dependence. This is because, during inflation, the energy scale (set by H) will slightly change together with time and the inflaton mass is non-zero, although being very small (order η). In order to parametrize the deviation from scale-invariance, we introduce the *spectral indexes* n_s and n_t defined by

$$n_s - 1 \equiv \frac{d \ln \Delta_\zeta^2}{d \ln k}, \qquad n_t \equiv \frac{d \ln \Delta_h^2}{d \ln k}, \tag{56}$$

respectively for scalar and tensor perturbations. In terms of the slow-roll parameters, they read

$$n_s - 1 = 2\eta - 6\varepsilon, \qquad n_t = -2\varepsilon. \tag{57}$$

Furthermore, since observations probe just a limited range of k, we can express the deviation from scale-invariance by means of the power laws

$$\Delta_\xi^2(k) = \Delta_\zeta^2(k_0)\left(\frac{k}{k_0}\right)^{n_s-1}, \qquad \Delta_h^2(k) = \Delta_h^2(k_0)\left(\frac{k}{k_0}\right)^{n_t}, \tag{58}$$

where k_0 is a normalization point called *pivot scale*. Note that we have only included the first coefficients of scale-dependence; higher-order effects lead to a scale dependence of these coefficients themselves (referred to as running). Finally, the *tensor-to-scalar ratio* is defined by

$$r \equiv \frac{\Delta_h^2(k_0)}{\Delta_\zeta^2(k_0)} = 16\varepsilon, \tag{59}$$

and indicates the suppression of the power of tensor with respect to scalar modes.

5.2 Planck Data

The Planck satellite [20, 21] has mapped the Universe with unprecedented accuracy. In this way it has set stringent constraints on the parameters related to the inflationary dynamics. First of all, at $k_0 = 0.05\,\text{Mpc}^{-1}$, the experimental value for the scalar amplitude (first detected by COBE [22]) is

$$\Delta_\zeta^2(k_0) = (2.14 \pm 0.10) \times 10^{-9}. \tag{60}$$

Secondly, the deviation from perfect scale-invariance has been definitively confirmed and the scalar spectral index n_s has been measured to be

$$n_s = 0.968 \pm 0.006. \tag{61}$$

On the other hand, the value of the tensor-to-scalar ratio has been observationally bounded to be

$$r < 0.11. \tag{62}$$

These values can be read from Fig. 12 of [21] where Planck 2015 results for the spectral index and tensor-to-scalar ratio with the predictions of different inflationary models are superimposed.

5.3 Universality at Large-N

As we saw in Sect. 5.1, the window we can probe by means of CMB observations corresponds to a small portion of the inflationary trajectory. The measured values of the cosmological parameters (61) and (62) constrain the form of the scalar potential just on a limited part. This sensitive region is located around 50–60 e-folds before the end of inflation, when the modes relevant for the CMB power spectrum left the region of causal physics. The practical situation is that several scenarios can give rise to the same predictions despite the details of specific model. This situation is visually explained in Fig. 9.

In Sect. 3, we have described the inflationary background dynamics in terms of the canonical normalized field ϕ. A valid alternative description is the one in terms of the number of e-folds N, provided the relation

$$\frac{d\phi}{dN} = \sqrt{2\varepsilon}. \tag{63}$$

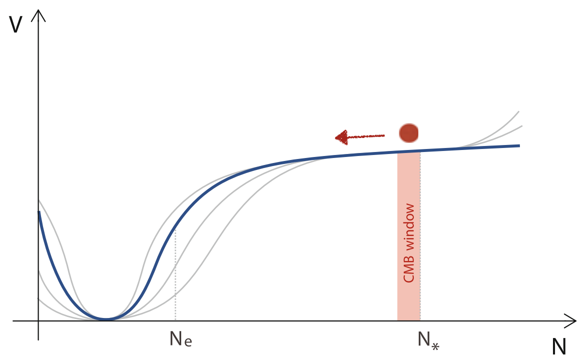


Fig. 9 Cartoon of a typical inflationary scalar potential (*blue line*) with different deviation (*grey lines*). The details of the models are different but they agree on the CMB window thus yielding identical observational predictions

This can be interpreted as a background field redefinition from ϕ, with canonical kinetic terms, to the field N with Lagrangian

$$\mathcal{L} = \sqrt{-g}\left[\tfrac{1}{2}R - \varepsilon(N)(\partial N)^2 - V(N)\right]. \tag{64}$$

Once switched to the N-formulation, we can expand the cosmological variables at large number of e-folds N, in order to keep the relevant features for observations. This approach is also motivated by the percentage-level deviation of the Planck reported value for the spectral index (61) from unity which can be interpreted as

$$n_s = 1 - \frac{2}{N}, \tag{65}$$

with N being equal to the number of e-folds between the points N_* of horizon crossing and N_e where inflation ends, that is

$$N = N_* - N_e. \tag{66}$$

These arguments naturally lead to assume the first slow-roll parameter scaling as [23–25]

$$\varepsilon = \frac{\beta}{N^p}, \tag{67}$$

where β and p are constant and we have neglected higher-order terms in $1/N$ as not relevant for observations. This simple assumption (67) yields to

$$r = \frac{16\beta}{N^p}, \qquad n_s = \begin{cases} 1 - \frac{2\beta+1}{N}, & p = 1, \\ 1 - \frac{p}{N}, & p > 1, \end{cases} \tag{68}$$

where we have discarded the case $p < 1$ as it generically not compatible with the current cosmological data.

The analysis at large-N allows us to identify the generic predictions of the cosmological scenarios with a first slow-roll parameter scaling as (67) (implications on the inflaton excursion $\Delta\phi$ studied in [26, 27]). Most of the inflationary models in literature have this property and many examples are listed in [24, 25]. Specifically, by means of (68), we can exclude a consistent region of the (n_s, r) plane and make definite predictions for our cosmological variables [24, 28]. The allowed regions can be seen in Fig. 1 of [24] where are shown the predictions of the inflationary scenarios with equation of state parameter given by (67) superimposed over the Planck data. Given the favored value of the spectral index (65), one has generically a forbidden region for value of the tensor-to-scalar ratio r. In particular, given the best fit value for n_s and the strict bound on r, we will generically expect a very low value for the tensor-to-scalar ratio, probably order 10^{-3}.

6 Inflation, Supergravity and Attractors

In the last chapter of these lecture notes, we change gears somewhat and will discuss a more theoretical underpinning of inflationary models. In particular, we consider inflation in the context of supersymmetry. Due to the presence of gravity, this naturally implies the framework of supergravity [29]. Although not observed (yet) at the energies of particle colliders, i.e. up to 1 TeV, supersymmetry is a natural ingredient of many theories of UV physics such as string theory. Given that inflation takes place at far higher energies than the Standard Model, this appears as a theoretically natural framework. Moreover, supersymmetry helps in protecting the inflaton mass from a very large contribution which would render inflation inviable: the inflaton mass is protected from being raised above the Hubble scale. This reduces the amount of necessary finetuning/modelbuilding by a few orders of magnitude. Finally, supergravity naturally includes (many) scalar fields, yielding a magnitude of possible inflaton candidates. In this chapter we will address the type of scalar potentials that arise (or can be embedded) in this set of theories, and extract inflationary predictions from these.

6.1 Flat Kähler Geometry

We will start from the simplest possible supergravity models, with $\mathcal{N} = 1$ and a single superfield Φ. Moreover, we take a flat geometry for this superfield: it is given by $ds^2 = d\Phi d\bar{\Phi}$. Note that it has an $ISO(2)$ isometry group. We will assume that inflation proceeds along the real part of Φ, which is one of the isometry directions. The canonical Kähler potential reads

$$K = \Phi\bar{\Phi}. \tag{69}$$

However, the scalar potential will be of the form $V = e^K \times \cdots$, where the dots are determined by the superpotential. For generic choices of the latter, the present Kähler potential will therefore induce order-one contributions to the second slow-roll parameter η of inflation [30]. The reason for this is the particular choice of Kähler potential: it has a rotational invariance but breaks the translational symmetry along the inflationary direction.

To remedy this, one can invoke a Kähler transformation

$$K \rightarrow K + \lambda + \bar{\lambda}, \quad W \rightarrow e^{-\lambda} W, \tag{70}$$

with holomorphic parameter λ, which leaves the entire $\mathcal{N} = 1$ theory invariant. A bringsbrins one to [31]

$$K = -\tfrac{1}{2}(\Phi - \bar{\Phi})^2, \tag{71}$$

which does respect the shift symmetry of the inflaton. As a consequence, the scalar potential does not receive order-one contributions from the Kähler potential: we have evaded the η-problem. Additional simplifications arise as both K and its first derivative K_Φ vanish along the real inflationary direction.

In this simple set-up with a single superfield, one can introduce a superpotential

$$W = f(\Phi). \tag{72}$$

Provided the function f is a real holomorphic function, it is consistent to truncate to the real part of Φ. We have therefore succeeded in identifying a possible single-field inflationary trajectory. However, its scalar potential reads

$$V = -3f(\Phi)^2 + f'(\Phi)^2, \tag{73}$$

which makes it difficult to realize e.g. the simplest inflationary model with a quadratic scalar potential in this set-up.

At this point we will follow [31] and extend the field content. In addition to the chiral superfield Φ that contains the inflaton, we introduce a second superfield S. Its role will be to "soak up" the effects of supersymmetry breaking, leaving no constraints on the inflationary potential. Indeed we will see that one can introduce arbitrary inflationary models in this way [32].

The two-superfield model reads

$$K = -\tfrac{1}{2}(\Phi - \bar{\Phi})^2 + S\bar{S}, \quad W = Sf(\Phi), \tag{74}$$

where we have added an additional piece to the Kähler potential, and moreover we have assumed that the superpotential is linear in the new field S. As inflation will take place along $\Phi - \bar{\Phi} = S = 0$, the F-term contributions read

$$D_\Phi W = 0, \quad D_S W = f, \tag{75}$$

confirming that indeed supersymmetry breaking takes place in the S-superfield. Since both K and W vanish during inflation, the potential is given by

$$V = f(\phi)^2, \tag{76}$$

where ϕ is the real part of Φ. At this point one can choose $f = m\Phi$ in the original superpotential, thus reproducing the quadratic inflationary potential from a supergravity theory. This was the original motivation and result of [31]. However, as was pointed out in [32], the same set-up allows for arbitrary real functions $f(\Phi)$. This shows that one can build an arbitrary scalar potential in this simple scenario. This implies that the predictive power of supergravity is rather limited! However, we will see in the next subsection that this conclusion changes dramatically when including curvature.

6.2 Hyperbolic Kähler Geometry and α-Attractors

Instead of a flat geometry, we now turn to the other maximally symmetric possibility. This is the hyperbolic space of the Poincaré half-plane (or disc). We will use half-plane coordinates with $Re(\Phi) > 0$. In this case the metric takes the form

$$ds^2 = 3\alpha \frac{d\Phi d\bar{\Phi}}{\left(\Phi + \bar{\Phi}\right)^2}, \tag{77}$$

whose curvature is given by

$$R_K = -\frac{2}{3\alpha}. \tag{78}$$

Note that it is negative (corresponding to hyperbolic space), and maximal symmetry implies it to be constant over moduli space. Its isometries are given by the Möbius group, which contain

- Nilpotent symmetry: $\Phi \to \Phi + ic$, corresponding to a vertical shift,
- Non-compact symmetry: $\Phi \to e^{\lambda}\Phi$, corresponding to a horizontal shift,
- Compact symmetry with a more complicated action.

The usual Kähler potential for this space is given by

$$K = -3\alpha \log(\Phi + \bar{\Phi}). \tag{79}$$

Note that it breaks all but one of the isometries: it is only invariant under the nilpotent generator. Therefore it is not invariant under shifts of the inflaton, which again we will take along the real axis of Φ. Similar to the flat case, one can however do a Kähler transformation to make this isometry explicit in the Kähler potential. In this case one finds [33]

$$K = -3\alpha \log\left[\frac{\Phi + \bar{\Phi}}{(\Phi\bar{\Phi})^{1/2}}\right], \tag{80}$$

which is invariant under the non-compact generator. Again both K and K_{Φ} vanish along the inflationary trajectory. This therefore seems to be the most natural starting point for our discussion of the curved case.

Inclusion of the supersymmetry breaking sector leads to

$$K = -3\alpha \log\left[\frac{\Phi + \bar{\Phi}}{(\Phi\bar{\Phi})^{1/2}}\right] + S\bar{S}, \tag{81}$$

while we retain the simple superpotential of the flat case:

$$W = Sf(\Phi). \tag{82}$$

Again this allows us to restrict to the real axis of Φ: the truncation to $\Phi - \bar{\Phi} = S = 0$ is consistent provided the function f is real. The single-field inflationary potential in this case reads

$$V = f^2 \left(e^{-\sqrt{\frac{2}{3\alpha}}\varphi} \right), \tag{83}$$

where φ is the canonically normalized scalar field that is related to the real part of the superfield Φ by

$$\phi = e^{-\sqrt{\frac{2}{3\alpha}}\varphi}. \tag{84}$$

Note that the curvature has a dramatic effect on the inflationary potential: the argument of the arbitrary function f is now given by an exponential of the inflaton. For a generic function f that, when expanded around $\phi = 0$, has a non-vanishing value and a slope, the resulting inflationary potential reads

$$V = V_0(1 - e^{-\sqrt{\frac{2}{3\alpha}}\varphi} + \cdots). \tag{85}$$

The potential therefore attains a plateau at infinite values of φ and has a specific exponential drop-off at finite values. At smaller values of φ, higher-order terms will come in whose form depends on the details of the function f. However, when restricting to order-one values of α, none of these higher-order terms are important for inflationary predictions: in order to calculate observables at $N = 60$, one only needs the leading term in this expansion. This means that all dependence of the function f has dropped out: the only remaining freedom is the parameter α.

In more detail, the inflationary predictions of this model are given by

$$n_s = 1 - \frac{2}{N} + \cdots, \quad r = \frac{12\alpha}{N^2} + \cdots. \tag{86}$$

The dots indicate higher-order terms in $1/N$, whose coefficients depend on the details of the function f; however, at $N \sim 60$, none of these higher-order terms are relevant for observations. The leading terms are independent of the functional freedom and only depend on the curvature of the manifold. This is what is referred to as α-attractors [34–40]: as α varies from infinity (i.e. the flat case) to order one or smaller, the inflationary predictions go from completely arbitrary (in the flat case) to the very specific values above. Turning on the curvature therefore "pulls" all inflationary models into the Planck dome in the (n_s, r) plane. The specific predictions include the magnitude of the tensor-to-scalar ratio, which naturally comes out at the permille level, as well as the scale dependence of the spectral index of scalar perturbations:

this is referred to as the running parameter, and takes the expression

$$\alpha_s = -\frac{d}{dN} n_s = -\frac{2}{N^2} + \cdots . \qquad (87)$$

Future observations will hopefully shed light on these crucial inflationary observables, and thus can (dis)prove the α-attractors framework.

7 Discussion

The topic of these lecture notes has been dual: both to provide the reader with an understanding of recent CMB observations, as well as a theoretical proposal to explain these data. We hope to have given a flavour of the excitement on the present status of observations and the theoretical expectations for possible future observations. First and foremost amongst the latter are tensor perturbations: a crucial signature of inflation, a detection of these would prove the quantum-mechanical nature of gravity as well as provide the inflationary energy scale. Moreover, depending on its value, such a detection would either disprove or lend further evidence to the inflationary models known as α-attractors.

Acknowledgments We are grateful to our collaborators John Joseph Carrasco, Mario Galante, Juan Garcia-Bellido, Renata Kallosh and Andrei Linde, who have all contributed in a major way to the results described in the last chapters. Moreover, DR would like to thank the organization of the school on "Theoretical Frontiers in Black Holes and Cosmology" in Natal, Brasil, from June 8 to 12, 2015, for a stimulating atmosphere.

References

1. A.D. Linde, Particle physics and inflationary cosmology. Contemp. Concepts Phys. **5**, 1 (1990). arXiv:hep-th/0503203
2. S. Dodelson, *Modern Cosmology* (Academic Press, Amsterdam, 2003)
3. V. Mukhanov, *Physical Foundations of Cosmology* (Cambridge University Press, Oxford, 2005)
4. S. Weinberg, *Cosmology* (Oxford University Press, Oxford, 2008)
5. D. Baumann, Inflation, in *Physics of the large and the small, TASI 09, proceedings of the Theoretical Advanced Study Institute in Elementary Particle Physics, Boulder, Colorado, USA, 1–26 June 2009* (2011) pp. 523-686, arXiv:0907.5424 [hep-th]
6. E. Hubble, A relation between distance and radial velocity among extra-galactic nebulae. Proc. Natl. Acad. Sci. **15**, 168 (1929)
7. A.H. Guth, The inflationary universe: a possible solution to the horizon and flatness problems. Phys. Rev. D **23**, 347 (1981)
8. A.D. Linde, A new inflationary universe scenario: a possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems. Phys. Lett. B **108**, 389 (1982)
9. A. Albrecht, P.J. Steinhardt, Cosmology for grand unified theories with radiatively induced symmetry breaking. Phys. Rev. Lett. **48**, 1220 (1982)

10. U. Seljak, Measuring polarization in cosmic microwave background. Astrophys. J. **482**, 6 (1997). arXiv:astro-ph/9608131
11. M. Kamionkowski, A. Kosowsky, A. Stebbins, A Probe of primordial gravity waves and vorticity. Phys. Rev. Lett. **78**, 2058 (1997). arXiv:astro-ph/9609132
12. U. Seljak, M. Zaldarriaga, Signature of gravity waves in polarization of the microwave background. Phys. Rev. Lett. **78**, 2054 (1997). arXiv:astro-ph/9609169
13. M. Zaldarriaga, U. Seljak, An all sky analysis of polarization in the microwave background. Phys. Rev. D **55**, 1830 (1997). arXiv:astro-ph/9609170
14. M. Kamionkowski, A. Kosowsky, A. Stebbins, Statistics of cosmic microwave background polarization. Phys. Rev. D **55**, 7368 (1997). arXiv:astro-ph/9611125
15. W. Hu, M.J. White, A CMB polarization primer. New Astron. **2**, 323 (1997). arXiv:astro-ph/9706147
16. T.S. Bunch, P.C.W. Davies, Quantum field theory in de sitter space: renormalization by point splitting. Proc. R. Soc. Lond. A **360**, 117 (1978)
17. A.H. Guth, S.Y. Pi, Fluctuations in the new inflationary universe. Phys. Rev. Lett. **49**, 1110 (1982)
18. A.A. Penzias, R.W. Wilson, A measurement of excess antenna temperature at 4080-Mc/s. Astrophys. J. **142**, 419 (1965)
19. W. Hu, Lecture Notes on CMB Theory: From Nucleosynthesis to Recombination. arXiv:0802.3688 [astro-ph]
20. P.A.R. Ade et al., Planck Collaboration, Planck 2015 results. XIII. Cosmological parameters. arXiv:1502.01589 [astro-ph.CO]
21. P.A.R. Ade et al., Planck Collaboration, Planck 2015 results. XX. Constraints on inflation. arXiv:1502.02114 [astro-ph.CO]
22. G.F. Smoot et al., Structure in the COBE differential microwave radiometer first year maps. Astrophys. J. **396**, L1 (1992)
23. V. Mukhanov, Quantum cosmological perturbations: predictions and observations. Eur. Phys. J. C **73**, 2486 (2013) arXiv:1303.3925 [astro-ph.CO]
24. D. Roest, Universality classes of inflation. JCAP **1401**, 007 (2014). arXiv:1309.1285 [hep-th]
25. J. Garcia-Bellido, D. Roest, Large-*N* running of the spectral index of inflation. Phys. Rev. D **89**(10), 103527 (2014) arXiv:1402.2059 [astro-ph.CO]
26. J. Garcia-Bellido, D. Roest, M. Scalisi, I. Zavala, Can CMB data constrain the inflationary field range? JCAP **1409**, 006 (2014). arXiv:1405.7399 [hep-th]
27. J. Garcia-Bellido, D. Roest, M. Scalisi, I. Zavala, Lyth bound of inflation with a tilt. Phys. Rev. D **90**(12), 123539 (2014). arXiv:1408.6839 [hep-th]
28. P. Creminelli, S. Dubovsky, D.L. Nacir, M. Simonovic, G. Trevisan, G. Villadoro, M. Zaldarriaga, Implications of the scalar tilt for the tensor-to-scalar ratio. arXiv:1412.0678 [astro-ph.CO]
29. D.Z. Freedman, A. Van Proeyen, *Supergravity* (Cambridge University Press, Cambridge, 2012)
30. E.J. Copeland, A.R. Liddle, D.H. Lyth, E.D. Stewart, D. Wands, False vacuum inflation with Einstein gravity. Phys. Rev. D **49**, 6410 (1994). arXiv:astro-ph/9401011
31. M. Kawasaki, M. Yamaguchi, T. Yanagida, Natural chaotic inflation in supergravity. Phys. Rev. Lett. **85**, 3572 (2000). arXiv:hep-ph/0004243
32. R. Kallosh, A. Linde, T. Rube, General inflaton potentials in supergravity. Phys. Rev. D **83**, 043507 (2011). arXiv:1011.5945 [hep-th]
33. J.J.M. Carrasco, R. Kallosh, A. Linde, D. Roest, Hyperbolic geometry of cosmological attractors. Phys. Rev. D **92**(4), 041301 (2015). arXiv:1504.05557 [hep-th]
34. R. Kallosh, A. Linde, Universality class in conformal inflation. JCAP **1307**, 002 (2013). arxiv:1306.5220 [hep-th]
35. S. Ferrara, R. Kallosh, A. Linde, M. Porrati, Minimal supergravity models of inflation. Phys. Rev. D **88**(8), 085038 (2013). arXiv:1307.7696 [hep-th]
36. R. Kallosh, A. Linde, D. Roest, Superconformal inflationary α-attractors. JHEP **1311**, 198 (2013). arXiv:1311.0472 [hep-th]
37. R. Kallosh, A. Linde, D. Roest, Large field inflation and double α-attractors. JHEP **1408**, 052 (2014). arXiv:1405.3646 [hep-th]

38. M. Galante, R. Kallosh, A. Linde, D. Roest, Unity of cosmological inflation attractors. Phys. Rev. Lett. **114**(14), 141302 (2015). arXiv:1412.3797 [hep-th]
39. D. Roest, M. Scalisi, Cosmological attractors from a-scale supergravity. Phys. Rev. D **92**, 043525 (2015). arXiv:1503.07909 [hep-th]
40. M. Scalisi, Cosmological α-Attractors and de Sitter Landscape. arXiv:1506.01368 [hep-th]

Index

R. Kallosh and E. Orazi (eds.), *Theoretical Frontiers in Black Holes and Cosmology*, Springer Proceedings in Physics 176,
DOI 10.1007/978-3-319-31352-8