MODIFIED GRAVITY AND COSMOLOGY: An Update by the CANTATA Network

Emmanuel N. Saridakis, Ruth Lazkoz, Vincenzo Salzano, Paulo Vargas Moniz, Salvatore Capozziello, Jose Beltrán Jiménez, Mariafelicia De Laurentis, Gonzalo J. Olmo (Editors)

Abstract

General Relativity and the \( \Lambda \)CDM framework are currently the standard lore and constitute the concordance paradigm. Nevertheless, long-standing open theoretical issues, as well as possible new observational ones arising from the explosive development of cosmology the last two decades, offer the motivation and lead a large amount of research to be devoted in constructing various extensions and modifications.

All extended theories and scenarios are first examined under the light of theoretical consistency, and then are applied to various geometrical backgrounds, such as the cosmological and the spherical symmetric ones. Their predictions at both the background and perturbation levels, and concerning cosmology at early, intermediate and late times, are then confronted with the huge amount of observational data that astrophysics and cosmology are able to offer recently. Theories, scenarios and models that successfully and efficiently pass the above steps are classified as viable and are candidates for the description of Nature.

This work is a Review of the recent developments in the fields of gravity and cosmology, presenting the state of the art, high-lightening the open problems, and outlining the directions of future research. Its realization was performed in the framework of the COST European Action “Cosmology and Astrophysics Network for Theoretical Advances and Training Actions”.

arXiv:2105.12582v2 [gr-qc] 19 May 2023
Contents

Introduction

1 Cosmophysics of modified gravity 13

2 General Relativity 26

3 Foundations of Gravity – Modifications and Extensions 32
   3.1 Preliminaries 32
   3.2 Matter Couplings 34
   3.3 The Einstein-Hilbert Action – Linear Extensions 35
   3.4 The Einstein-Hilbert Action – Nonlinear Extensions 38

Part I: Theories of Gravity 41

4 Introduction to Part I 41

5 A Flavour on $f(R)$ Theories: Theory and Observations 43
   5.1 Historia, Lux Veritatis 43
   5.2 Scalar-Tensor Theories 45
      5.2.1 Field Equations of Scalar-Tensor Gravity 45
      5.2.2 Brans-Dicke Theory 46
   5.3 Introduction to $f(R)$ Gravity 47
      5.3.1 $f(R)$ Formalisms 48
      5.3.2 $f(R)$ Gravity From a Scalar-Tensor Perspective 50
      5.3.3 Viability 51
   5.4 Background Cosmology in the Metric Formulation 56
   5.5 Scalar Perturbations: the 1+3 Formalism 57
      5.5.1 Fluid Sources 58
      5.5.2 Geometry 59
      5.5.3 Propagation and Constraint Equations 60
   5.6 Geodesic Deviation in $f(R)$ Gravity 63
      5.6.1 Formalism 63
      5.6.2 Past-directed Null Geodesics and Area Distance in $f(R)$ Gravity 64
   5.7 Gravitational Attractiveness in $f(R)$? 65
   5.8 Conclusions 66
| Chapter | Title                                                                 | Page |
|---------|----------------------------------------------------------------------|------|
| 6       | **Horndeski/Galileon theories**                                       | 68   |
| 6.1     | From Brans-Dicke to Horndeski                                        | 68   |
| 6.2     | Background Cosmology                                                  | 70   |
| 6.3     | Cosmological Perturbations                                            | 71   |
| 6.4     | Gravitational Waves Constraints                                      | 72   |
| 7       | **Massive Gravity and Bimetric Gravity**                              | 74   |
| 7.1     | Massive and Bimetric Gravity                                          | 74   |
| 7.2     | Cosmological Applications                                             | 77   |
| 8       | **Gravity in Extra Dimensions**                                      | 79   |
| 8.1     | Kaluza-Klein Model                                                    | 79   |
| 8.2     | Large Extra Dimensions                                                | 80   |
| 8.2.1   | Brane Worlds                                                          | 80   |
| 8.2.2   | Universal Extra Dimensions                                            | 83   |
| 8.2.3   | Mixed Models                                                          | 83   |
| 9       | **Non-local Gravity**                                                 | 85   |
| 9.1     | UV Nonlocal Gravity                                                   | 86   |
| 9.2     | IR Nonlocal Gravity                                                   | 91   |
| 10      | **Metric-Affine Gravity**                                             | 95   |
| 10.1    | Geometrical Objects: Torsion, Curvature and non-Metricity             | 95   |
| 10.2    | Geometrical Meaning of Torsion and Non-metricity                      | 97   |
| 10.2.1  | Geometrical Meaning of Torsion                                        | 97   |
| 10.2.2  | Geometrical Meaning of Non-Metricity                                  | 99   |
| 10.3    | Identities of non-Riemannian Geometry                                 | 100  |
| 10.3.1  | The Sources of Metric-Affine Gravity                                  | 101  |
| 10.4    | Field Equations of Metric-Affine Gravity                              | 101  |
| 10.5    | The Differential Form Formulation of Metric-Affine Gravity            | 102  |
| 10.6    | Conservation Laws and Hyperfluid Models                               | 102  |
| 10.6.1  | Hyperfluids in Cosmology                                              | 103  |
| 11      | **Geometric Foundations of Gravity**                                  | 105  |
| 11.1    | Metric-affine geometry                                                | 105  |
| 11.2    | The Geometrical Trinity                                               | 107  |
| 11.3    | Purified Gravity                                                      | 109  |
| 11.3.1  | Field Equations                                                       | 110  |
| 11.3.2  | Energy and Entropy                                                    | 110  |
| 11.3.3  | On Quantum Theory                                                    | 112  |
| 11.3.4  | Matter Coupling                                                       | 113  |
| 11.4    | Modified Gravity                                                     | 115  |
| 12      | **Palatini Theories of Gravity and Cosmology**                       | 118  |
| 12.1    | Smoothing out Cosmological Singularities                              | 119  |
| 12.2    | Inflationary Models                                                  | 120  |
| 12.3    | Background Evolution, Late-time Acceleration, and Observational Constraints | 122  |
Part III: Cosmology and Observational Discriminators

28 Introduction to Part III

29 Phenomenological Tests of Gravity on Cosmological Scales

29.1 Cosmological Tests of Gravity
29.1.1 Large Scales and the Linear Regime: Phenomenological Departures from GR
29.1.2 Cosmological Observables and Phenomenological Constraints
29.1.3 Einstein-Boltzmann Codes: from Theoretical Predictions to Data Analysis
29.1.4 Small Scales and Nonlinearities

29.2 Existing Constraints and Tensions

29.3 Upcoming Surveys and the Road Ahead

30 Relativistic Effects

30.1 Number Counts
30.2 Correlation Function
30.2.1 Estimators
30.2.2 Even and Odd Multipoles
30.3 Test of the Equivalence Principle
30.4 Conclusions

31 Cosmological Constraints From the Effective Field Theory of Dark Energy

31.1 The Effective Field Theory for Dark Energy in a Nutshell
31.2 Einstein Boltzmann Codes
31.3 Cosmological Constraints on Horndeski and GLPV Models
31.4 Astrophysical Constraints

32 The $H_0$ Tensions to Discriminate Among Concurring Models

32.1 The Effective Number of Relativistic Degrees of Freedom
32.2 Dark Energy Equation of State
32.3 Multi-parameters Extension
32.4 Early Dark Energy
32.5 Interacting Dark Energy
32.6 Modified Gravity
32.7 More specific models
32.8 Requirements: Hubble Hunter’s Guide
32.9 Standard Sirens

33 $\sigma_8$ Tension. Is Gravity Getting Weaker at Low $z$? Observational Evidence and Theoretical Implications

33.1 The $f\sigma_8$ Tension and Modified Gravity
33.1.1 Observational Evidence
33.1.2 Theoretical Implications
33.2 Evolving $G_{eff}$ and the Pantheon SNeIa Dataset
33.3 Constraints on Evolving $G_{eff}$ from Low $l$ CMB Spectrum and the ISW Effect
33.4 Conclusions
| Chapter                                                                 | Page |
|------------------------------------------------------------------------|------|
| 34 Testing Gravity with Standard Sirens: Challenges and Opportunities  | 339  |
| 34.1 Gravitational Wave Propagation Beyond General Relativity          | 339  |
| 34.2 Standard Sirens                                                   | 340  |
| 34.3 The Speed of GWs                                                  | 341  |
| 34.3.1 Constraints After GW170817                                      | 342  |
| 34.4 GW Luminosity Distance                                            | 344  |
| 34.5 GW Oscillations                                                   | 346  |
| 34.6 Future Prospects                                                  | 347  |
| 34.6.1 Theoretical Challenges                                           | 348  |
| 34.6.2 Observational Opportunities                                     | 349  |
| 35 Testing the Dark Universe with Cosmic Shear                         | 351  |
| 35.1 2D, Tomographic and 3D Weak Lensing                               | 351  |
| 35.2 Current Data and Forecasts on Horndeski Gravity                   | 354  |
| 35.3 Higher-order Statistics and Lensing Peak Counts                   | 357  |
| 35.4 Machine Learning and the Dark Universe                            | 357  |
| 36 Galaxy Clusters and Modified Gravity                                | 360  |
| 36.1 What Makes Galaxy Clusters Interesting for Testing Gravity?       | 360  |
| 36.2 Consistency Conditions Based on the Mass Profiles of Galaxy Clusters | 360  |
| 36.2.1 Generalities                                                    | 360  |
| 36.2.2 Probes Based on Mass Profiles from Galaxy Kinematics and Lensing| 361  |
| 36.2.3 Probes Based on Thermal and Lensing Mass Profiles               | 363  |
| 36.3 A Brief Discussion on Systematics                                 | 364  |
| 36.4 Future Outlook                                                    | 367  |
| 37 Probing Screening Modified Gravity with Non-linear Structure Formation | 368   |
| 37.1 Theoretical Models                                                | 369  |
| 37.1.1 Chameleon-$f(R)$ Gravity                                        | 370  |
| 37.1.2 Symmetron                                                       | 370  |
| 37.2 Efficiency of Screening Mechanisms                                | 371  |
| 37.2.1 Solar System Constraints                                        | 371  |
| 37.2.2 Simulations                                                     | 372  |
| 37.2.3 Results                                                         | 373  |
| 37.3 Distribution of Fifth Force in Dark Matter Haloes                 | 374  |
| 37.4 The Matter and the Velocity Power Spectra                         | 374  |
| 37.5 The Dynamical and Lensing Masses                                  | 375  |
| 37.6 Thermal Versus Lensing Mass Measurements                          | 376  |
| 37.6.1 Including the Non-thermal Pressure Component                   | 377  |
| 37.7 Modelling Void Abundance in Modified Gravity                     | 379  |
| 37.7.1 Linear Power Spectrum                                           | 379  |
| 37.7.2 Spherical Collapse                                              | 381  |
| 37.7.3 Void Abundance Function                                         | 382  |
| 37.7.4 Voids from Simulations                                         | 385  |
| 37.7.5 Results                                                         | 386  |
| 37.8 Conclusions and Perspectives                                     | 389  |
Preface

The dawn of the 21st century came with very positive prospects for gravity, cosmology and astrophysics. Technological progress made it possible for cosmology to enter to its adulthood and become a precision science, both for its own sake as well as for being the laboratory of gravity, which can now be accurately tested and investigated in scales different than the earth ones. As a result, the opinion that cosmology is one of the main directions that will lead to progress in physics in the near future, is now well established.

“Cosmology and Astrophysics Network for Theoretical Advances and Training Actions” (CANTATA) is a COST European Action established in 2015 in order to contribute to the front of research in the fields of gravity, cosmology and astrophysics. It involves Institutions from 26 European countries, as well as from 5 countries abroad. CANTATA Collaboration has a variety of interests, which include: i) the classification and definition of theoretical and phenomenological aspects of gravitational interaction that cannot be enclosed in the standard lore scheme but might be considered as signs of alternative theories of gravity, ii) the confrontation of the theoretical predictions with observations at both the background and the perturbation levels, iii) the production of numerical codes to simulate astrophysical and cosmological phenomena, iv) the construction of self-consistent models at various scales and the investigation of the features capable of confirming or ruling out an effective theory of gravity, v) the study of how extended and modified theories of gravity emerge from quantum field theory and how mechanisms produced by the latter may explain cosmological dynamics. This Review presents the recent developments in the above fields.
### Conventions

| Greek small letters $\alpha, \mu, \nu$, ... | space-time coordinates indices |
| Latin small letters $i, j, k$, ... | space coordinates indices |
| Latin capital indices $A, B$, ... | tangent space indices |
| (only in chapters 8 and 9) | (only in chapters 8 and 9) |
| D-dimensional coordinate indices | D-dimensional coordinate indices |
| metric tensor | metric tensor |
| metric signature | metric signature |
| Levi-Civita connection | Levi-Civita connection |
| Riemann curvature tensor | Riemann curvature tensor |
| Ricci tensor | Ricci tensor |
| Ricci scalar | Ricci scalar |
| Einstein tensor | Einstein tensor |
| covariant derivative | covariant derivative |
| d’Alembertian operator | d’Alembertian operator |
| anti-symmetry | anti-symmetry |
| symmetricity | symmetricity |
| 4-dimensional Friedmann-Lemaître-Robertson-Walker (FLRW) line-element | 4-dimensional Friedmann-Lemaître-Robertson-Walker (FLRW) line-element |
| conformal time | conformal time |
| cosmic time derivative | cosmic time derivative |
| conformal time derivative | conformal time derivative |
| maximally symmetric 3-dimensional space-like hyper-surfaces metric | maximally symmetric 3-dimensional space-like hyper-surfaces metric |
| grad operator on the 3-dimensional space-like hyper-surfaces | grad operator on the 3-dimensional space-like hyper-surfaces |
| Laplacian operator | Laplacian operator |
| Newtonian gauge scalar metric perturbations | Newtonian gauge scalar metric perturbations |
| energy-momentum tensor | energy-momentum tensor |
| of the Lagrangian density $\mathcal{L}$ | of the Lagrangian density $\mathcal{L}$ |
| gravitational constant | gravitational constant |
| natural units | natural units |

List of notational conventions used in this manuscript, unless otherwise stated.
(cont.) List of notational conventions used in this manuscript, unless otherwise stated.
3. Foundations of gravity – modifications and extensions

Christian G. Böhmer

3.1 Preliminaries

Shortly after the formulation of General Relativity was completed in 1915, it became clear that this theory could be extended in various different ways. The theory of General Relativity (GR) is formulated using the language of differential geometry, which - in itself - was a relatively new topic of research in mathematics at that time. The geometrical setting used by Einstein consists of a four-dimensional Lorentzian manifold, equipped with a metric structure \( g \) and a covariant derivative \( \nabla \), or equivalently, a connection \( \Gamma \). This derivative is assumed to be metric compatible and torsion free, which then uniquely determines the connection coefficients to be the Christoffel symbol components \( \Gamma \).

Let us briefly dissect these assumptions to get an immediate idea of how one could modify GR. To begin with one does not have to restrict the geometry to four dimensions. Kaluza and Klein are credited with suggestions along those lines [66,67]. The use of four dimensions relies on our experience of three spatial dimensions and a sense of time, which acts as the fourth dimension, commonly denoted as the zeroth coordinate. One could now assume that there exist other spatial dimensions that have not yet been observed. Hypothetically, there could also be other time-like dimensions. These initial ideas were motivated by the idea of geometrically unifying the different physical theories known at that time; see [68,69] for a comprehensive review on so-called unified field theories. It is probably fair to say that String Theory has followed that path, point particles (points are zero dimensional objects) being replaced by strings (strings or curves are one-dimensional objects). Bosonic string theory is formulated in a 26-dimensional Lorentzian manifold, while superstring theory is formulated in 10 dimensions. These extra dimensions are dealt with by compactification, which means ‘rolling up’ those dimensions in such a way that they are very small, hence, effectively leading to a four-dimensional space in which Special Relativity and General Relativity are formulated.

The next generalisation concerns the connection \( \hat{\Gamma} \) which neither has to be metric compatible nor torsion free. Both, non-metricity and torsion have neat geometrical interpretations [6,70]; one speaks of an affine connection. Let us consider an infinitesimal parallelogram that is constructed by parallelly transporting two given vectors along each other. There is no \textit{a priori} guarantee that this process does give a closed parallelogram. Indeed, torsion represents the failure of this infinitesimal parallelogram to close. In order to understand the effect of non-metricity on the manifold, let us consider a null vector \( u^\mu \), which means it satisfies \( g_{\mu \nu}u^\mu u^\nu = 0 \). If the covariant derivative of the metric tensor does not vanish, then this vector may no longer be null when parallelly transported. In particular, the light cone structure would no longer be invariant.
under parallel transport. However, neither the lack of closed infinitesimal parallelograms nor the non-invariance of the light cone structure under parallel transport are reason enough to discard these geometrical concepts from a physical point of view. In the end, any theoretical model of the gravitational field will make certain predictions that an experiment can either verify or falsify.

The entire discussion up to now was independent of the Einstein field equations; it merely assumed that there exists a gravitational theory that can be formulated using differential geometry. Let us now start making some connections between the mathematical formulation and the physical content of our theories. It is a well-established everyday fact that light travels along straight lines, and so do massive particles in the absence of external forces. In classical physics one would refer to these as Fermat’s principle and Newton’s first law, respectively. In the context of differential geometry things start to get interesting now, as a manifold equipped with a metric structure and an affine connection gives rise to two distinct curves: geodesics and autoparallels. Geodesics are the shortest possible curves between two fixed end points, autoparallels are the straightest possible curves between two points. Geodesics are generally introduced by studying curves \( C \) with tangent vectors \( T^\mu = \frac{dX^\mu}{d\lambda} \) such that the quantity

\[
s = \int_{\lambda_1}^{\lambda_2} \sqrt{g_{\mu\nu} T^\mu T^\nu} \, d\lambda,
\]

is extremised. Here, \( X^\mu(\lambda) \) are the local coordinates of the curve and \( \lambda \) is the (affine) parameter of the curve. This yields the familiar geodesic equations

\[
\frac{d^2 X^\mu}{d\lambda^2} + \Gamma^\mu_{\sigma\tau} \frac{dX^\sigma}{d\lambda} \frac{dX^\tau}{d\lambda} = 0 \quad \Leftrightarrow \quad \frac{dT^\mu}{d\lambda} + \hat{\Gamma}^\mu_{\sigma\tau} T^\sigma T^\tau = 0.
\]

(3.1)

(3.2)

It needs to be emphasised that the geodesic equation, defined via this variational approach, depends on the Christoffel symbol components \( \Gamma^\mu_{\sigma\tau} \) only. This follows from the fact that (3.1) is independent of the affine connection - that is, it depends on the metric tensor and the curve.

On the other hand, we can introduce the straightest possible curves or autoparallels. Let us again consider a curve \( C \) with tangent vector \( T^\mu \), then the vector \( V^\sigma \) is parallelly transported along this curve if \( T^\mu \nabla_\mu V^\sigma = 0 \). The notion of parallel transport allows us to consider curves (defined indirectly) whose tangent vectors are parallelly transported along themselves, the tangent vector is kept as parallel as possible along the curve, hence autoparallel. Using the chain rule and the definition of covariant differentiation, the autoparallel equations are given by

\[
T^\mu \nabla_\mu T^\sigma = 0 \quad \Leftrightarrow \quad \frac{dT^\mu}{d\lambda} + \hat{\Gamma}^\mu_{\sigma\tau} T^\sigma T^\tau = 0.\]

(3.3)

The key difference between (3.2) and (3.3) is that two different connections appear in these equations, while their form is identical. It is clear that (3.3) depends on the symmetric part of the connection, since one can exchange \( T^\sigma \) and \( T^\tau \); however, it is important to state

\[
\hat{\Gamma}^\mu_{(\sigma\tau)} \neq \Gamma^\mu_{\sigma\tau},
\]

(3.4)

which means that the symmetric part of the affine connection is not the Christoffel symbol. This symmetric part contains the Christoffel symbol, but it also depends on torsion and non-metricity, should these be present.

General Relativity is special in the sense that the shortest possible lines coincide with the straightest possible lines. These considerations have practical implications. By studying the

\[\text{If the affine connection differs from the Christoffel symbol components by a totally skew-symmetric piece, geodesics and autoparallels would also coincide.}\]
geometric properties of trajectories of test particles one can, in principle, determine whether the connection contains contributions other than those from the Christoffel symbol; see the footnote again.

In its standard formulation, the dynamical variables of General Relativity are the 10 metric functions $g_{\mu\nu}$, which are the solutions of the ten Einstein field equations

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \kappa^2 T_{\mu\nu} .$$

(3.5)

Here, $G_{\mu\nu}$ is the Einstein tensor, $R_{\mu\nu}$ is the Ricci tensor, $R$ is the Ricci scalar, and $T_{\mu\nu}$ stands for the metric energy-momentum-stress tensor. This is a true tensor equation in the sense that it is valid for all coordinate systems and hence diffeomorphism invariant. In four spacetime dimensions one has four coordinates, which can be arbitrarily changed, which implies that the Einstein field equations can be viewed as six independent equations. When a Hamiltonian analysis is performed on these equations, one finds four primary constraints, thereby reducing the number of propagating degrees of freedom of this theory to two (10 metric components minus four coordinate transformations minus four primary constraints). When General Relativity is introduced in a first course - see for instance [71] - the Einstein field equations are motivated by comparing Newton’s equations with geometrical equations, which may describe the same physics.

A more elegant approach, which somewhat lacks physical motivation from first principles, is the variational approach. The field equations can also be derived from the so-called Einstein-Hilbert action

$$S_{\text{EH}} = \frac{1}{2\kappa^2} \int g^{\mu\nu} R_{\mu\nu} \sqrt{-g} \, d^4x = \frac{1}{2\kappa^2} \int R \sqrt{-g} \, d^4x .$$

(3.6)

$$S_{\text{matter}} = \int L_{\text{matter}}(g, \phi, \nabla \psi) \, d^4x = \int L_{\text{matter}}(g, \phi, \nabla \psi) \sqrt{-g} \, d^4x ,$$

(3.7)

$$S_{\text{total}} = S_{\text{EH}} + S_{\text{matter}} ,$$

(3.8)

where one varies with respect to the dynamical variable $g_{\mu\nu}$. Here, $g$ is the determinant of the metric tensor $g_{\mu\nu}$, so that $\sqrt{-g} \, d^4x$ is the appropriate volume element when integrating over the manifold. The matter fields are denoted by $\psi$ and the matter Lagrangian can depend on derivatives of the matter fields. The gravitational coupling constant $\kappa^2$ is given by $\kappa^2 = 8\pi G/c^4$.

It is through this variational approach that one can introduce and motivate various gravitational theories, which can be seen as extensions or modifications of the original theory. In the following sections, some of these ideas will be discussed.

### 3.2 Matter Couplings

Before discussing other gravitational theories, let us briefly mention the issue of matter couplings. This is, of course, of crucial importance as gravity is universal and is the dominant interaction in the macroscopic world [72]. Lagrangians, which describe scalars (spin 0 particles) or spinors (spin 1/2 particles), typically depend on the fields and their first derivatives, thereby giving rise to equations of motion of at most second order. This also holds for Yang-Mills theories; however, we will focus our discussion on scalars and spinors for now. When the scalar or spinor field actions, formulated in Minkowski space, are formulated on an arbitrary manifold, one replaces the Minkowski metric $\eta$ with an arbitrary metric $g$. The partial derivatives are replaced with covariant derivatives. In the scalar field case one simply has $\nabla_\mu \phi = \partial_\mu \phi$, while for spinorial fields the covariant derivative also depends on the connection and we have $\nabla_\mu \psi \neq \partial_\mu \psi$, with $\psi$ standing for a spinor field. The immediate consequence of this is that theories in which
variations with respect to the connection are considered will contain source terms when spinor fields are taken into account. Since protons, neutrons and electrons are all spin $1/2$ particles, this is an important issue to keep in mind. Finally, when considering Yang-Mills theories we recall that the currents which act as the source terms are conserved and couple to the gauge fields. General Relativity can also be formulated as gauge theories; however, it is not in the form of a typical Yang-Mills theory, see [72–74]. The above mentioned approach is often referred to as the principle of minimal coupling, however, many other coupling terms are in principle possible. There are Pauli-type terms and Jordan-Brans-Dicke-type terms where geometrical quantities like the Riemann tensor or the Ricci tensor couple to the matter fields; these mainly appear in phenomenological models or when models with symmetry breaking are concerned. Note that any coupling term which involves curvature will necessarily vanish in Special Relativity, therefore such terms require strong gravitational fields to affect the theory and to be, in principle, observable.

3.3 The Einstein-Hilbert Action – Linear Extensions

On manifolds where the connection is metric compatible and torsion free, the Einstein-Hilbert action is the unique action that is linear in a curvature scalar, and the Ricci scalar is the unique linear curvature scalar. In more general spaces with torsion and non-metricity, one can also construct the scalar $\varepsilon^{\mu\nu\kappa\lambda} \hat{R}_{\mu\nu\kappa\lambda}$, which does not vanish in general. This term appears in the so-called Palatini action of General Relativity or the Holst action. It becomes important in the context of Loop Quantum Gravity, where it appears in Ashtekar’s choice of variables, which allows the formulation of GR as a Yang-Mills type theory [75–80]. Let us return to the Einstein-Hilbert action (3.8) for now.

The Riemann curvature tensor and the Ricci tensor can be defined by the affine connection $\hat{\Gamma}$ alone, without requiring the metric tensor. To make this explicit, it is often written as $\hat{R}_{\mu\nu}$. Hence, in affine spacetimes the Einstein-Hilbert action can be generalised simply by writing

$$S = \frac{1}{2\kappa^2} \int g^{\mu\nu} \hat{R}_{\mu\nu} \sqrt{-g} \, d^4x, \quad (3.9)$$

$$S_{\text{matter}} = \int L_{\text{matter}}(g, \phi, \nabla \psi) \sqrt{-g} \, d^4x, \quad (3.10)$$

$$S_{\text{total}} = S_{\text{EH}} + S_{\text{matter}}. \quad (3.11)$$

One now considers the metric tensor $g$ and the connection $\hat{\Gamma}$ as a priori independent dynamical variables. The matter action also depends on the connection through the covariant derivative; this is completely consistent with the principle of minimal coupling used in General Relativity. This principle states that first one writes all equations covariantly in a four-dimensional Lorentzian manifold, flat Minkowski space, then one replaces all partial derivatives with covariant derivatives and all Minkowski metric tensors with arbitrary metric tensors; see Sec. 3.2.

If we assume that the matter part of the action does not depend on the connection and we make independent variations with respect to the metric and connection, we arrive at Einstein’s theory of General Relativity. This is often referred to as the Palatini variation; however, as we will soon discuss, things become more subtle when geometries are more general.

In many ways the most natural generalisation of General Relativity is constructed when beginning with (3.11) and allowing the matter part of the action to depend on the matter fields, the metric and an independent connection. When we now compute the variations with respect to the metric and the independent connection, we arrive at two sets of field equations. Variations with respect to the metric yield equations that resemble the Einstein field equations,
while variations with respect to the connection give a new set of field equations which determine
the connection. The source term that appears in the latter is often referred to as the hyper-
momentum $\Delta^\lambda_{\mu\nu}$, following a commonly used notation [6]. As the affine connection has no
symmetries, the hyper-momentum tensor has, in general, 64 independent components in four
dimensions.

Let us now discuss how we can connect these different theories back to General Relativity,
using a mathematically consistent approach. The perhaps most elegant way to do it is through the
introduction of Lagrange multipliers in the total action (3.11), so that this action is subsequently
extremised subject to constraints. These constraints are introduced so that the geometrical
properties of the manifold are controlled. More explicitly, let us, for the time being, extract
General Relativity within the framework of metric affine theories. Recall that the two key
geometrical assumptions are a metric compatible and torsion-free covariant derivative. In the
language of constraints we would write

$$S_{\text{GR}} = \frac{1}{2\kappa^2} \int \left\{ g^{\mu\nu} \hat{R}_{\mu\nu} + \lambda^{\mu\nu\lambda}_{(1)} T_{\mu\nu\lambda} + \lambda^{\mu\nu\lambda}_{(2)} Q_{\mu\nu\lambda} \right\} \sqrt{-\hat{g}} \, d^4 x, \quad (3.12)$$

$$S_{\text{total}} = S_{\text{GR}} + S_{\text{matter}}, \quad (3.13)$$

where $T_{\mu\nu\lambda}$ is the torsion tensor and $Q_{\mu\nu\lambda}$ is the non-metricity tensor. Here, $\lambda_{(1)}$ and $\lambda_{(2)}$ are two
Lagrange multipliers which ensure that the affine connection will become the usual Christoffel
symbol. Clearly, variations with respect to $\lambda_{(1)}$ give $T_{\mu\nu\lambda} = 0$, while variation with respect to
$\lambda_{(2)}$ yields $Q_{\mu\nu\lambda} = 0$.

A popular extension of General Relativity is the so-called Einstein-Cartan theory, which was
proposed in the 1920s by Cartan; see [70]. Within the above framework, Einstein-Cartan theory
is simply defined by

$$S_{\text{EC}} = \frac{1}{2\kappa^2} \int \left\{ g^{\mu\nu} \hat{R}_{\mu\nu} + \lambda^{\mu\nu\lambda}_{(2)} Q_{\mu\nu\lambda} + \right\} \sqrt{-\hat{g}} \, d^4 x, \quad (3.14)$$

$$S_{\text{total}} = S_{\text{EC}} + S_{\text{matter}}. \quad (3.15)$$

The only difference with respect to General Relativity is the possible presence of torsion, which
is no longer assumed to be zero. A natural source term for torsion would be fermions; their
action depends on the connection, and hence variations with respect to the connection lead
to source terms for torsion. A peculiar property of Einstein-Cartan theory is that the field
equations for torsion are algebraic; they do not contain derivatives of the torsion tensor. This
immediately implies that torsion cannot propagate, and consequently, only regions of spacetime
that contain sources of torsion can contain torsion. This is in stark contrast to curvature, a fact
well-known in GR. The Schwarzschild solution, for instance, is a vacuum (source-free) solution of
the Einstein field equations, yet contains curvature. Likewise, gravitational waves can propagate
through otherwise empty regions of space-time; torsional waves in this sense do not exist in
Einstein-Cartan theory.

If we recall that Minkowski space is the setting of Special Relativity, we can of course also
include this using the above approach, namely we consider the following theory

$$S_{\text{Mink}} = \frac{1}{2\kappa^2} \int \left\{ g^{\mu\nu} \hat{R}_{\mu\nu} + \lambda^{\mu\nu\lambda}_{(0)} \hat{R}_{\mu\nu\lambda\kappa} + \lambda^{\mu\nu\lambda}_{(1)} T_{\mu\nu\lambda} + \lambda^{\mu\nu\lambda}_{(2)} Q_{\mu\nu\lambda} \right\} \sqrt{-\hat{g}} \, d^4 x. \quad (3.16)$$

Minkowski space is the unique space that has vanishing torsion, vanishing non-metricity and is
globally flat.

However, what makes this approach, using constraints, particularly useful is the ability to
systematically study a variety of theories in a uniform setting; see also [4]. Let us now discuss
a theory, which is equivalent to General Relativity but is formulated rather differently. It was
noted in the 1920s by Einstein and others that there exists a formulation of General Relativity
based solely on the torsion tensor; this theory is now known as the Teleparallel Equivalent of
General Relativity (TEGR).

Start as before, within the setting of metric-affine theories where the connection is assumed
to be fully independent of the metric, we can define the Teleparallel Equivalent of General
Relativity by

\[ S_{TEGR} = \frac{1}{2\kappa^2} \int \left\{ g^{\mu\nu} \tilde{R}_{\mu\nu} + \lambda_{(0)}^{\mu\nu\lambda\kappa} \tilde{R}_{\mu\nu\lambda\kappa} + \lambda_{(2)}^{\mu\nu\lambda\kappa} Q_{\mu\nu\lambda\kappa} \right\} \sqrt{-g} \, d^4x. \]  

(3.17)

We note that the constraints force the connection to be metric compatible (no non-metricity,
\( Q_{\mu\nu\lambda} = 0 \)) and make the manifold globally flat, \( \tilde{R}_{\mu\nu\lambda\kappa} = 0 \) everywhere. There is now a conceptual
issue to understand: Is the theory so defined non-trivial? This is a natural question to ask, as
we know that only Minkowski space satisfies \( R_{\mu\nu\lambda\kappa} = 0 \) everywhere in the usual GR setting. To
answer this, let us begin by introducing the so-called contortion tensor \( K \), defined by

\[ \tilde{\Gamma}_{\sigma\tau}^\mu = \Gamma_{\sigma\tau}^\mu + K_{\sigma\tau}^\mu. \]  

(3.18)

The contortion tensor simply contains all the information of the connection that is not part
of the Christoffel symbol. In other words, it contains the deviations from the standard GR
framework. When we compute the Riemann curvature tensor for the full connection \( \tilde{\Gamma}_{\sigma\tau}^\mu \) and
express the result using the Christoffel symbol and the contortion tensor, we arrive at the neat
result

\[ \tilde{R}_{\nu\mu\lambda\kappa} = R_{\nu\mu\lambda\kappa} + \nabla_{\nu} K_{\mu\lambda}^\kappa - \nabla_{\mu} K_{\nu\lambda}^\kappa + K_{\nu\rho}^\kappa K_{\mu\lambda}^\rho - K_{\nu\rho}^\kappa K_{\mu\lambda}^\rho, \]  

(3.19)

which means that the curvature tensor splits into two parts: the Levi-Civita part \( R_{\nu\mu\lambda\kappa} \), which
is constructed using the Christoffel symbols components only, and a second part that depends
on the contortion tensor. One normally defines the torsion tensor to be the skew symmetric
part of the affine connection \( T_{\nu\mu}\kappa = (\tilde{\Gamma}_{\nu\mu}^\kappa - \Gamma_{\nu\mu}^\kappa) / 2 \) so that, in spaces where \( Q_{\mu\nu\lambda} = 0 \), one has the simple relation

\[ T_{\nu\mu}\kappa = \frac{1}{2} \left( K_{\nu\mu}\kappa - K_{\nu\mu}\kappa \right). \]  

(3.20)

Therefore, there is a linear relation between torsion and contortion. Going back to (3.17), which
imposes the constraint \( \tilde{R}_{\mu\nu\lambda\kappa} = 0 \), we can now attempt to understand this condition in view of
(3.19). Is it always possible to choose a contortion or torsion tensor for a given curvature tensor
\( R_{\nu\mu\lambda\kappa} \) such that the full metric-affine curvature vanishes?

The answer to this question is affirmative; this choice can be made but it requires a little bit
more mathematics. Let \( e^A_\mu \) be four linearly independent orthonormal co-vector or co-frame fields,
often called tetrads, which allow us to write the metric tensor at any point of the manifold as

\[ g_{\mu\nu} = e^A_\mu e^B_\nu \eta_{AB}. \]  

(3.21)

These fields can be used to define the frame components of any vector via \( V^A = e^A_\mu V^\mu \). Moreover,
one needs to introduce the spin connection \( \tilde{\omega}_{\mu}^A B \), which is defined through the vanishing of the
covariant derivative of the tetrad

\[ \nabla_{\mu} e^A_\nu = 0 \iff \partial_\mu e^A_\nu + \tilde{\omega}_{\mu}^A B e^B_\nu - \tilde{\Gamma}_{\mu\nu} e^A_\sigma = 0. \]  

(3.22)
We can now express the complete (Levi-Civita plus torsional contributions) Riemann curvature tensor using either the connection or the spin connection. In the latter case we have

\[ \hat{R}^{AB\mu\nu} = \partial_\mu \hat{\omega}^{A\mu\nu} + \partial_\nu \hat{\omega}^{A\mu\nu} - \partial_\mu \hat{\omega}^{A\mu\nu} + \hat{\omega}^{A\mu\nu} \hat{\omega}^{C\mu\nu} - \hat{\omega}^{A\mu\nu} \hat{\omega}^{C\mu\nu}, \]  

(3.23)

which now allows us to make the following useful observations. If we choose \( \hat{\omega}^{A\mu\nu} = 0 \) everywhere, then \( \hat{R}^{AB\mu\nu} = 0 \) everywhere. This means that, applying the decomposition (3.18) to the spin connection, we can write

\[ \hat{\omega}^{A\mu\nu} = \omega^{A\mu\nu} + K^{\mu\nu AB}, \]  

(3.24)

which implies that \( \hat{\omega}^{A\mu\nu} = 0 \), or equivalently \( \omega^{A\mu\nu} = -K^{\mu\nu AB} \). Therefore, for any Levi-Civita spin connection \( \omega^{A\mu\nu} \) there exists a contortion tensor \( K^{\mu\nu AB} \) such that \( \hat{\omega}^{A\mu\nu} = 0 \). This result was found by Weitzenböck, who noted that this connection can always be constructed, given a tetrad. Putting this result back into (3.19) allows us to rewrite the complete Ricci scalar in terms of the Levi-Civita part and a torsion part; this gives

\[ \hat{\omega}^{A\mu\nu} = \omega^{A\mu\nu} + K^{\mu\nu AB}, \]  

(3.24)

\[ \hat{R} = 0 \iff R + 2\nabla_\nu K^{A\mu\nu} + K^{\mu\nu AB} - K^{\mu\nu AB} K^{A\mu\nu} = 0. \]  

(3.25)

Consequently, we have an alternative formulation of the Einstein-Hilbert action (3.8) using the previous identity, namely

\[ S_{EH} = \frac{1}{2\kappa^2} \int R\sqrt{-g} \, d^4x \iff S_{TEGR} = \frac{1}{2\kappa^2} \int \left\{ K^{\mu\nu \lambda} K^{\mu\nu \lambda} - K^{\mu\nu \lambda} K^{\mu\nu \lambda} \right\} \epsilon \, d^4x, \]  

(3.26)

where we neglected the boundary term that does not contribute to the equations of motion. Here, \( \epsilon \) denotes the determinant of the tetrad field, which satisfies \( \epsilon = \sqrt{-g} \) due to (3.21). This is the standard formulation of the Teleparallel Equivalent of General Relativity where the tetrad \( e^\alpha_\mu \) is the independent dynamical variable and the spin connection vanishes identically [81].

The issue of matter couplings was mentioned earlier, and the Teleparallel Equivalent of General Relativity is a good case study when it comes to matter couplings, especially for spin 1/2 particles. The Lagrangian for a Dirac field contains the term \( \nabla \psi \), where, as before, \( \psi \) stands for the spinor field. Its covariant derivative depends explicitly on the connection, which is no longer a dynamical variable in the teleparallel formulation. This leads to issues regarding the coupling prescription of Dirac fields; see in particular the series of papers [82–86] and the references given therein. It was pointed out in [86] that this problem is generic and affects all Poincare gauge theories that admit a teleparallel formulation.

### 3.4 The Einstein-Hilbert Action – Nonlinear Extensions

All theories considered so far were based on an action linear in the curvature scalar as the key ingredient to formulate gravitational theories; however, we already noted in (3.26) that such actions are quadratic in the contortion tensor. From a theoretical point of view it is well motivated to consider more general theories, which depend on other scalars constructed out of the Riemann curvature tensor or the Ricci tensor. There is no reason to exclude terms like \( c_1 R^{\mu\nu} R_{\mu\nu} \), for example, in a gravitational action. Alternatively, one can consider theories where an arbitrary function of the Ricci scalar is considered. In the following we will focus on the latter approach.
The key idea of this scheme is to consider the action

\[ S_{f(R)} = \frac{1}{2\kappa^2} \int f(R) \sqrt{-g} \, d^4x, \tag{3.27} \]

\[ S_{\text{matter}} = \int L_{\text{matter}}(g, \phi, \nabla \psi) \sqrt{-g} \, d^4x, \tag{3.28} \]

\[ S_{\text{total}} = S_{f(R)} + S_{\text{matter}}, \tag{3.29} \]

where \( f(R) \) is a sufficiently regular function of the Ricci scalar. When choosing \( f(R) = R \), one recovers General Relativity, while the choice \( f(R) = R - 2\Lambda \) introduces the cosmological constant into the field equations. Such a model was studied in the context of cosmology by [87]; however, it was only after the observation of the accelerated expansion of the Universe that models of this type became more mainstream and were subsequently thoroughly studied [88,89]; for reviews on \( f(R) \) gravity the reader is referred to [90–92]. The basic idea underlying this approach is to view General Relativity as the lowest order theory. To see this, recall that Minkowski spacetime is the geometrical framework for Special Relativity that satisfies \( R_{\mu\nu\lambda\kappa} = 0 \), the space being globally flat. Let us consider a series expansion of \( f(R) \) around Minkowski spacetime, then

\[ f(R) = f(0) + f'(0)R + f''(0)R^2/2 + \ldots \]

so that a term linear in \( R \) emerges quite naturally.

However, if one wishes to modify General Relativity for cosmological applications in particular, this expansion might not be ideal. Over very large scales the curvature becomes small, which motivates modifications that contain inverse powers of the Ricci scalar; clearly such terms pose problems when considering Minkowski space. Other models contain nonlinear functions of total derivative terms, like the Gauss-Bonnet term, for example. The Gauss-Bonnet term is related to a topological number, the Euler characteristic of the manifold. However, when any nonlinear function of any topological quantity is added to the action, it will yield some non-trivial field equations. Of course, one can also introduce new couplings between the geometry and the matter, different from the minimal coupling. Theories of this type have also received substantial attention; see in particular [93] for a comprehensive reference of such models. Let us add a small sceptic’s remark: A function \( f \) contains uncountably many degrees of freedom, so it is perhaps not too surprising that various models are able to fit a variety of observational data.

Going back to the Teleparallel Equivalent of General Relativity, one could apply the same ideas to the action (3.26) and consider nonlinear models. The scalar that appears in the integrand of \( S_{\text{TEGR}} \) is often denoted by \( T \), so that (3.25) can be written in the convenient form \( R = -T + B \), where \( B \) stands for the boundary term. Consequently, one would consider the model

\[ S_{f(T)} = \frac{1}{2\kappa^2} \int f(T) \sqrt{-g} \, d^4x, \tag{3.30} \]

\[ S_{\text{matter}} = \int L_{\text{matter}}(g, \phi, \nabla \psi) \sqrt{-g} \, d^4x, \tag{3.31} \]

\[ S_{\text{total}} = S_{f(T)} + S_{\text{matter}}, \tag{3.32} \]

which was first suggested in [94] and also led to a surge of interest; for reviews see [15,95]. These models allow for cosmological solutions with accelerated expansion without the need to introduce dark energy. From a conceptual point of view, modified teleparallel theories of gravity are interesting, as these are no longer invariant under local Lorentz transformation in their standard formulation; this means the choice \( \omega^A_{\mu} B = 0 \), where the spin connection vanishes identically. To see this, one only has to note that neither \( T \) nor \( B \) are Lorentz scalars; the combination \( R = -T + B \) is the unique Lorentz scalar that can be constructed, implying that General Relativity and \( f(R) \) gravity are both locally Lorentz invariant, while any nonlinear theories based on combinations of \( T \) and \( B \) are not; see [96]. We note that a fully invariant
formulation of modified teleparallel gravity models has been proposed [97]. It is very interesting to study the degrees of freedom in $f(T)$ gravity in four dimensions; see [98–100]. While in $f(R)$ gravity the extra degree of freedom is easily interpreted as a scalar due to the function $f$, an analogue interpretation in $f(T)$ cannot be made and the precise meaning of the extra degrees of freedom is an open question in the field.

Breaking local Lorentz invariance can be well motivated by considering physics on very small scales. Quantum theory implies that positions and momenta cannot be measured simultaneously with unlimited accuracy. Consequently, one would expect a certain length scale at which local Lorentz transformations break down.

Let us finish this section with another sceptic’s remark: Once one starts to consider nonlinear theories based on various scalar quantities, motivated either by the geometry or the matter content, one is able to create a plethora of theories. The entirety of such models is so large that it is (practically) impossible to study all of them. Clearly, many of these models can be built to pass a variety of observational tests due to their generality. What appears to be missing at the moment is an overarching guiding principle, which would allow us to restrict our attention to a small class of models based on some neat theoretical argument.